Doctoral theses at NTNU, 2021:369

Ola Aarøen

Coalescence in oil-in-water emulsions characterized using optical tweezers and rare event simulations

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

Ola Aarøen

Coalescence in oil-in-water emulsions characterized using optical tweezers and rare event simulations

Thesis for the Degree of Philosophiae Doctor

Trondheim, November 2021

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Natural Sciences
Department of Biotechnology and Food Science

ISBN 978-82-326-5715-5 (printed ver.)
ISBN 978-82-326-6444-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2021:369

Printed by NTNU Grafisk senter

Acknowledgements

Firstly, I would like to thank my supervisor, Professor Marit Sletmoen, for giving me the opportunity to work on this project. Thank you for your guidance on the experimental work, which I had little experience in, the door to your office being literally always open for discussions, and much help provided during the writing of papers and thesis.

I would also like to thank my co-supervisor Dr. Enrico Riccardi, for his expertise in simulations and coalescence modelling. Additionally, I would like to thank him for suggesting this project to me, and for helping in the process of publishing my own work in my very first paper.

Furthermore, I need to thank both my supervisors so much for their commitment to me and this project, allocating their precious time to me even after accepting new positions, co-operating from afar and throughout an ongoing pandemic.

I would also like to thank co-authors for their contribution to the papers in this thesis, and to Swapnil Vilas Bhujbal a thanks for helping me with the microfluidic set-up and for supplying the custom chips. A thank you also goes to Michael Lenuweit, for helping whenever the tweezer instrument acted up and I needed tech support.

Finally I have to thank my family, your encouragement and support helped me greatly during this process, especially this last year where both distance and distancing were felt. To my father: thank you yet again for agreeing to proofread my thesis. I hope that my spelling has improved since last time, both as a personal accomplishment and to ease your work. To my mother: thank you for the long conversations, discussions about work and everything else. Thank you for listening the times I simply needed to talk it out and clear my head. To my grandmother: thank you for our long-distance coffee phone calls, I would look forward to the weekend for this alone. Lastly, to my sister: you came to be the most amazing flatmate, and I will always look back at our time sharing the same roof fondly.

Preface

This thesis is submitted as a partial fullfilment of the requirements for the academic title *Philosophiae Doctor* at the Norwegian University of Science and Technology (NTNU).

The work of this thesis was carried out at the Department of Biotechnology and Food Science under the supervision of Professor Marit Sletmoen, and co-supervised by dr. Enrico Riccardi.

The thesis consists of a general overview subject matter and the theory behind it, an introduction to the different topics and fields contained within the work presented, the scope of the project, a summary and discussion of the results presented in the appended papers, and lastly concluding remarks.

Summary

Emulsions are thermodynamically unstable systems comprised of two immiscible liquid phases, with one dispersed into the other as small droplets of varying size. Coalescence is a separation mechanism for emulsion systems, in which two droplets make contact, and merge into one. It is critical to the emulsion stability and transport processes involving emulsions. Previous experiments have shown that a main parameter that controls the dynamics of the coalescence event is interactions between interfacial layers surrounding the approaching droplets, in the liquid thin film that establishes between them. Additionally, the current state-of-the-art continuous models describe the interface region very crudely, without concern for the interactions that lead up to and facilitate coalescence.

The interactions leading up to and facilitating coalescence between oil-in-water emulsion droplets have been studied, using a combination of optical tweezers and molecular dynamics. Both pristine (un-stabilized) and (ionic surfactant) stabilized emulsion droplets were used, in separate experiments.

Studies with the pristine emulsions were focused on measuring the effects of approach velocity, and droplet size on the work of separation and the time of coalescence. Using optical tweezers allowed for exploring an area of droplet diameters rarely visited using other experimental techniques. And insight on ionic effects to the production, handling and use of emulsions in optical traps was therefore also obtained during this work as well.

Stabilized emulsions were used to explore the interactions of droplets with both ionic and surfactant presence, and how changing these two parameters affected depletion force and coalescence times. Experiments were combined with molecular dynamics simulations to clarify the dynamics of the thin film breakage, using model systems to represent the different stages of droplet interface approach. And this combination provided crucial explanations to the unexpected observed behavior leading up to film breakage in the optical tweezers. Specifically, they revealed how the solvated ions interacted with the ionic headgroups to change the stability of the thin film between droplet interfaces, and how the electrical double layer formed on interfaces was disrupted in steps leading up to coalescence.

Although the MD simulation scheme was crude, the work provides an excellent transition to a proper rare event simulation, using software that have been improved within this work as well. Improvements to already well-established rare event simulation software PyRETIS were also made. To the simulation scheme, this is done by the improvement of sampling efficiency, interfacing with external dynamics, and refining the phase space exploration by implementing new shooting

moves and simulation initialization. Furthermore, the PyRETIS software gained a user-friendly post-processing and visualization package, PyVisA, answering the call for more accessible and customizable software to a wider user base. Functionalities of the PyRETIS–PyVisA combination include, but are not limited to, correlation of order parameters with other descriptors, identifying latent variables, interfacing with external molecular visualizer, and a variety of plotting tools that can aid in identifying possible meta-stable states. For a system as large as the ones used to model the thin film, these tools would unquestionably be of great value in analysing of the results, providing further insight into the effects of parameters to resulting film thickness and coalescence times.

Contents

List of Papers

Paper I

Riccardi, E., Lervik, A., Roet, S., Aarøen, O., and van Erp, T. S. (2020b). PyRETIS 2: An improbability drive for rare events. *Journal of Computational Chemistry*, 41(4):370–377

Paper II

Aarøen, O., Kiær, H., and Riccardi, E. (2020). Pyvisa: Visualization and analysis of path sampling trajectories. *Journal of Computational Chemistry*, 42(6):Pages 435–446

Paper III

Aarøen, O., Riccardi, E., and Sletmoen, M. (2021a). Exploring the effects of approach velocity on depletion force and coalescence in oil-in-water emulsions. *RSC Advances*, 11(15):8730–8740

Paper IV

Aarøen, O., Riccardi, E., Sletmoen, M., and van Erp, T. S. (2021b). Multidisciplinary study of effects on thin film breakage in oil–in–water emulsions. *Preprint*

Papers not included in thesis

Dahlen, O., Lervik, A., Aarøen, O., Cabriolu, R., Lyng, R., and van Erp, T. S. (2020). Teaching complex molecular simulation algorithms: Using self-evaluation to tailor web-based exercises at an individual level. *Computer Applications in Engineering Education*, 28(4):779–791

List of Symbols

Symbol	Description	[Constant values]	unit
	Colloids		
η	Viscosity of a liquid		kg/ms
ψ	Potential energy from interface charges		
R_S	Radius of spherical interface of particle		m
r	Radius from center of particle		m
κ	Debye-Hückel parameter		m^{-1}
e	Elementary charge,	1.6022×10^{-19}	C
N_A	Avogadros number,	6.022×10^{23}	mol^{-1}
ϵ	Specific permittivity of material		F/m
z_i	Valency of ion i		
M_i	Concentration of ion i		mol/L
$F(h)$	Force between colloidal particles		N
d_{eq}	Equilibrium diameter of two particles		m
h	Height of liquid thin film separating colloidal particles		m
t_d	Drainage time of thin film		s
h_c	Critical height of thin film		m
	Rare events		
k_{AB}	Rate of reaction $A \rightarrow B$		s^{-1}
$x_i(\vec{r}, \vec{p})$	System state positions $\vec{r}$ at a given time frame, i, with current momentum $\vec{p}$		
$\mathcal{X}(\mathcal{T})$	Trajectory of a system propagation, with length $\mathcal{T}$		
λ_i	Dividing surfaces in phase space separating states A and B		[OP]

$\mathcal{P}(\lambda_{i+1}\|\lambda_i)$	Conditional probability of trajectory crossing λ_{i+1} given λ_i has been crossed	
f_A	Flux out of stable state A	s^{-1}

Optical tweezers

ϵ_0	Permittivity of vacuum,	8.85418781×10^{-12}	F/m
ϵ	Specific permittivity of material		F/m
E	Electrical field vector		
v_f	Phase velocity of wave		m/s
λ	Wavelength of light		nm
f	Frequency		s^{-1}
$n(\lambda)$	Refractive index of material at wavelength λ		
θ_i	Angle of a beam i		
P_{total}	Radiation pressure from momentum change of incident light		N/m^2
I_f	Irradiance of light, power per unit of area		W/m^2
F_k	Spring force		N
k	Spring constant		N/m
F_s	Drag force on spherical particle in viscous medium		N
γ	Drag (friction) coefficient of spherical particle		kg/s
v	Velocity of particle		m/s
$S(f)$	Power spectrum density		V^2/Hz
D	Diffusion coefficient		m^2/s
T_i	Trap number (i = 1 or 2)		
F_d	Depletion force		pN
J	Work of separation		pN/μm
v_{app}	Approach velocity of traps		μm/s
d_{app}	Moving distance of traps during approach		μm

Chapter 1

Emulsions

An emulsion is a type of *colloidal system*, that is, a system consisting of two or more parts, where one or more of these are dispersed in the first as smaller particles, but not completely to form a solution, and systems where one liquid phase is dispersed in another liquid, are defined as *emulsions*. Two otherwise immiscible liquid phases are forced to mix, in such a way that the one of smallest volume is dispersed as small *droplets* in the larger, *bulk* liquid. Their immiscibility comes from the fact that one of theses phases usually consists of a polar medium, like water, while the other consists of a non-polar one, like hydrocarbons.

The use of emulsions is wide-spread in many different fields both academic and commercial. In food science, some common known emulsions include mayonnaise, milk and vinaigrette, and in medicine they have shown great promise as drug delivery systems(Chime and Onyishi 2013, Sengupta and Chatterjee 2017) and for applying skin pharmaceuticals(Khan et al. 2011). In the petroleum industry the formation of specific water–in–oil emulsions during crude oil extraction(Goodarzi and Zendehboudi 2019, Umar et al. 2018) is important for environmental and economical reasons.

Emulsions created as a desired product or for use in study are generated by the use of *homogenizers*. Homogenizers are a common term for instruments that impart a mechanical force upon the separated phases to break up and disperse one in the other. The most common types of homogenizers can be classified into three different types:

- Rotor generators or colloid mills, that apply a shearing force to the material by submerging a spinning shaft with angled blades attached into the sample preparation vessel.

- Sonic disruptors that generate ultrasonic waves in the range of 20–50 Hz, which when amplified into the sample are able to break or cut particles. As such they are mostly used on pre-homogenized emulsions to further split droplets into smaller ones.
- High-pressure, where sample liquid is forced through a narrow tube at very high pressure in order to break phases into droplets once they escape through the nozzle of the tube.
- Microfluidics, similar to high-pressure homogenizers, in that sample is produced by forcing liquids through narrow tubes. Except the scale of production is vastly reduced, in both volumes and pressure, making the system of tubes and nozzles small enough to fit on a lithographic chip design.

The type of homogenizer used depends on the target emulsion parameters, total volume, and overall dispersion droplet size and the distribution of sizes produced with the method. For instance, microfluidics have proven excellent at producing *monodispersed* (narrow) droplet sizes, but will only generate a tiny amount of emulsion over time, while a colloid mill may mix a much larger sample volume, but with product being much more *polydispersed*(wide range of sizes).

1.1 Stability

Although the two immiscible phases of an emulsion have been successfully mixed, this state is not stable, and over time will separate back to two pure states via different processes, depending on the overall stability of the emulsion in question. With regards to thermodynamics, an emulsion is unstable due to the tendency for a liquid–liquid system to separate in order to reduce interfacial area and thus interfacial energy. Emulsions do however resist separation and are considered stable over a period of time, i.e. they demonstrate *kinetic stability*, where the dispersed phase remains as small droplets dispersed in bulk up to and exceeding hours.

1.1.1 Separation mechanisms

Alas, over time the emulsion is subject to several mechanisms that will succeed in separating the phases. These mechanics can be divided into: (i) *sedimentation*, (ii) *aggregation* or *flocculation* (iii) *phase inversion*, and finally (iv) *coalescence*. Sedimentation occurs as the two liquid phases usually have very different densities, causing dispersed droplets to sink to the bottom of the bulk phase if heavier, or float to the top/surface if lighter (otherwise known as *creaming*). Aggregation or flocculation is a process in which dispersed droplets connect and form larger clusters upon contact, to further sediment or precipitate from bulk, and is often promoted by adding *clarifying agents* to the bulk solution. Phase inversion occurs

in very specific cases where the dispersed phase and bulk phases switch roles abruptly, so that dispersed phase now is bulk, and vice versa. Coalescence, while a separation mechanism alone, most often follows subsequent to the previous mechanicsm, where two droplets collide so that their interfaces come into contact, the liquid *thin film* of the bulk medium separating them breaks, and the droplets merge into one, reducing interfacial area.

1.1.2 Thermodynamic stabilization

Prolonging an emulsions stability and life-time (*shelf-life*) can be achieved by hindering the mechanics above, such as constantly agitating the emulsion or adjusting phase densities to prevent sedimentation, or carefully control its composition to avoid flocculation and phase inversion. Random collisions will still occur, and coalescence events will over time combine smaller droplets to larger ones, promoting all other separation mechanics. In order to prevent droplet coalescence, emulsions need to be *thermodynamically stable*. This may be achieved if the entropic gain of dispersing the one liquid phase outweighs the thermodynamically unfavorable formation of new liquid–liquid interfaces, for instance by having the dispersed phase as very small droplets, *micro-emulsions* (10–100 nm)(Windhab et al. 2005)), or by reducing the *interfacial tension* between phases.

1.2 Surface active agents

The interface layer between the two phases of an emulsion, presents a transition from a polar liquid to a non-polar one. At this transition, the relatively strong intermolecular forces in the polar liquid result in a net force normal to the interface, into the polar region. This force difference along the interface causes the droplets to form spherical shapes, that reduces the interfacial area, and the force necessary to deform the spherical interface is the interfacial tension.

With the difference in polarity between phases, some compounds have structures that allow them to gather in this region. Compounds with both polar and non-polar parts may effectively be dissolved by both types of liquids, and are thus *amphiphilic* molecules, both water-loving (*hydrophile*) and fat-loving (*lipophile*). Typical amphiphiles are single or multiple chained hydrocarbon "tails" with polar "head" functional groups, -OH, -COOH, etc., and such molecules are referred to as *surfactants*. By covering the liquid interface with a surfactant, the net force difference causing interfacial tension is reduced, thus stabilizing the emulsion thermodynamically. Furthermore, surfactant coverage on emulsion interfaces also stabilizes emulsions by introducing electrostatic and steric repulsion to the interfaces when droplets collide, hindering interfacial contact and subsequent film breakage. Surfactant effectiveness is dependant on the structure of the surfactant,

more specifically the form and relation of its head-group and tail(s), and is described by hydrophilic-lipophilic balance (HLB)(Pasquali et al. 2009), or Davies' method(Guo et al. 2006).

1.3 Electrostatic interactions

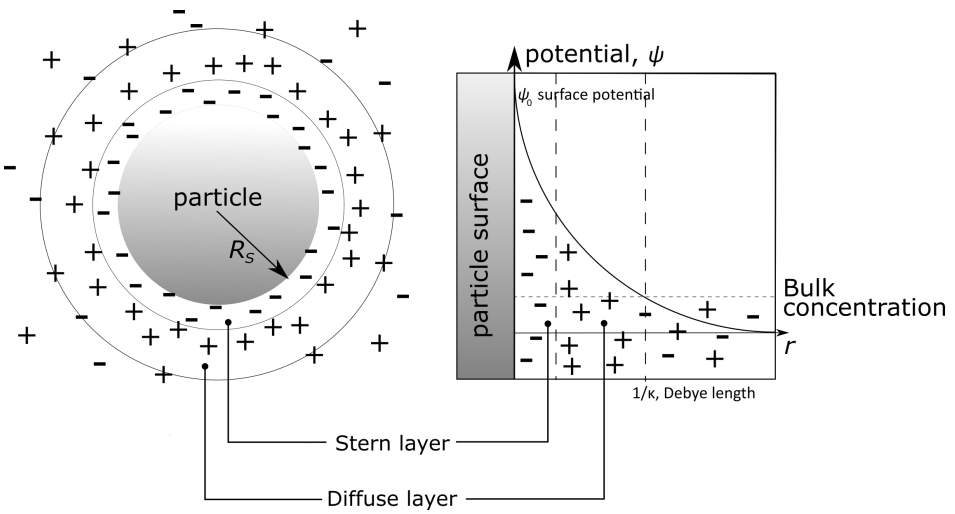


Figure 1.1: Illustration of the electrical double layer forming on a colloidal particle of radius R_S, dispersed in an ionic solution. The innermost Stern layer of negatively charge ions, and the diffuse outer layer with counterions, and the distance $1/\kappa$ where potential arising from charges become equal to the bulk concentration.

Once any surface exposed to a fluid, a structure called the *electrical double layer* (EDL) forms, an illustration of this layer is shown in figure 1.1. This layer consists of two parallel layers of opposite charges, and for oil–water systems this layer occurs at the interface between phases, on water-side, using either ions present in water phase or hydroxide–hydronium from autoprotolysis.

The first layer (*Stern* layer) will consist of ions adsorbed to the interface due to chemical interactions, and the second (*diffuse*) layer consists of oppositely charged ions attached to the first layer by Coloumb force, screening the first layer. Electrostatic forces from this EDL contribute to the stability of emulsions, by introducing equally charged layers around the emulsion droplets, producing a repulsive force that arise when these begin to overlap.

The potential of charges that produce the repulsive force can be described by the Poisson-Boltzmann equation, and when assuming low potentials, the Debye-

Hückel approximation(2016, p. 508–512):

$$\frac{d^2\psi}{dx^2} = \kappa^2\psi$$

which for spherical surfaces has the solution: (1.1)

$$\psi = \psi_0 \frac{R_S}{r} \times e^{-\kappa(r-R_S)}$$

where R_S is particle radius, r is distance from center of particle. κ is the Debye-Hückel Parameter, defined as

$$\kappa = \sqrt{\frac{1000\mathrm{e}^2 N_A}{\epsilon k_B T} \sum_i z_i^2 M_i} \tag{1.2}$$

where e is elementary charge, N_A is Avogadros constant, ϵ the permittivity of the liquid, z_i valency and M_i ionic concentration of ions i. This parameter can be used to give the extension — or thickness -– of the double layer, the Debye length: $\frac{1}{\kappa}$. While not some "true width" of the repulsive ion layer, the Debye length provides a theoretical distance from the interface at which the potential energy from the charges in the double layer becomes equal to those of the bulk solution, as seen in figure 1.1. In table 1.1 are some sample calculations of this distance for cases of pure water (hydroxide–hydronium ions) and NaCl solutions.

Table 1.1: Concentration of hydroxide–hydronium ions in pure water (pH 7) and the corresponding Debye length κ^{-1}, calculated with eq. (1.2).

Ions	Concentration [mol/L]	κ^{-1} [nm]
H^+ OH^-	10^{-7}	960.28
Na^+ Cl^-	0.1	0.96
	1.0	0.30

From this table it is possible to see that with increasing ionic strength, the EDL can be reduced – theoretically – to the scale of individual atoms.

1.3.1 DLVO theory

Combining the long-range repulsive effects of the EDL with short-range attractive Van der Waals forces is the idea of DLVO theory for colloidal dispersion stability, outlined by Lyklema et al. (1999). Here the force between to colloidal particles can be defined by:

$$F(h) = \pi d_{eq} E(h) \tag{1.3}$$

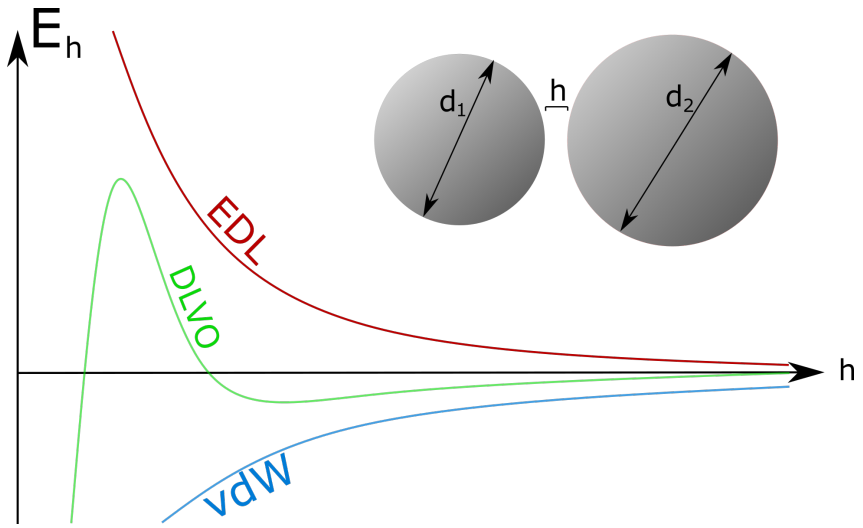


Figure 1.2: DVLO energy and components of EDL and vdW, as function of distance, h (height), between the interfaces of two colloidal particles of diameters d_1 and d_2

with force acting along the separation distance between interfaces, or *height* of the thin film, h, and the equilibrium particle size given as

$$d_{eq} = \frac{2d_1 d_2}{d_1 + d_2} \tag{1.4}$$

The energy term $E(h)$ is further split into double layer (*dl*) and Van der Waals (*vdW*) components

$$E(h) = E_{vdW}(h) + E_{dl}(h) \tag{1.5}$$

In rough terms, the theory describes the total potential energy as two particles or interfaces approach each other, giving first an energy barrier where repulsion exceeds attractive forces, and second an energy well at very close distance where attractive forces dominate. This energy, and the components of EDL and vdW, are illustrated in figure 1.2.

Thus, in order for interfaces to make contact, the particles, or droplets, need sufficient kinetic or thermal energy to pass the energy barrier, or the repulsive potential part must be reduced by eq. (1.2). And this makes the theory well suited to approximate aggregation of colloidal particles in suspension. Overall the theory outlines an important bedrock to model the electrostatic interactions in colloidal suspensions.

A drawback of this theory, however, arise from its dependency on the assumptions made in the theory of the EDL, and the Debye-Hückel approximation, as it only

allows for low charge potentials, meaning only dilute ionic solutions. Furthermore, individual charges that make up the Stern layer, and to an extent the diffuse layer (see figure 1.1), are considered static at the colloidal interface, as opposed to having proper freedom of movement.

1.3.2 Depletion force

A different effect of ionic (or any other particles relatively smaller than the colloids in question) presence in the bulk solution is the *depletion force*. This force is entropic, and arises from a concentration difference between the bulk solution and the thin liquid film. When two colloidal particles in a solution approach, the bulk liquid between the droplet interfaces drains away to a thin film, thus creating a local depletion of whichever ions or molecules or sub-micro particulate matter within this volume. This will in turn create an osmotic pressure towards the thin film, effectively pushing the colloidal particles together with a certain force. This force has been shown to be crucial to the overall colloidal stability(Mao et al. 1995; 1997). Additionally, surfactants present at the interfaces could in theory also contribute to the overall depletion felt by colloidal particles, whilst still stabilizing the interfaces (Ji and Walz 2015).

Comparatively to DLVO, the force of depletion between particles occurs at a longer separation distance than the film drainage and rupture explained in the following section, and is therefore not relevant to models describing coalescence by the film drainage alone. It is, however, important in the overall process of coalescence, as an attractive force keeping droplets stuck to one another. By keeping interfaces at sub-micrometer separation, avoiding droplet bouncing on collision or other separation, it is facilitating the later film drainage and rupture.

1.4 Film drainage modelling

With the basic understanding of the colloidal interactions between droplets from DLVO theory, one would expect that a reduction of the repulsive forces of the EDL should increase the overall probability for spontaneous coalescence when droplets are brought to close proximity. This de-stabilization of emulsions would be of interest especially within petroleum industries, where the production of *wastewater* emulsions containing sea-water and crude oil would be of concern, both for economic and environmental intentions(Kundu and Mishra 2018, Pan et al. 2018, Vegas Mendoza et al. 2019). Techniques to control this stability by the use of surfactants(Kamal 2016, Negin et al. 2017), polymers or particles(Sun et al. 2017) to promote aggregation and coalescence are classified as (green) enhanced oil recovery ((G)EOR). And modelling the effect of such additives to the emulsion stability and droplet behavior similarly becomes of great importance.

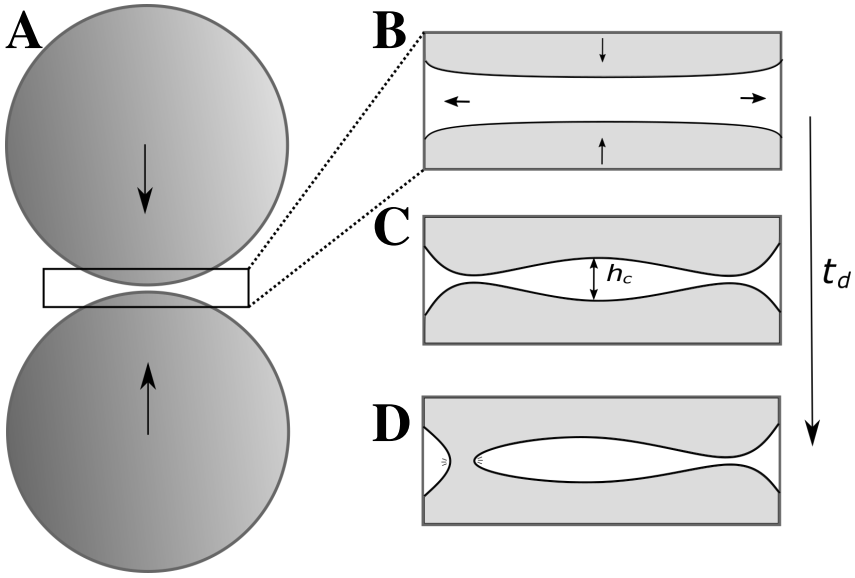


Figure 1.3: Illustration of the film drainage between two droplets approaching **A**, forming the thin liquid film between them. The insert **B** is zoomed in on the droplet interfaces and thin film, which over time t_d begins to drain, and deform **C** at a critical height h_c, before the interfaces make contact **D**, and the film is ruptured and the droplets coalesce.

Despite many years of investigations, resulting in vast experimental studies and models proposed to describe them, coalescence in emulsions and foams is still far from being completely understood(Langevin 2019). The coalescence event itself can be characterized as very rapid, which negatively influences the accuracy of experimental studies observing the event progression. Furthermore, thin film rupture depends on several properties of the emulsion system, including hydrodynamics, surface rheology, surface forces, and thermal fluctuations, all of which require extensive knowledge to successfully being implemented in theoretical models.

In contrast to the experimental studies of film drainage when including the effect of surfactants, theoretical models primarily employ simple descriptors for their behavior, basing the model systems on London dispersion forces in for W/O emulsions and DLVO theory for O/W emulsions.

Klaseboer et al. (2000) proposed a model for the thin film breakage in droplets and bubbles for constant approach velocity. The model predicts specific steps from initial collision and leading up to the coalescence event. On initial contact, a small deformation of the interface occurs, establishing a flat liquid thin film separating the interfaces. Later, after a given time t_d the thin film drains to a critical height of the thin film h_c, and the interfaces further deform, causing a "dimpling" of the interfaces which cause contact between the two droplets, making the thin film

subsequently break, coalescing them. A brief illustration of this process is shown in figure 1.3, where two droplets on approach form, drain, and subsequently break, the liquid thin film as the lead up to the coalescence event itself.

Ozan and Jakobsen have made improvements to the model idea of Klaseboer(2000), by attempting to estimate the effect of approach velocity, surface viscosity and viscoelasticity to the film drainage and thus coalescence time (2019a, 2019b, 2020). Their findings suggest clear dependencies on droplet approach velocity, even a theoretical lower bound for the drainage time and a upper limit to approach velocity, resulting in droplet "bounce" rather than film breakage.

1.4.1 Current state

In order to describe the interface region under different conditions of surfactant and ionic presence, the system has to be described at molecular resolution. This region between the two distinct phases is very sensitive to system modification, which must be accounted for. Solvent molecules determine the behavior of the thin film by interacting both with surfactants on interface and the interface region surrounding emulsion droplets itself. Structural rearrangements in the thin film occur as a result of altered conditions to this region, by changing surfactant concentrations and type, or ionic strength and type. And the importance of these structural and conformational rearrangements have been investigated at molecular level(Jang and Goddard 2006, Rekvig and Frenkel 2007). Furthermore, studies to provide comprehensive understanding by unifying macroscopic behavior with these underlying mechanisms on molecular scale have been performed(Kovalchuk et al. 2014, Riccardi and Tichelkamp 2019), but remain lacking.

Thus, at contiuous resolution, current models lack proper description of the film rupture mechanics to predict the overall system behavior of emulsion systems and coalescence. Furthermore, experimental techniques to fill the gap are being limited by the required control of the critical distance between droplet interface regions, and measurement of the inter-droplet forces to permit a validation of the models for film drainage. As these models also consider interfaces already at a separation distance sufficiently low for the sequence of events to occur, they also disregard any interactions prior to the thin film formation, which on an experimental scale will influence the perceived coalescence time.

Chapter 2

Rare event simulation

Conventional molecular dynamics (MD) gives its user the ability to compute many structural and conformational properties of the studied system. Unfortunately, there are several dynamical phenomena that are impossible to be studied this way. In general, conventional MD can not be applied when studying an activated process, or rather, a dynamical change which is considered to be a *rare event*. In chemical reaction simulations, a rare event is usually considered the crossing of a high energy barrier between stable states. By only using conventional MD for such a crossing, the probability for the event to happen at all would be extremely low. Reactions like protein folding, formation of clusters and diffusion of gas molecules in solids are basically impossible to model with a brute-force MD. This is due to its limitations regarding the barrier that needs be overcome, and probable bottlenecks in the potential energy surface in phase space that the system in addition must happen upon to cross.

However, in the extremely unlikely case that the rare event actually occurs during a simulation, it will in general progress quickly enough – be short enough – that the reaction itself could be simulated within MD time scale; in the order of ns to ps. Therefore, the optimal solution would be to have a method of avoiding any needless wandering around in either stable state around the energy barrier, and somehow enhance the probability of getting a rare event transition.

2.1 Transition state

One way of imagining how such rare event reactions might happen is called *Transition State Theory* (TST)(Wigner 1938). TST states that for a given, reversible reaction between two stable states A and B, the reactants in A must have a sufficient

kinetic energy in order to overcome the potential energy barrier that separates them from product state B.

$$A \rightleftharpoons B \tag{2.1}$$

Whatever height of this barrier, it is the *activation energy* for the transition, and the theory imagines an unstable *Transition State*(TS) of the system at the very peak of the barrier. TST is illustrated for the stable states A and B in figure 2.1.

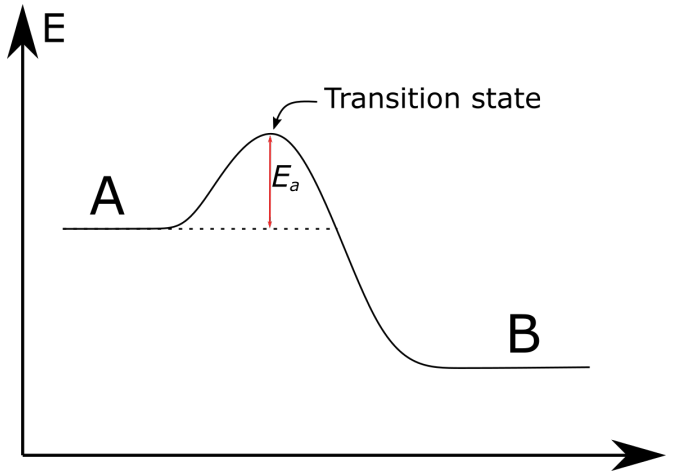


Figure 2.1: Illustration of the transition from state A to state B, with a energy barrier E_a separating the two and the transition state indicated at the top of this barrier.

The rate of reaction for a the transition in eq. (2.1), k_{AB}, can be defined by Arrhenius' law:

$$k_{AB} = A \cdot e^{-E_a/RT} \tag{2.2}$$

Here E_a is the activation energy, R is the universal gas constant, T temperature in Kelvin, A is a frequency factor. The constant k_{AB} has units s^{-1} and represents the number of reactions, or transitions, per unit of time. However while this rather simple case of transitioning A–to–B models well with TST, the theory has its limitations when applied to more complex systems. Now the transition state might no longer be as a singular point on a 1-Dimensional potential energy surface, but rather a winding path on a multidimensional surface in the free energy landscape, with a trajectory of transitioning that may cross the imagined peak several times back and forth before settling in either stable states. This is a problem as TST also assumes the trajectory across the energy landscape to be relatively short, with only a single crossing through the TS from A to B along a quite direct and straight path. The free energy barrier of TST is also limited by the degrees of freedom used to describe the free energy surface (Van Erp 2012).

2.1.1 Path sampling

So rather than focusing on the transition as the crossing of a single point or plane in a multi-dimensional landscape, it makes more sense to think of the transition as a trajectory, or path, of least resistance between states A and B. In doing so, there is no need for any information about the system outside of the two states, and this is achievable by defining a *reaction coordinate* (RC) or *order parameter* (OP). This coordinate or parameter can at any time be used to determine whether the system is currently residing within the areas that now define either stable states. For its method of finding multiple pathways through free energy landscapes, this simulation method is called *Transition Path Sampling* (TPS) (Bolhuis et al. 2002). By using a Monte Carlo approach to generate pathways, TPS is further not limited to single point or surface crossings (Dellago et al. 2002). And when connecting the two states by a trajectory or path, TPS provides the opportunity of exploring all likely transition pathways. Trajectories or paths are defined as sequences of system states, like an ordered collection of still frames of the moving system, taken along the path of the RC/OP and time. For a given path of length $\mathcal{T}$, each frame x_i is separated by its neighbours in time by a small increment Δt

$$\mathcal{X}(\mathcal{T}) \equiv \{x_0, x_{\Delta t}, x_{2\Delta t}, ..., x_{\mathcal{T}}\} \tag{2.3}$$

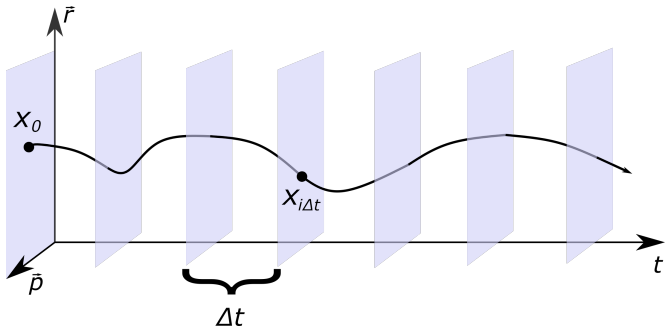


Figure 2.2: Illustration of a trajectory $\mathcal{X}$ over time in the space of particle positions $\vec{r}$ and momentum $\vec{p}$, containing snapshots at intervals of Δt.

In a molecular system evolving with discrete equations of motion, $x_{i\Delta t}(\vec{r}, \vec{p})$ is the full representation of the system state at time i, with positions $\vec{r}$ and momentum $\vec{p}$ for each and every particle. As modelling the rare transition event of the system is the simulations goal, reactive paths are the main focus, and are usually grouped in a *path ensemble*. Reactive is here defined as those paths that successfully traverse between states A and B, or specifically, paths that have been started in the region

of A, and end in the region of B.

$$h_{A,B}(x) = \begin{cases} 1 & \text{if } x \in A, B \\ 0 & \text{if } x \notin A, B \end{cases} \tag{2.4}$$

Thus, any path not starting within A and ending within B has zero statistical weight on the result.

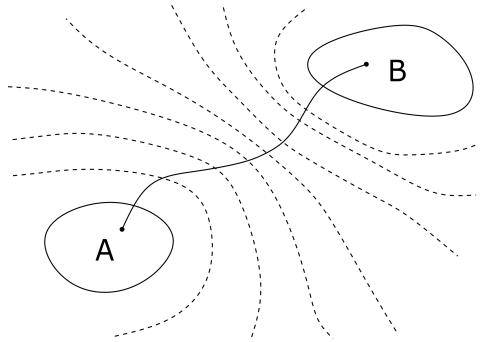


Figure 2.3: Illustration of a reactive trajectory (or transition path) connecting the stable states A and B through some undefined energy landscape.

Initializing a TPS simulation involves generating a primary pathway that, as required, connects reactant and product state. Such a path could be generated running a very long MD simulation, however such a rare event occurring was previously mentioned as practically impossible within a reasonable time scale. Increasing system temperature, and therefore kinetic energy, could result in a somewhat higher probability transition, but the resulting trajectory would be very different from one at proper energy levels. Additionally, a systematic cooling of the system by small steps would be necessary in order to bring it down to normal conditions. A proper way of achieving transition would be the re-implementation of a RC or OP to define regions A and B and the space between them. From there, the system can be driven through transition to obtain an initial path.

2.2 Transition Interface Sampling

Transition interface sampling (van Erp et al. 2003) (TIS) is a sampling method based on the principles of TPS. Unlike TPS, however, TIS measures the effective positive flux of the reaction, restricting the *reactive paths* to only positive transititons, paths beginning in A and ending in B, for the rate calculation. Additionally, several dividing surfaces, or *interfaces* λ_i, between A and B are introduced, as illustrated in figure 2.4. And rather than the single point or surface approximating the transition state as in TST, the flux is now defined by the crossing of said in-

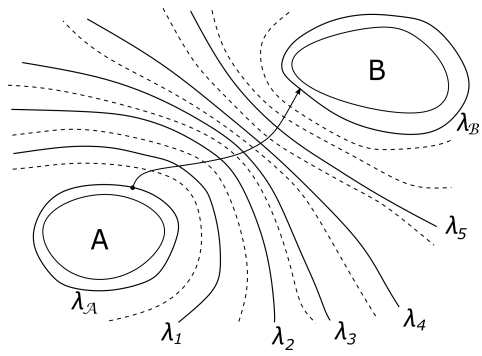


Figure 2.4: Illustration of reactive path, with interfaces λ_i inserted along reaction pathway. Path shortened to just connect interfaces λ_A and λ_B, which are positioned just outside of what has been regarded as the stable states of A and B.

terfaces. The interfaces are separated in phase space by re-introducting the order parameter, and ordered like $\lambda_1, \lambda_2, ...\lambda_n$.

To formulate the flux from one state to another properly, the phase space is re-defined to make it so that the system is never considered to be in an intermediate region, but is directly related to either A or B. By now introducing the overall states $\mathcal{A}$ and $\mathcal{B}$, so that in addition to including phase space points within A and B as originally intended, state $\mathcal{A}$ also covers states that have visited A before reaching B when integrating backwards in time.

2.2.1 Rate calculation

Using the newly defined $\mathcal{A}$ and $\mathcal{B}$ overall states, the expression of the TIS rate constant is defined as the derivative of a time correlation function related to the transition from A to B,

$$k_{AB} = \frac{d}{dt} C(t)\big|_{t=0}, \quad C(t) = \frac{\langle h_{\mathcal{A}}(x_0) h_{\mathcal{B}}(x_t)\rangle}{\langle h_{\mathcal{A}}\rangle}$$

$$k_{AB} = \frac{\langle h_{\mathcal{A}}(x_0)\dot{h}_{\mathcal{B}}(x_0)\rangle}{\langle h_{\mathcal{A}}\rangle} = \frac{\langle \Phi_{A,\lambda_B}\rangle}{\langle h_{\mathcal{A}}\rangle} \tag{2.5}$$

where $\langle \Phi_{A,\lambda_B}\rangle$ is positive flux out of state A and through interface λ_B. The effective flux through any given interface λ_i is $\langle \Phi_{A,\lambda_i}\rangle$, and is directly related to the flux through interface before it, λ_{i-1}, by

$$\langle \Phi_{A,\lambda_i}(x_0)\rangle = \mathcal{P}(\lambda_i|\lambda_{i-1}) \cdot \langle \Phi_{A,\lambda_{i-1}}(x_0)\rangle \tag{2.6}$$

where $\langle ... \rangle_{\Phi_{A,\lambda_{i-1}}}$ is the ensemble average over all points in phase space for x_0 for which $\Phi_{A,\lambda_{i-1}}(x_0) \neq 0$ and $\mathcal{P}(\lambda_i|\lambda_{i-1})$ is the conditional probability of a trajectory crossing λ_i, given it already has passed λ_{i-1} earlier in time. Applying this condition on all interfaces, the rate constant expression of eq. (2.5) is reduced to:

$$k_{AB} = \frac{\langle \phi_{A,\lambda_1} \rangle}{\langle h_{\mathcal{A}} \rangle} \prod_{i=1}^{n-1} \mathcal{P}(\lambda_{i+1}|\lambda_i) = \frac{\langle \phi_{A,\lambda_1} \rangle}{\langle h_{\mathcal{A}} \rangle} \mathcal{P}(\lambda_B|\lambda_1). \tag{2.7}$$

This new expression for the rate constant relies almost exclusively on calculation of the continuous conditional probabilities $\mathcal{P}(\lambda_{i+1}|\lambda_i)$ of crossing from from state A to B. The positive effective flux term from eq. (2.5) has simply been reduced to the positive flux over the first interface $\langle \phi_{A,\lambda_1} \rangle$. And now this property can be calculated by counting the number of effective crossings over λ_A, or more commonly λ_0. With a proper choice of order parameter and location of interfaces, this calculation is possible with a simple brute-force MD. Further, the power of the TIS scheme is that the probability of a total crossing event $\mathcal{P}(\lambda_B|\lambda_A)$, which as discussed can be extremely low, is separated into smaller and more manageable probabilities $\mathcal{P}(\lambda_{i+1}|\lambda_i)$, that individually have a greater probability of occurring.

2.2.2 The TIS algorithm

A complete TIS scheme goes as follows: An MD simulation is started from the initial system configuration in order to generate a trajectory within stable state A that, that at some point, crosses the first interface λ_0 out of A then to later return. This generates an initial path of length $\mathcal{T}$ that start in A, proceeds to cross λ_0 and return to A or continues to cross interfaces before ending in either A or B. From this initial path the point x_0 is defined as the first point in phase space after crossing λ_0, and $x_{\mathcal{T}}$ the last point before crossing into a stable state. This new trajectory is now defined to belong to the $[0^+]$ ensemble. More trajectories can be generated from this initial pathway via a process known as a *shooting move*, which is a Monte Carlo move on the system configuration. This move is usually performed by perturbing the momentum of every particle in the system at a time t, to generate a new system state $x_t^{(new)}$ and then progressing the system forwards and backwards in time in order to see which interfaces the system crosses, creating a new trajectory $\mathcal{X}^n(\mathcal{T}^*)$.

$$\begin{aligned} x_t^{(new)}(\vec{r},\vec{p}) =& x_t^n = x_t^{(old)}(\vec{r},\vec{p}+\vec{\Delta p}) \\ & \downarrow \\ \{x_0^n, x_{\Delta t}^n, ..., & x_t^n, ..., x_{\mathcal{T}^*}^n\} = \mathcal{X}^n(\mathcal{T}^*) \end{aligned} \tag{2.8}$$

The integration of equations of motion is continued until the backwards and forwards progression of the trajectory reaches and crosses into a stable state, creating

a valid path out of state A and either reaching B or returning to A. Should either the forwards or backwards progressions not reach a stable state within a certain time, however, the integration has to be terminated and the trajectory rejected not to waste unnecessary time. Rejection of trajectories occurs if the new pathway generated either:

- does not reach the edge of A backwards within a user defined time, $\mathcal{T}^{max}$
- does not reach the edge of either A or B forwards within $\mathcal{T}^{max}$ steps
- does not pass a probability check: $\min\left[1, \frac{\mathcal{T}^{(o)}}{\mathcal{T}^{(n)}}\right]$, where $\mathcal{T}^{(o)}$ is the length of the old path and $\mathcal{T}^{(n)}$ is the length of the new one.

Seeing how deterministic dynamics are defined to propagate, and by scaling the perturbation of the random momentum to the shooting point correctly to system state, the probability to accept any new generated path can be written as:

$$P_{acc}[x^{(o)}(\mathcal{T}) \to x^{(n)}(\mathcal{T})] = h_A(x_0^{(n)})h_B(x_{\mathcal{T}}^{(n)}) \min\left[1, \frac{\rho(x_{t'}^{(n)})}{\rho(x_{t'}^{(o)})}\right] \tag{2.9}$$

implying that any new path connecting region A and B is accepted with a probability that depends only on the shooting points of the old and new paths(Dellago et al. 2002).

Accepted paths are collected in the *path ensemble*, and the new path takes the place as the current accepted in its ensemble. And from it, the process repeats in order to generate more valid paths. In sampling techniques where Monte Carlo moves are used in this way, any rejection of the trial path implies that the old path should be counted again in the path ensemble and the process is restarted by choosing another shooting point (van Erp et al. 2003).

Accepted paths are grouped in the path ensemble associated with the interface which they attempt crossing, so that trajectories that cross λ_i belong to the $[i^+]$ path ensemble. Any of these paths that proceed further to also cross λ_{i+1} are still kept in $[i^+]$, and they give the relative probability of any accepted path in $[i^+]$ to cross the next interface, crucial for the conditional probabilities in the rate calculation of eq. (2.7.) To achieve the highest efficiency of rate calculations, it has been shown that interfaces should be positioned so that the general crossing probability of each is as close as possible to 0.2 (van Erp 2006).

2.2.3 Developments to RETIS

Over the years, advancements to transition interface sampling has provided several Monte Carlo moves to path generation, and in particular, two computationally

cheap but very rewarding moves added were the *time reversal* and *path swapping* moves. The time reversal move was initially defined in the TPS scheme(Bolhuis et al. 2002), and works simply by taking the current accepted path, and reversing the order of time frames of the system and all particle momenta.

$$\begin{aligned} x^n_{\mathcal{T}}(\vec{r},\vec{p}) &= x^o_0(\vec{r},-\vec{p}) \\ \mathcal{X}^n(\mathcal{T}) &= \{x^o_{\mathcal{T}}(\vec{r},-\vec{p}),...,x^o_0(\vec{r},-\vec{p})\} \\ &= \{x^n_0,...,x^n_{\mathcal{T}}\} \end{aligned} \tag{2.10}$$

Time reversal almost always gives an accepted path, as it simply is a path already accepted in the ensemble, so it fulfills the criteria of acceptance. The only exception to this statement is in the case of a time reversal on a reactive pathway connecting A and B, as the reversed pathway would begin in B, which implies an immediate fail.

The swapping move, or replica exchange (RETIS), works quite differently from the first two, by attempting to swap the current paths of two adjacent path ensembles(van Erp 2007), over all path ensembles in the simulation, potentially providing new paths to several ensembles in one step. As described earlier, accepted paths in ensemble $[i^+]$ must cross λ_i, with hopefully a 20% chance of crossing λ_{i+1}. If the current path in $[i^+]$ does, it is also eligible to be accepted in $[i{+}1^+]$, and of course *any* path in $[i{+}1^+]$ is eligible for $[i^+]$, so a swap of these two paths is accepted, providing an additional accepted path to each ensemble cheaply. The RETIS scheme additionally substitutes the MD simulation of TIS with a path ensemble $[0^-]$, where from a point x_0 just past λ_0 they cross *into* the region of A and away from λ_0, before crossing on their way *out* of A. This allows for some flexibility to also swap paths between $[0^-]$ and $[0^+]$ ensembles, but with the added step of integrating the dynamics of motion for both paths, using x_0 as a starting point going forwards in time. Furthermore, this ensemble substitutes for the term for the flux calculation of eq. (2.6) with

$$f_A = \left(\left\langle \mathcal{T}^{[0^-]}_{\text{path}} \right\rangle + \left\langle \mathcal{T}^{[0^+]}_{\text{path}} \right\rangle\right)^{-1} \tag{2.11}$$

where $\langle \mathcal{T}^{[0^{+/-}]}_{\text{path}} \rangle$ is the average length of paths in ensemble$[0^{+/-}]$.

Standard procedure for the choice of moves on the path ensemble is usually performed with a random number generation to determine whether to swap on the entire path ensemble. If not swapping, an additional check is performed to do either a shooting or a time reversal on every ensemble individually. If a swapping is chosen however, a second check would determine whether to begin the neighboring swaps at ensemble $[0^-]$ or $[0^+]$ and upwards. In the case where $[0^-]$ is

omitted from the swap, or possibly the final ensemble – as one does not always have a even number of ensembles – a *null move* is performed, which returns the already accepted path, and recounts it for statistical weight. While the re-use of already existing paths might not intuitively increase the sampling efficiency, studies show that the sharing of paths between ensembles greatly increases the probing of the energy landscape in phase space(van Erp 2008), and greatly increases the convergence of the rate calculation(van Erp 2007).

2.3 Current developments to sampling techniques

In recent years the development of sampling techniques have shifted towards widening the overall user base. Previously, in-house software was usually constructed in the preferred computing language of the small group of people that developed and used the software almost exclusively. With the power and flexibility of sampling techniques to fit a practically infinite set of systems and reactions, one wishes to a greater extent to reach scientists within other fields.

The base program for the defined TIS/RETIS schemes in previous sections was an in-house software, written in Fortran90, and was initially developed without using any sort of version control system. These choices result in a very cumbersome development process. The lack of version control, i.e. Git/GitHub, limits the amount of parallel development available to the users. And while a compiled imperative language like Fortan90 works well for numerical and scientific computing, it is now severely outdated, lacking any sort of object-oriented features or parallel computing - restricting its use to mainly act as an inclusion in order to perform numerical calculations.

Thus, the RETIS algorithm was re-made in a Python library, PyRETIS(Lervik et al. 2017a). Similarly, other packages to perform path sampling techniques have been contructed, also in Python: WESTPA(Zwier et al. 2015) (Weighted Ensemble Simulation toolkit with Parallelization and Analysis) and OPS(Swenson et al. 2018) (OpenPathSampling). This transition to Python allowed an object-oriented restructuring of the program, streamlining the simulation scheme in an overall much more user-friendly syntax, which in addition to the version control opened up development for more users not accustomed to the old software. Although Python as a scripting language requires translation to machine code for each run, increasing its computing time for simple tasks when compared to a compiled language, it makes up for this by including external modules. Such pre-compiled "imports" like NumPy(Harris et al. 2020), handle the computationally intensive tasks, reducing computational time greatly, and freeing the user from the need of creating a similar module in an often improper fashion.

In line with this, PyRETIS also makes use of external modules to handle the intensive calculation of system dynamics to generate trajectories, while keeping path ensembles, statistics and analysis as class objects in the Python structure. This opens up for interfacing the path/interface sampling algorithm with established software packages like CP2K(Kühne et al. 2020) and GROMACS(Abraham et al. 2015) that may model the system and its dynamics with more sophisticated methods.

Most recently, the improvements to RETIS/PyRETIS(Riccardi et al. 2020b) involve introducing new Monte Carlo methods (Riccardi et al. 2017) to improve sampling efficiency, and integrating machine learning into the algorithm to identify reactive pathways(Roet et al. 2021). With the RETIS approach now being made more open to non-developers (2019), more user-friendly tools for analysing and understanding the results of sampling techniques are required, which will be discussed in a later chapter.

Chapter 3

Optical Tweezers

Optical tweezers is a scientific name for an instrument performing *micromanipulation* using a focused laser to grab onto and control microscopic objects, similar to regular tweezers. Other similar techniques include microneedles, atomic force microscopes (AFM) and magnetic tweezers, and they all excel in force microscopy, the measurement of interaction forces between individual molecules. The physics of optical tweezers and their usage in science will be expanded upon in this chapter.

3.1 Nature of light

To understand how optical tweezers work, it is necessary to first briefly discuss the properties of light. In nature, some particles have a charge property, allowing them to interact via *electromagnetic waves*. Light is just a description that such interactions between charged particles is referred to as. An electromagnetic wave consist of an two coupled oscillating fields, one electric and one magnetic, perpendicular to one another. The direction of the electric field is referred to as *polarization*, and most sources of light produce waves with equal mixture of polarization, lasers being a notable exception.

3.1.1 Light–Matter interactions

The way light interacts with matter, results in some key behaviors. When light travels through a material, it may interact with its electrical field caused by electrons in atoms in the material. With this, light has been observed to slow down and "bend" when passing through different transparent mediums. When light first hits the surface of a transparent medium, electrons in the medium are oscillated by the electromagnetic wave of light. This oscillation generate a new electromag-

netic wave inside the medium, which combines with the light wave to form a third wave, moving slower through the medium than the incident light was travelling previously, outside of it. To explain the bending of the light, two of Maxwell's equations are needed.

$$\nabla \times E = -\frac{\delta B}{\delta t}$$
$$\epsilon \nabla \cdot E = \rho \tag{3.1}$$

Imagining the case of a single beam of light (or photon) traveling in air, hitting the surface of water and bending, illustrated in figure 3.1. Incident and bent beams travel in air and water respectively, both transparent mediums with different specific *permittivity*, ϵ, a measure of how polarizable the material is. A more polarizable material will have a stronger interaction between its electrical field and that of the light. In addition to the vector indicating the direction of light, both beams each have a perpendicular vector indicating the direction of the electrical field of the light wave, E. This vector can be defined as the vector sum of two components, perpendicular and parallel to the air-water surface.

$$E_{water/air} = E^{parallel}_{water/air} + E^{perpendicular}_{water/air} \tag{3.2}$$

Since the air-water surface belongs to both regions of air and water, at the surface

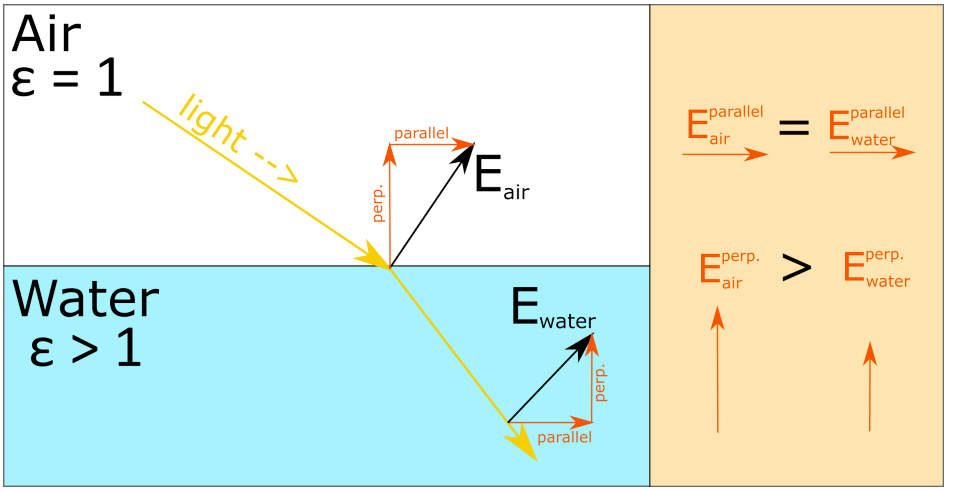


Figure 3.1: Illustration of refraction of light beam going from air to water, with added vectors indicating electrical field of light, E_{air} and E_{water}, including components parallel and perpendicular to the air-water surface.

the equations can be written:

$$\begin{aligned} \nabla \times E_{air} &= -\frac{\delta B}{\delta t} = \nabla \times E_{water} \\ \epsilon_{air} \nabla \cdot E_{air} &= \rho = \epsilon_{water} \nabla \cdot E_{water} \end{aligned} \tag{3.3}$$

Giving two restrictions:

$$\begin{gathered} E^{air}_{parallel} = E^{water}_{parallel} \\ \text{and} \\ \epsilon_{air} E^{air}_{perpendicular} = \epsilon_{water} E^{water}_{perpendicular} \end{gathered} \tag{3.4}$$

Meaning that as the light cross the surface into the water, the parallel component of the electrical field vector stays equal, and the product of the permittivity of the material (ϵ) and the perpendicular component is equal. Because $\epsilon_{water} > \epsilon_{water} = 1$, this means the perpendicular component must be reduced by $\frac{\epsilon_{air}}{\epsilon_{water}}$, which in turn change the angle of the electrical field vector and thus, the angle of the light through the water.

When considered an electromagnetic wave, light also has the following relation:

$$v_f = \lambda f \tag{3.5}$$

where v_f is the phase velocity of the wave, λ is the wavelength or the distance over which the wave repeats, and f the frequency of repetition. In vacuum, v_f is the speed of light, c. However as shown above when entering some transparent medium this velocity decreases. Due to boundary conditions, the frequency does not change, so both phase velocity and wavelength decreases. The decrease depends on the material and the wavelength $v = c/n(\lambda)$, and $n(\lambda)$ is defined as the materials *refractive index*.

Knowing the refractive index of the material traversed by light and the angle of the incident beam, bending angles of the light beam in figure 3.1 can be found by Snell's law:

$$n_{air} sin(\theta_{air}) = n_{water} sin(\theta_{water}) \tag{3.6}$$

where θ_{air} is the incident angle of light in air, and θ_{water} the bent angle through water.

3.1.2 Radiation Pressure

In addition to passing through transparent materials and bending at interphases, light may also be absorbed or reflected. Absorption concerns the transfer of energy to the material, exciting the atoms of the material to a higher energetic state.

This is not relevant for this project, however, and will as such not be discussed any further. Reflection occurs when the incident light beam hits the boundary surface or interphase between two transparent materials, but stays in the original media rather than transferring into the second one. For a flat reflective surface, incident beam angle and reflected beam angle are equal, respective to the surface. In the case of spherical or uneven surfaces, reflected beams will scatter in several directions depending on the local area they reflect off. When the beam is reflected, it causes a change of direction of the photon, meaning a change in its momentum. While photons are mass-less, any change in momentum adhere to the law of conservation of momentum, meaning an equal but opposite change in momentum must occur for the reflective material. This force over the reflective area is described as a *radiation pressure*, and is defined by

$$P_{total} = P_{incident} + P_{emitted} = 2\frac{I_f}{c} \tag{3.7}$$

with P is the pressure, I_f the *irradiance* or power of light over an area (W/m^2 usually) and c the speed of light in vacuum. Because the beam is reflected either completely or partially, depending on the incident angle, $\theta_{incident} \geq 0$, total momentum change becomes the sum of incident and emitted momenta. Radiation pressure also occurs when light is absorbed or bent by the material, but P_{total} is then either equal to $P_{incident}$, or some function of the angle difference due to the bending, respectively. The forces generated by radiation pressure are in general too small to even notice for humans. Although we are constantly bombarded by sunlight, we do not have to account for the radiation pressure when moving about the same way as we sometimes do with strong winds. The forces become more noticeable, however, when acting on objects in outer space. Here they may be the main force acting on objects besides gravity and viscous movements of our atmosphere, resulting in some interesting applications(Rios-Reyes and Scheeres 2005). Similarly, the forces can be observed in microscale, where they become even more apparent.

3.2 Optical Manipulation

The first proper demonstration of optical micromanipulation in laboratory experiments was done by Ashkin (1970). Radiation pressure was here proven to impart sufficient momentum alone on microscopic particles to push them along the path of light propagation. Later, in 1986, he improved on the concept by introducing a counteracting gradient force to the radiation pressure, and effectively trapping a small particle in a small spacial volume. By having a Gaussian beam profile and focusing the beam into a small volume of high intensity, a gradient force towards the beam center arise from the Lorentz force, moving the particle towards it, as

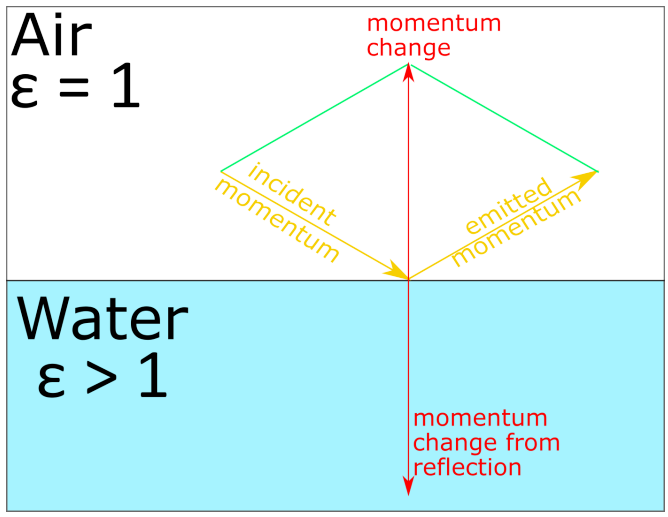


Figure 3.2: Illustration of radiation pressure from light hitting a reflective surface, with momentum of incident and emitted light, the total momentum change and opposite imparted momentum on material from reflection.

illustrated in figure 3.3. By further focusing this beam profile through a high numerical aperture (NA) optic set-up, as shown in figure 3.4, a focal point of high intensity is created. The gradient forces illustrated in figure 3.3 now point inwards in all 3 dimensions, creating an *optical trap*. The gradient forces are now able of moving interactable particles towards focal point. Ashkin named this discovery "*single-beam gradient force optical trap*"(1986), although it is more commonly referred to now as *optical tweezers* (OT), due to the ability it has of picking up and holding microscale particles, like a pair of tweezers.

3.2.1 Optical traps

To summarize, when approaching an optical trap, a dielectric particle will interact with the light of the focused beam, and experience a force to the transfer of momentum from incident photons. The forces of this optical trap consist of the two components Ashkin proved: (i) a scattering force from the radiation pressure, eq. (3.7), from photons scattered by the particle surface, acting along the light propagation path, commonly defined as z-direction, (ii) and a gradient force acting along the intensity gradient to the traps focal point (figure 3.3 and 3.4). In order to have stable trapping of a particle, gradient force (ii) must be sufficient to balance the constant "push" of the scattering (i), achieved by sharply focusing the trapping laser beam. The result of this is that the actual trap center is positioned slightly behind the focal point along light propagation, where the z-components of acting

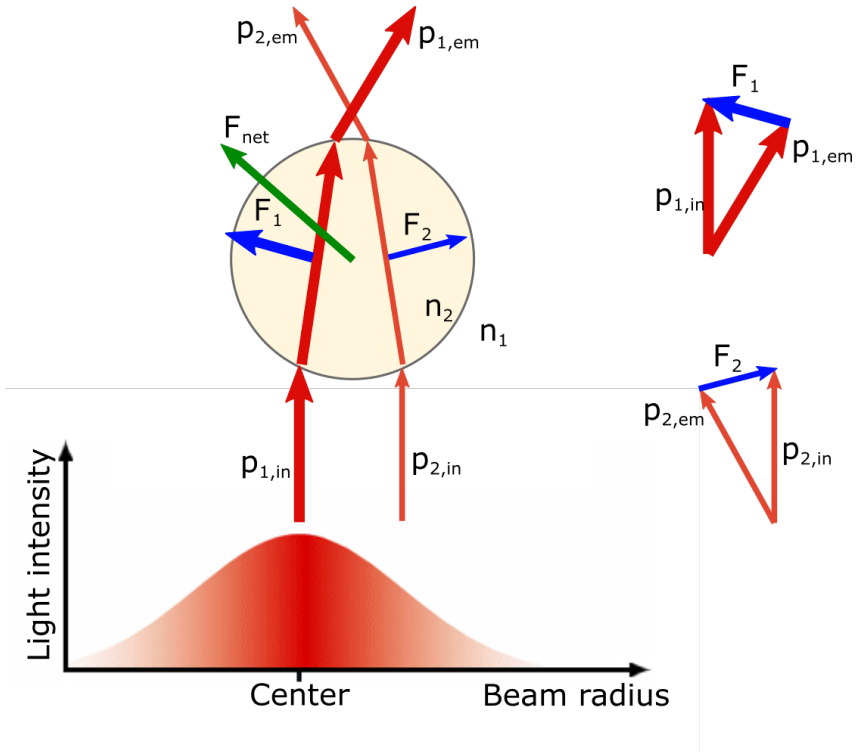


Figure 3.3: Light beam with Gaussian profile acting on a particle. Illustration with 2 vectors of different intensities (1 and 2) from the beam profile, their propagation paths (p, red) of incident (in) and emitted (em) light. Forces (F, blue) associated with the momentum change caused by the light–matter interactions, and the net force (F, green) moving the particle towards beam center and along light propagation. Thickness of vectors indicate intensity, and refractive index of particle, $n_2 > n_1$, that of the surrounding medium.

forces balance.

Trapped particles are not, however, completely stationary in their respective traps. External forces acting on a trapped particle are described as a harmonic oscillator(Moffitt et al. 2008). Thus, traps are defined to have a *spring constant* k, like that of a mechanical spring following Hooke's law:

$$F_k = -kx \tag{3.8}$$

where the spring force F_k scales with the distance from the trap center, x. In OT set-ups, the spring constant is also referred to as the *trap stiffness*, which gives the forces in each direction as a function of the particles displacement from trap center. Now the particle is held relatively still, with some freedom to diffuse by Brownian

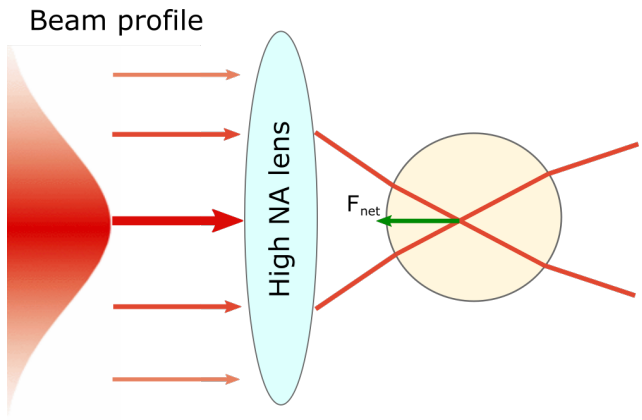


Figure 3.4: Illustration of optical trap, where a light source with Gaussian profile is focused through a high NA objective to create a small volume of highest intensity light at focal point with a gradient force towards it, trapping a particle.

motion(Huang et al. 2011), Observing this restriction of motion, combined with knowledge of its environment like viscosity of bulk medium and temperature, is usually how the spring constant is obtained(Deng et al. 2007). The outgoing trapping beam, after focal point, can be re-aligned to parallel waves using a second high NA lens, returning to a 2D Gaussian profile that can be analyzed using a quadrant photodiode (QPD)(Pralle et al. 1999) to detect any changes in position relative to the trap center. Change in relative position will shift the area of light by

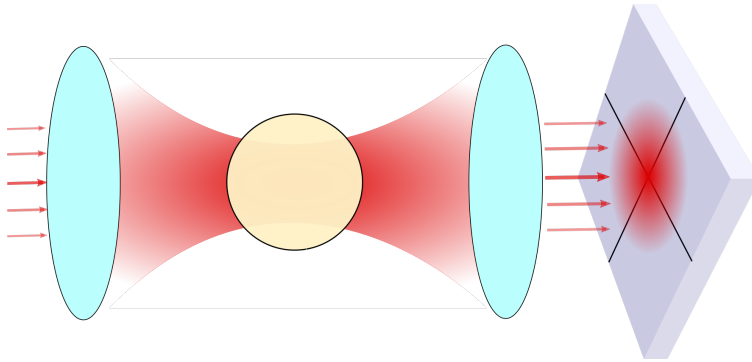


Figure 3.5: Illustration of trapped particle in optical trap (*left*) with lenses along light propagation before and after trap, and the now parallel light casting a Gaussian profile on a quadrant photodiode (*right*).

which the trapped particle casts as a "shadow" on the QPD, giving different voltaic outputs in the 4 regions of the detector, illustrated in figure 3.5. Combining this

detection with high precision software controlled optics gives an optical tweezer set-up the ability to move and position trapped objects with high precision, while simultaneously record any changes in position relative to the trap, converted into forces via the trap spring constant.

This (combined with accurate software controlled optics) gives the optical tweezer set-up the ability to exert and quantify forces up to and beyond 10^{-10} N on particles ranging in size from nanometers to micrometer, while simultaneously measuring relative displacements from the trap in 3 dimensions with sub-micrometer and sub-millisecond accuracy.

3.2.2 Calibration

As with all wide field microscopes, to evenly illuminate the sample cell the first step is to adjust for Köhler illumination(Goldberg 1980) – assuming that samples have been produced – injected into relevant sample cells and mounted in the microscope sample stage. In an OT set-up, this is achieved by first focusing the trapping objective to show the inside of the sample cell, then closing the back field diaphragm and focus the condenser up and down until the edges of the diaphragm is shown sharply in the image, before opening it to show the whole sample plane again.

After successfully trapping particles in each of the traps, a rough size estimation can be performed with a built in software measure tool. While the tool reports size in nanometers, the Rayleigh Criterion causes error in resolution, limiting accurate size determinations due to diffraction of light. For visible light, a general rule gives the limit of resolution at roughly 200nm, giving all measurements this as a positive and negative error when reporting size.

This estimated size is then used for calibrating the optical traps, calculating the trap stiffness and force sensitivity based on particle size, system temperature and bulk viscosity. From here, a known force applied to the trapped particle could be used to calculate the corresponding signal from the QPD to calibrate the trap, using Stokes law:

$$F_s = \gamma v, \qquad \gamma = 3\pi\eta d \tag{3.9}$$

where force F_s depends on the moving velocity v and the *drag coefficient* γ, that further relies on known (spherical) particle diameter d and bulk medium viscosity, η. However, a more accurate way to calibrate traps is to record the Brownian motion of trapped particles. The recorded diffusive motion of the particle plotted as a power spectrum presents a Lorentzian shape, and a curve $S(f)$ can be fitted to

the experimental data.

$$S(f) = \frac{D}{2\pi^2(f_c^2 + f^2)} \tag{3.10}$$

With a particle in a trap with a Hookean-like (eq. (3.8)) spring constant k stiffness, the characteristic roll-off corner frequency of the spectrum is given by $f_c = k/2\pi\eta$. The *diffusion coefficient* D, comes from the Einstein-Stokes relation,

$$D = \frac{k_B T}{\gamma} \tag{3.11}$$

This gives the spectrum density in eq. (3.10) a dependency on both bulk viscosity and trapped particle size from eq. (3.9). By comparing the measure power spectrum in V^2/Hz with the expected movement in nm^2/Hz, the detector sensitivity can then be calculated in m/V for all three spacial dimensions, giving the trap a stiffness and QPD sensitivity, which usually reports in pN/μm and V/μm, respectively.

A proper walk-through of the calibration procedure for the Nanotracker 2 is supplied by JPK/Bruker.

3.2.3 Nanotracker 2

The Nanotracker 2(JPK Instruments, Berlin, Germany) is the commercially available OT instrument used for all work in this project. It is based on an inverted microscope, Zeiss Axio Observer, which simply means that samples are mounted over the light beam instead of under.

Trapping and *detection* objectives (microscope objective and condenser, respectively) are both oil/water immersion objectives, removing aberrations from light bending on passing from objectives to air to cell. Combining this with the wide field scheme of the microscope gives the instrument clear images of the sample cell, evenly illuminated and to a very thin plane perpendicular to light path.

The tweezers are equipped with a 3W TEM00 laser, and a software controlled beam splitter plus frequency shift on one beam and optical filters along light path after sample gives the instrument two distinct optical traps, T_1 and T_2, that can be controlled individually and do not interfere on each others output signal QPD.

In addition to QPDs on the back focal plane, the instrument is also fitted with a digital camera, giving a video feed of 30 frames per second to the control software for ease of trapping and control.

Also in addition to the specifications above, the instrument also supports control using a wireless gamepad (Logitech, Xbox) with dual analog thumbsticks steering each trap.

3.3 Relevance

New methods for trapping small particles and even single molecules while measuring forces acting on them have been developed during the last decades. In studies on molecular processes where force has a crucial role, micromanipulation techniques such as atomic force microscopes, optical tweezers, and magnetic tweezers have excelled(Neuman and Nagy 2008).

Optical tweezers have been shown to be very successful for a wide range of fields of research since its invention(Polimeno et al. 2018), and within physics(Anderegg et al. 2019, Pettit et al. 2019), biology and medicine(Hadjialirezaei et al. 2017, Thammawongsa et al. 2012, Wang et al. 2011) its application has been well covered. In the studies of emulsions and microdroplets, however, its use have been limited. Several studies have at some point included a proof-of-concept, showing the OT abilities within respective fields, and for the most part the studies have focused on: (i)formation of nanofluidic networks interconnecting droplets(Woods et al. 2011), (ii) deformation of ultralow IFT droplets(Hargreaves et al. 2015, Ward et al. 2006) (iii) pH dependant interactions between individual emulsion droplets(Bauer et al. 2011), and (iv) forces acting on droplet of micro- and macromolecular stabilization(Nilsen-Nygaard et al. 2014). Nilsen-Nygaard et al. measured the interaction between stabilized emulsion droplets with OT, and similar studies on interaction forces between colloidal particles have naturally been explored with OT. Recent advancements contributed by Chen and coworkers include quantification of interaction forces as functions of varying NaCl and SDS concentrations(2018), later with non-ionic surfactant coating(2019), and a switchable surface-active colloids system(2020).

Studies focused on the dynamics of coalescence using OT are rather lacking, which this project aims to expand upon. Some relevant studies have looked at droplet coalescence, however, with both Otazo et al. (2019) looking into the aggregation of particles and following coalescence in partially crystalline droplets, and Mitsunobu et al. (2017) that performed a study with focus on induced coalescence on droplets trapped in a custom temperature controlled fluid cell. Coalescence has also been thoroughly studied for aerosol particles trapped(Power et al. 2012) and holographic optical tweezer set-ups(Bzdek et al. 2016, Haddrell et al. 2017).

Chapter 4

Scope of thesis

The work presented in this thesis aim to improve the understanding of the coalescence event, on a micro- and atomistic scale, as a bridge between theoretical and practical, using optical trapping and manipulation of droplets in combination with rare event MD simulations.

Coalescence in foams and emulsions is far from being fully understood, however it is known that the phenomenon is controlled by the drainage of a thin liquid film separating approaching droplet interfaces, and its eventual rupture. This rupture is influenced by the system hydrodynamics, droplet surface rheology, and thermal fluctuations of the system. This understanding of the underlying mechanisms of coalescence is of great importance when trying to understand emulsion behavior and stability.

Current models of coalescence depend heavily on some specific parameters of the system such as exact droplet size and separation distance, in addition to approach velocity. The models also predict certain droplet behaviors occurring during approach, like the dimpling of interfaces at some critical thin film height, and an absolute maximum approach velocity where droplets begin to bounce off each other rather than coalesce. With commonly used experimental approaches, these parameters are difficult to control, and the described behaviors hard to observe.

As coalescence is also characterized as an extremely rapid event, its rapidness makes experimental studies challenging, and demands specialized techniques and tools in order to capture it occurring. With these challenges, most previous studies have been working with droplet sizes ranging from 20 μm – 200 μm. This excludes many emulsions of different applications, that all have droplet diameter below 10 μm, that have shown behaviors that deviate from the larger droplets, especially

when considering their deformation and internal pressure, and hydrodynamic effects. Although rarely regarded in models, droplets with diameters below 10 μm are considered the real size scale of many emulsions in wide use.

Experimental studies regarding coalescence also focus less on the effect of approach velocity to coalescence times, which has previously been explained by the challenge of controlling said velocity, compared to other important parameters (diameters, temperature, viscosity, interfacial tension, etc.). So as opposed to studies using flotation or flow-based methods of approaching emulsion droplets, presenting challenges to the velocity, optical tweezers allows for extremely precise control of the velocity of trapped objects. Optical tweezers are not limited to the sedimentation velocity or flow arising from individual pumps of a fluid cell setup. The tweezers may also well emulate the random droplet collision otherwise used in fluidic-based experiments.

While being mainly used for the purpose of rate of reaction calculations, rare event simulations have been used to explore the possible phase space of systems. Path sampling gives its user the opportunity to model reactions which would never be practically attainable otherwise with molecular dynamics, making a simulated thin film breakage within reach. Implementation of more sophisticated Monte Carlo moves and collective variable have made the RETIS algorithm capable of finding otherwise hidden mechanisms and parameter correlation. With the addition of a custom plotting and analysing tool, information gathering both after and during simulations is made even better, identifying meta-stable states and descriptor importance, which when used correctly could provide much needed information regarding the dynamics of coalescence.

Chapter 5

Summary of results and discussion

This chapter and sections included provide insight into the additions made to the PyRETIS software in papers **I** and **II**, and a more detailed explanation of the software improvements of **Paper II** along with a model example study. Papers **I** and **II** additionally facilitated in teaching the complex theory of rare event simulations to students at BSc and MSc levels, which are laid out in an additional paper, it is not however included in the bulk work of this thesis.

Included is also a summary of the results and discussion of the papers **III** and **IV**, containing the experimental study using optical tweezers, individually, and in combination with MD simulations. Experimental details are contained in respective papers in appendices, and will be briefly repeated in text when relevant.

5.1 Emulsions in optical traps

5.1.1 OT emulsion requirements

When producing emulsions for use in an optical tweezer set-up, the emulsion must meet some key requirements. Firstly, as explained in section 3.2 on optical manipulation, in order to effectively displace a small particle to the focal point of a beam with Gaussian profile, the refractive index of the particle must be higher than of the surrounding medium. While in theory, any positive change in refractive index from bulk to trapped particle generates the gradient force necessary to move the particle, common practice dictate a lower bound for this at 0.2. With this, most W/O emulsion systems are excluded as model oils usually have much higher indexes than water. Furthermore, as the OT laser traps exert forces in the range of pN

on trapped droplets, drag force and sedimentation entail an upper limit in droplet size, as the droplets tend to escape trap at higher moving velocities or with slight displacements from trap center due to collision or Brownian motion. Naturally, with a greater difference in refractive indexes, traps are able to exert greater force on trapped droplets, which would allow for trapping larger droplets.

The optimal size range for trapped particles diameter in OT is 2–5 μm, as this range is the most suited when calibrating trap stiffness and sensitivity. Going lower in droplet diameter is absolutely possible, but one has to keep in mind that the relative error in visual size estimation becomes greater due to the diffraction limit of optics.

As seen in later results, an upper trapping size limit for our chosen emulsion system was at $\sim$ 25 μm. At this point, traps were still able to hold onto large droplets by changing the relative power output between the two traps; any movement of the trap or collision with other smaller dispersed droplets would knock the larger droplet out of the trap center, and quickly float it to the upper glass plate of the fluid cell.

5.1.2 Preparation of macroemulsions

Several preparation schemes for producing O/W macroemulsions (0.5 – 100 μm) for trapping in OT were tested, using dodecane ($C_{12}H_{26}$) as model oil.

Microfluidic chips have shown great potential for creating custom emulsions with pre-defined size range requirements, in very narrow distributions. Initially we had planned to work in co-operation with students at the Department of Physics at NTNU, were they would prepare microfluidic cells and create emulsions with specific droplet diameters for use in OT. Instead, we were able to use a modified X-junction design provided by Swapnil Vilas Bhujbal, with a nozzle width 15μm, to attempt to produce them ourselves. As we were at this stage trying to create *pristine* emulsions (only water and oil), the microfluidic chips showed problems when trying to shear the oil stream with water to create droplets, as this requires a much lower interfacial tension in order to deform the oil–water interface at the intersection. Furthermore, the instability of emulsion droplets post-nozzle caused immediate coalescence of a large fraction of the generated droplets, moving the size distribution up, increasing the spread of droplet diameters. Emulsions produced with microfluidics were inserted in the OT for inspection, and although there were droplets present within desired range, there were so few of them that we spent an abundant amount of time looking for them. Finally, this technique, while excellent at producing narrow distributions under optimal conditions, produces tiny sample volumes. The pumps that control the flow of liquids in the cell operate in the range of μL/hr, meaning that to produce sufficient volumes to mount

two–three samples in OT would take several hours, which proved to be needlessly time-consuming for the amount and quality produced.

An easier and more reliable way of quickly homogenizing an O/W emulsion was by the use of an Ultra-Turrax colloidal mill. Initial observations showed that the colloidal mill was capable of generating a distribution of droplet diameters ranging from sub-μm to >100 μm and beyond, covering our goal range for optical trapping (1–30μm). Compared to microfludics, the colloidal mill also provided emulsions with higher a droplet density, because we had more control over the initial volumes of water and oil. This method of preparation required, however, relatively huge volumes of sample, with total volume of water and oil >50 mL, necessitating access to chemicals in greater quantities compared to all other approaches. Combined with the instability of the pristine emulsions prepared with the colloidal mill, this meant that large sample volumes that were stored in-between recordings on OT. These samples would over time sediment, cream and coalesce, and needed constant re-homogenization prior to new recordings in OT. In addition, this also made them prone to contamination, with the continued submersion of mill rotor in samples. So while providing a decent droplet population for measurements, the drawbacks of the colloidal mill for our use made us consider alternative methods.

In the interest of reducing time spent during sample preparation, and not forfeiting too much in the way of droplet diameter distribution, we tested a most simple method of homogenization. Detailed description of the method is reported in papers **III** and **IV** in Appendix, and involved short mixing of sample volumes $\sim$ 1000 μL in Eppendorf tubes on a vortex mixer. Although this approach produced a wider range of droplet diameters, it benefited from both reduced sample preparation time and no longer a need for re-homogenization, thus shifting the focus from sample preparation to actual time spent on OT, locating droplets and measuring interactions. Small volumes of these samples could be taken from mid-tube area, avoiding the larger diameter droplets that sedimented faster, allowing us to gather droplets with the ideal size for our intended studies.

While both microfluidics and colloidal mill approaches would be expected to behave more properly when applied to a stabilized emulsions system with surfactants present, they both present glaring issues regarding sample volumes that made us disregard them for use also once we began preparing our stable samples.

5.2 Effects to droplet interactions

Droplets in OT were approached with software control, using one of the two schemes laid out in **Paper III**. A brief summary follows.
Retract–Extend: the standard procedure for depletion force measurements on OT.

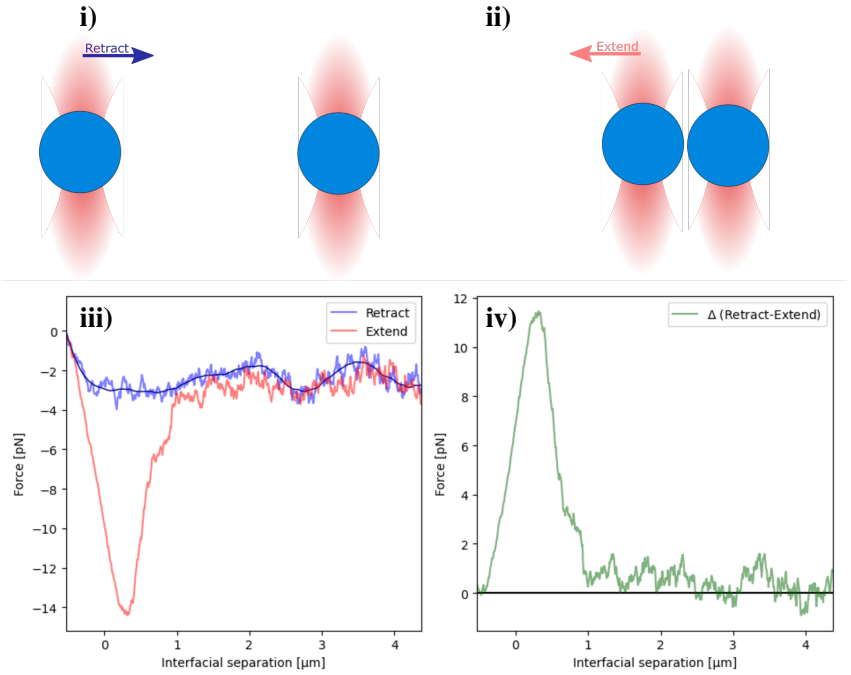


Figure 5.1: Illustrations on top show the two moving segments, retract **i)** and extend **ii)**. Arrows indicate the direction of movement of the leftmost trap. Force output from the two segments is shown in **iii)**, overlaid in the same plot against droplet separation distance [], and colored to match **i)** and **ii)**. Finally the difference in output between retract and extend is plotted in **iv)**, here presenting the depletion force as a peak, otherwise as the work of separation the area below the curve from numerical integration. Figure lifted from **Paper IV**

The naming scheme is a result of the naming of segments on the Nanotracker 2, where approaching traps at velocity is defined as a *retract segment*, and separating traps similarly is defined as a *extend segment*. An illustration of the droplet and trap positioning and movement, with subsequent force outputs, are shown in figure 5.1. Trapped particles are positioned in traps T_1 and T_2, their sizes determined visually in software, and trap stiffness calibrated. Traps are approached to a *retracted* state, by moving T_1 towards T_2 at a fixed velocity v_{app} for a user-defined distance d_{app} before stopping. Then the traps a kept in this state (a *retracted pause* in the Nanotracker 2's software) for a set time of 1 s, before extending T_1 back to initial position at v_{app}. During this, the output of the QPD is converted into forces in x,y,z directions and plotted continuously in *force–time* plots. The force–time curve of retract segment is then smoothed with a filter, to get a cleaner (noiseless) profile of the droplet movement, shown with a darker shade in figure 5.1 iii). In

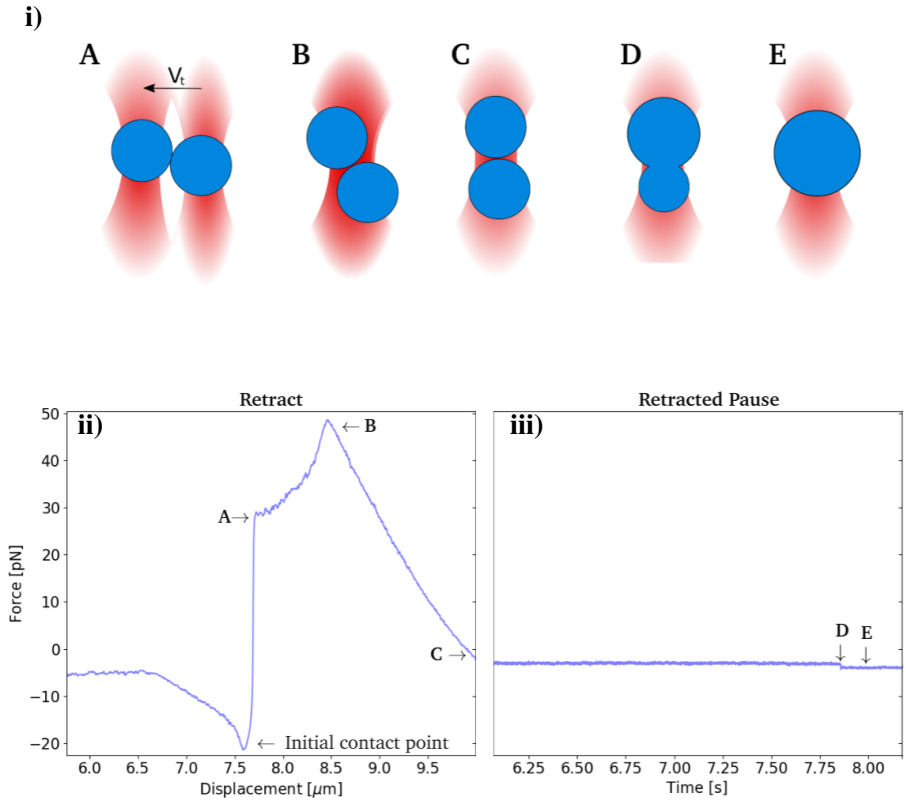


Figure 5.2: i): Top illustration show positions of the oil droplets when bringing them in contact and subsequently forcing them to overlap (retract–overlap trap movements). The traps approach each other at velocity V_t, until the contact point is reached. The contact between the droplets leads to their displacement from trap centers, beginning to show in **A**. **A–C**: The two optical traps approach each other until they completely overlap. This leads to droplet alignment along beam axis (**C**), followed by droplet coalescence (**D**, t_{coales}). **E** shows the resulting droplet formed after coalescence.
The two bottom panels show force recordings obtained during retract, **ii)**, and retracted pause, **iii)**, segments. Droplet positions in traps **A–E** are indicated in both panels. Only a narrow section of the X-axis is presented in order to focus on points of interest. *Left* plot shows force against displacement of traps, initial position with a distance of 10 μm. *Right* plot shows time [s] in overlapped state. Figure lifted from **Paper III**.

order to get the attractive force of depletion between droplets, the smoothed retract curve (5.1 iii), dark blue) is subtracted from the extend curve (5.1 iii), red), giving a positive force–distance curve (5.1 iv), green) with noise of retract curve included. In the event that this force difference did not present a clear peak of depletion force, the *work of separation*, J, was used. This work was calculated with the same force curve in 5.1 iv), by the numerical integration of the area below this positive curve, obtaining a value of the work required to separate the droplets [pN·μm] J.

Retract–Overlap: similar to the former scheme, except that calibrating trap stiffness is not necessary, as it is only used to find time of coalescence. T_1 is moved at v_{app}, but d_{app} is now the exact distance separating trap centers, so that during retraction the traps are brought to fully overlap, ensuring that droplets are properly in contact. Contact point is determined visually from the continuous and uncalibrated force–time curve, and the eventual time of coalescence too, giving the time difference from contact to coalescence event at the max time-resolution of the OT, 50kHz. Figure 5.2 i) illustrates the overlapping of droplets, with the manual analysis of force output graphs in the bottom half of the figure, that provide the time from contact to coalescence, shown in 5.2 ii) and iii) respectively, and indicated with corresponding illustrations of traps and particles from 5.2 i).

5.2.1 Ionic strength effect on droplet stability and work of separation

Initial tests on the pristine emulsions prepared showed us that a significant amount of ionic bulk strength was necessary in order to observe coalescence in the optical traps. Through purely qualitative analysis we determined an ionic concentration of 0.1 mol/L NaCl sufficient to coalesce trapped droplets.

5.2.2 Effect of size distribution to work of separation

Choice of droplet sizes for observation was done carefully, however in some situations the scarcity of droplets and overall randomness of sizes present immediately upon microscope focusing, made us pick sets of droplets with some deviations in size. With the wider distribution of droplet sizes, the effect of said distribution on depletion force and work of separation was studied. Furthermore, in order to study droplets approaching the upper limit of OT trapping, which often would escape from view and hit the upper glass cover plate, larger droplets were built by coalescing several smaller ones in each trap. This further obfuscated any determined choice in droplet sizes reported, as droplets were combined by droplets still present in population (within a narrow volume along focal plane of OT microscope). Additionally, force–time curves of the retract–extend curves were not always presenting a clear peak and cut-off force for depletion, with a wider curve, or "hump", rather than a nice maximum and cut-off.

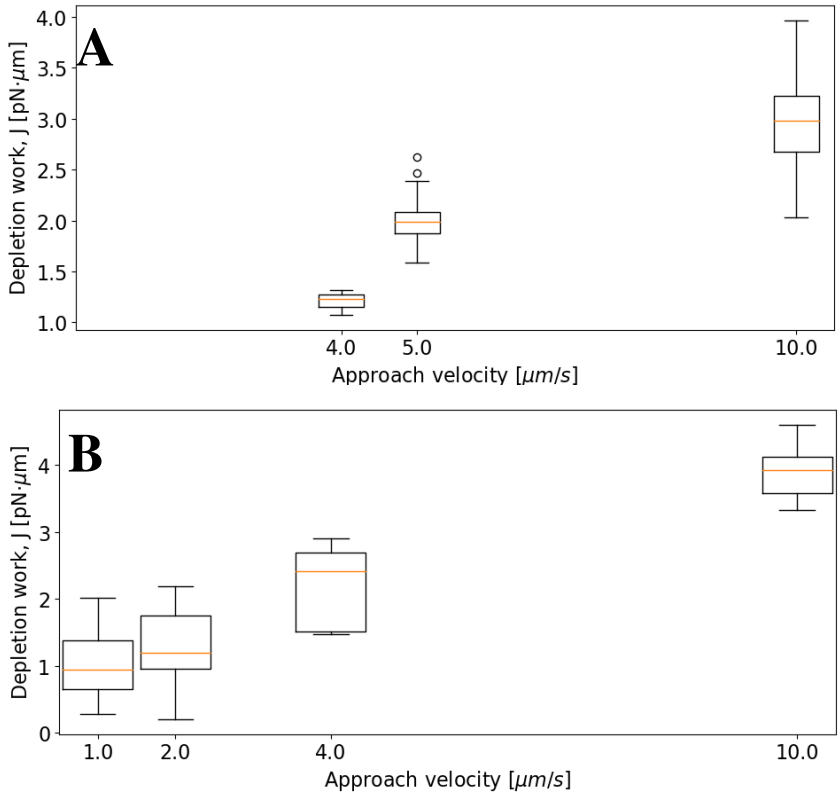


Figure 5.3: Boxplots of depletion work, J, as a function of increasing approach velocities. The linear regression lines are obtained from average values calculated based on the data sets. The NaCl concentration in the bulk was 0.05 mol/L **A** and 0.15 mol/L **B**, and the average droplet size was in the range 6–10 μm. Figure lifted from **Paper III**.

This was visible in the results shown in figure 5.3, where the work of separation for droplets in very dilute samples of 0.05 mol/L NaCl measured in the 1 pN/μm range for approach velocities higher than droplets in higher 0.15 mol/L concentrations. At these low levels, the calculated work was difficult to obtain as the force output was beginning to become indistinguishable from the background thermal noise of the traps.

Previous work by Nilsen-Nygaard et al. (2014) had showed that the depletion force and thus work of separation, J, increases with increasing droplet/particle size. This was verified with results for larger sets of droplet diameters reported in figure 5.4, where very large droplet pairs were approached at a fixed 10μm/s velocity in 0.1 mol/L NaCl solution.

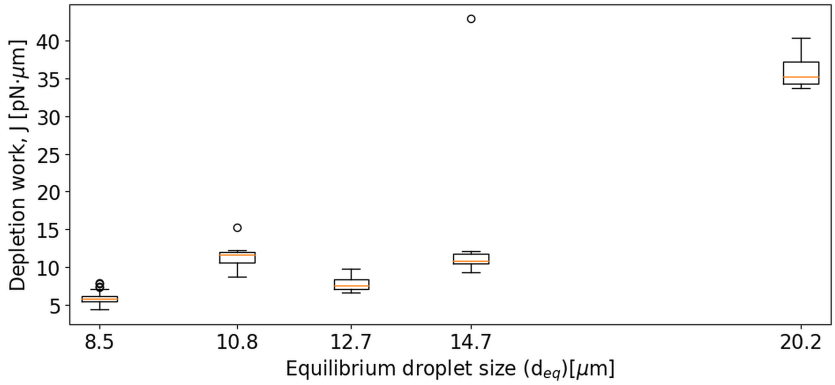


Figure 5.4: Boxplots showing depletion work J against equilibrium droplet size, d_{eq}, approaching at 10 μm/s, in 0.1mol/L NaCl concentration. Figure lifted from **Paper III**.

Similarly, our data shows that droplets retracting and extending at 10μm/s have a correlation between increasing size and increased work of separation, reaching quite high values of J even at the relatively normal velocity chosen.

Oil droplets of diameter >20.2μm are not presented in results in figure 5.4, as this was found to be approximately the upper limit of the tweezer trapping capability. Recording the data shown, during one retract segment the droplets coalesced into T_2 (stationary trap), and instantaneously escaped. Although this event may have been promoted by the added influence by the then retracted T_1 trap, exerting some force and displacing the newly coalesced droplet, thus causing its escape, this serves as a decent upper bound for trapping diameter.

5.2.3 Stokes drag and droplet trap displacement

On exploring the effect of approach velocity, induced drag force on moving droplets had to be accounted for. Increased moving velocity of droplet in T_1, individually or in combination with medium–large droplets, caused the drag force of eq. (3.9) to become noticeable and to affect the recorded retract curves. As the traps instantaneously accelerate to v_{app}, droplet inertia causes the droplet to lag slightly behind trap center position before moving. Since the movement usually was rather short, i.e. in the range of 1–2 droplet diameter separation between droplet interfaces, approaching droplet in T_1 did not experience a case where drag force F_d and spring force F_k equilibrated, which would record as a plateu of F_k. Rather, force curves showed a slope of force during the retract segment, dependant on droplet size and moving velocity. During the extend segment, depletion force would keep droplet interfaces in contact, hindering this slow distension behind moving trap center, un-

til trap force overcame depletion force. At this point, the droplet interfaces would separate suddenly, and the now "extended spring" of the trap would quickly draw the droplet towards it center, whilst still moving. The droplet would then catch up to the trap, and fluctuate a bit with the kinetic energy left from moment of separation, thus presenting a more standard depletion curve initially, followed by a sloped one, with possible small fluctuations along on the way back to initial trap positions. In practice, this meant that retract curves had to be corrected for this slope in addition to the overall smoothing, before subtracting from un-smoothed retract curves. Retract curves were not corrected for this sloping, as the cause of this effect would not be present until after interfaces separated, so the first part of the curve could be left unchanged for the subtraction, and the second part which presented sloping was neglected.

5.2.4 Effect of velocity on the work of separation

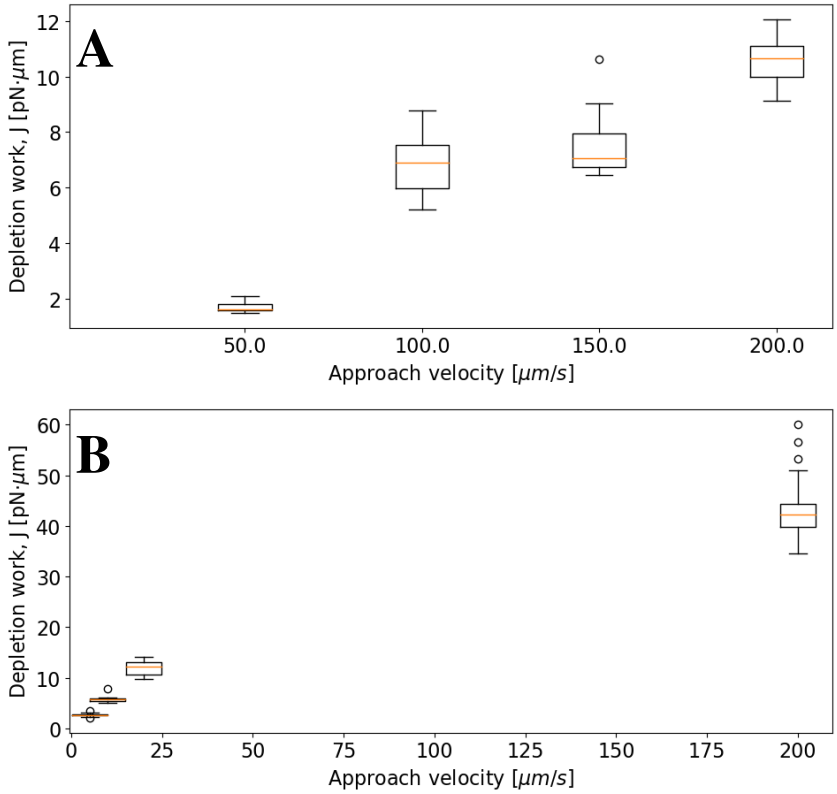


Figure 5.5: Boxplots of depletion work J at 0.1 mol/L NaCl concentration and different approach velocities. Linear regression lines from average values in data sets. **A** Average droplet sizes 3–4 μm, **B** average droplet sizes 8–9 μm. Figure lifted from **Paper III**.

Investigation of the effects retract and retract velocities had on the depletion force and work of separation J, was divided into two different size ranges, 3–4μm and 8–9μm, and presented in figure 5.5. While droplet choice was done carefully, some variation in sizes chosen still occurred. Drag force displacement explained in the previous section was most apparent for the larger of the size ranges.

Droplets were retracted and extended at velocities up to 200μm/s. The data presented in figure 5.5 show the work J as a function of approach velocity of trap, and for the lower velocities the work increases linearly with approach velocity. However, droplets of the larger diameter size range deviate from the linear relation between approach velocity and work J. Data still shows a positive correlation between approach velocity and work of separation, however, but rising quite as steep as for the lowermost velocities. Although we made efforts to remove the effect of drag force on approach from the force measurements, there may be some of results of this behavior still affecting the force output after analysis, too. Additionally, this displacement from trap center could cause reduced interfacial contact and thin film establishment during the retracted pause segment, further reducing the measured depletion and work. Overall the drag force complicated the determination of depletion force and work of separation at higher velocities. And, when combined with increase droplet diameters, the drag force makes the results using schemes where it can affect the result, unreliable.

5.3 Effects to coalescence times

Coalescence times were determined using the retract–overlap scheme defined in previous sections, for our pristine emulsions with ionic strength 0.1M NaCl, as function of both droplet sizes and their approach velocity. The advantage of the overlapping scheme here allowed for more precise determination of coalescence times, as the overlap assured proper contact and thin film establishment during the retract segment, and the displacement in traps caused by droplet drag could be neglected by inspection of the un-altered force–time curves. Figure 5.2 shows the real time force output during a retract–overlap, with clear indications where interfaces are in contact and when the coalescence event itself occurs, providing a time of film drainage and coalescence to the accuracy of 50kHz. Furthermore, any errors in size determination which would propagate into errors in force output are now neglected. As the contact points and coalescence times are determinable regardless of trap stiffness and calibration, they could even be determined with the base output of the QPD.

Coalescence times determination did, however, remove the ability to continuously record times for a single droplet pair. Whether the droplet pair coalesced during overlap or not, separation of initial droplet pair was practically impossible. Droplet

pairs would either coalesce during overlap, or remain connected in one of the two traps on extension to initial trap positions, or in rare cases be ejected and escape traps on separation, similarly to events outlined in 5.2.2.

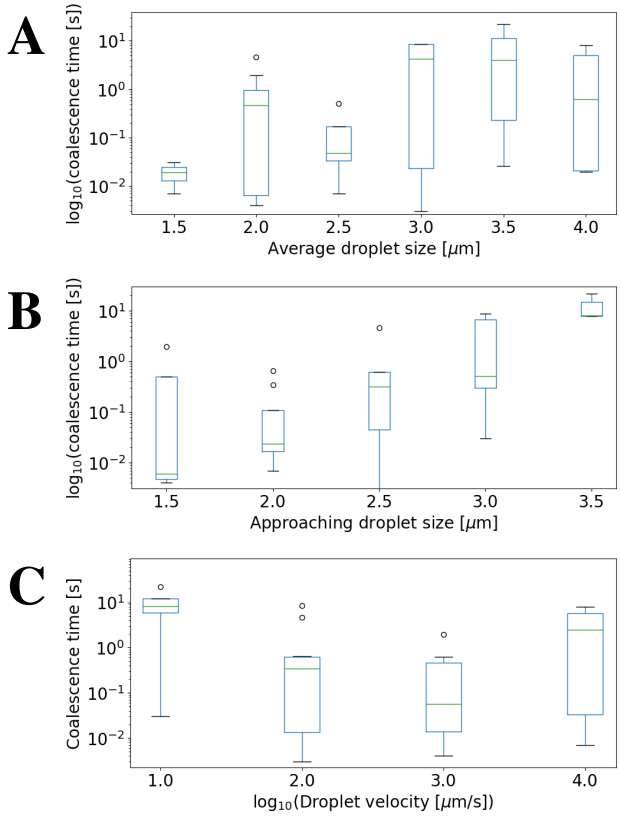


Figure 5.6: Coalescence times as a function of either droplet diameter or droplet velocity. The measurements are performed for O/W emulsions containing 0.1 mol/L NaCl in the polar phase. **A** Droplet coalescence time as a function of equilibrium droplet size d_{eq}, **B** Droplet coalescence time as a function of diameter of the droplet present in the approaching trap, and **C** Droplet coalescence time, as a function of the approach velocity of the trap ($\log_{10}$). Figure lifted from **Paper III**.

Data in figure 5.6 was collected with retract–overlap, and manually analysed for the coalescence times. Times are grouped and plotted with bins of average (equilibrium) droplet size d_{eq} (**A**), size of approaching droplet (**B**), and approach velocity (**C**). From this figure we see some indication of correlation between droplet size d_{eq} and coalescence time(**A**), and more clearly with approaching droplet size(**B**): droplets in the moving trap T_1. As expected, droplet diameter influence the coalescence times, increasing in proportion to both d_{eq} and diameter of approaching droplet, which agrees well with other studies(Gebauer et al. 2016)

Finally, coalescence times show a negative correlation with approach velocity, although with some large error intervals that obfuscate any conclusion from those results. While the lowest coalescence times were observed for the highest approach velocities, it is worth noting that coalescence times for these high velocities do seem to increase slightly, breaking somewhat from the earlier negative trend. Such behavior is postulated in the models of Ozan and Jakobsen (2019a), where a suspected minimum of coalescence times precede an increase before reaching a critical velocity, eventually resulting in bounce-events exclusively. Unfortunately this was at the very limit of what was achievable on the current OT set-up. Both the sampling rate and the approach velocity were at, or almost at, their limits, with the highest recorded approach velocities reported at 40 mm/s. While the Nanotracker 2 allows laser trap velocity up to 50 mm/s, previous detailed issues associated with high moving velocity hindered reliable measurements past 40 mm/s, and droplets escaped by drag force, displacement, ejection induced by neighbouring trap, or combinations of these. These escapes could be classified as the bounce-events over some critical velocity of Ozan and Jakobsen (2019a). However, at these rapid speeds of movement and events, nothing can be concluded, and would possibly require re-examination with high-speed imaging at equally high sampling rates.

In other coalescence studies, droplet diameters investigated are usually much larger, limited by the technique used. Thus, an important factor for the relative error observed in our results is the relative size difference between droplets arising from their significantly smaller sizes. A difference in size or error from diffraction limit becomes much more significant for our results than for i.e. microfluidics(Dudek et al. 2020), where droplets are one or two orders of magnitude larger, potentially affecting our results to a greater extent.

Nevertheless, the results presented fill a gap concerning droplet coalescence studies. Exploration into the effect of approach velocity on coalescence time is somewhat lacking, and some studies report opposing effects of velocity on coalescence time(Kamp and Kraume 2016, Kirkpatrick and Lockett 1974). Most importantly, however, this study explores the effect on much smaller droplets than others before it, below 10 μm, as pointed out by Chen et al. (2018). This further adds to understanding of how small droplets contribute to the behavior of the total, polydisperse emulsion.

5.4 Interfacial effects to depletion force and separation

As velocity and droplet diameter had been explored in the previous work, the aim of **IV** was shifted to focus on the interactions between ionic strength of solution and surfactant coverage on droplet interfaces. In addition to the experimental studies, MD simulations performed by dr. Enrico Riccardi aided in resolving these

interactions.

5.4.1 Stabilized emulsions in OT

The preparation of the stabilized emulsions for use in the following sections was as described in section 5.1.2, with the addition of sodium dodecyl sulfate (SDS, $CH_3(CH_2)_{11}SO_4^-$ Na^+) during the mixing stage.

Emulsions prepared with SDS showed greater stability than the pristine samples produced earlier. Still, no changes to the preparation scheme was done, as the advantages of the easy homogenization outweighed the slight increase in diameter distributions. Furthermore, as the added surfactant did prevent some coalescence prior to mounting sample in microscope stage, medium–large droplets had increased life-time. This increased their population in samples, making it easier to reliably find larger pairs for recording this, reducing some of the variation and reliance on size in measurements.

Visual size determination still involves some inaccuracies related to the diffraction limited optical microscopy. This limit results in a roughly 250 nm over-/underestimation of droplet diameter, the error is, however, not expected to be as significant for the diameter range used here, $\sim 9\mu$m. The visual determination of trap displacement d_{app} was by this stabilization also a bit easier, as the optimal contact point could be reached without too much probability of flocculation or coalescence events destroying the measurement pair.

A notable change in measurement on OT was during retract–extend scheme, where the now stabilized and generally larger droplets showed a much clearer peak at cut-off on separation, allowing for directly noting the height of said peak as depletion force. With stabilized emulsion droplets, coalescence times would also be expected to be longer, in addition to being approached at a fixed, lower velocity.

5.4.2 Surfactant effect on droplet separation

Depletion force measurements are shown in figure 5.7, where each datapoint is averaged of 20 runs per sample. It shows that depletion force (height of force peak upon separation) increased in correlation with SDS concentration of the sample, for all sets of ionic strengths investigated (*top*).

A similar trend appears when plotting depletion force vs ionic strength with SDS concentrations constant *bottom*. As with in our previous work, we would expect the measured depletion to be correlated with ionic strength, which initially looks valid, although some unexpected deviations did appear. In figure 5.7(*top*), one of our samples, containing 0.3μM SDS and 20 mM NaCl, shows a considerably low depletion force, in relation to the preceding and succeeding samples of ionic

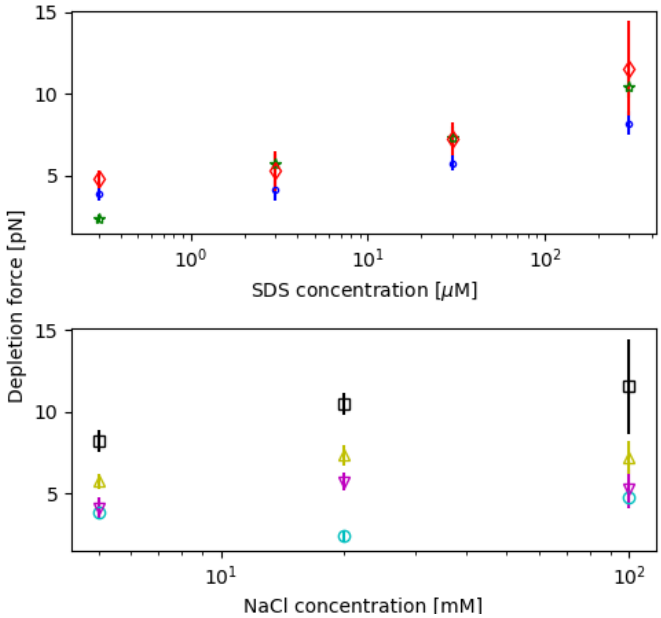


Figure 5.7: Depletion force as a function of increasing SDS concentration and increasing NaCl concentration. *Top:* Semi-logarithmic plot of depletion force as a function of increasing SDS concentration. The NaCl concentration was kept constant at 5 mM (blue datapoints), 20 mM (red datapoints) and 100 mM (green datapoints). *Bottom:* Semi-logarithmic plot of depletion force presented as a function of increasing NaCl concentration. The SDS concentration was kept constant at 0.3μM (cyan), 3.0μM (purple), 30.0μM (yellow) or 300.0μM (black). Each data point represent the average of 20 measurements. The error bars in the plots represent the standard deviation for each data point, and all droplets investigated had a diameter in the range 8–9 μm. Figure lifted from **Paper IV**.

strength. We reckon this is a result of the previously mentioned issue regarding underestimation of d_{app}, which, while less crucial for these samples, was not completely eliminated as a source of error.

On inspecting the trends of figure 5.7(*bottom*) an interesting progression is revealed, where samples of $\leq$ 30 μM SDS reach a maximum depletion force at intermediate NaCl concentration (20 mM), and plateauing for high NaCl concentration. Whether or not this is true for the 300 μM SDS sample too is indeterminate, due to the large error margins of the data set contained in this point. And although some surfactant influence on the depletion force has been shown (Ji and Walz 2015, Tulpar et al. 2007), it usually attributed more to a synergistic effect (the complexation with a polymer).

In **Paper IV** we initially propose several hypotheses to explain this observed be-

havior. One of these is that the presence of ionic surfactant (SDS) on interfaces and within thin film reduces the mobility of ions in solution. This reduction counteracts the formation of charge difference and the ensuing osmotic pressure, which generates the force of depletion. Observations made during molecular simulations later did provide this hypothesis with some merit.

5.4.3 MD simulations of thin film breakage

In total 36 molecular system boxes of varying ionic and surfactant composition were studied, to elucidate molecular interactions between surfactant molecule SDS and ions Na^+ and Cl^- in thin film prior to and during breakage. Each box consisted of two "slabs" of dodecane molecules, a thin film "slab" of water between them, containing a set number of surfactant molecules and NaCl ions. Initialization of systems and method of thin film breakage is further described in **Paper IV**.

An emulated "retraction" was performed for each system box, removing a fixed number of water molecules and re-equilibrating the system for 10 ns, checking for thin film breakage. If no thin film breakage was observed, i.e. the water thin film showed a uniform 2D density, the process was repeated. Once film breakage was observed, estimations of the critical height h_c was calculated from the simulated thin films of systems just prior to film breakage.

Surfactant interface coverage was calculated by number of surfactant molecules divided by total interface area; 2 cross sections of the system box. Ionic concentrations reported in table were calculated as the number of total ions in the initial water slab before step-wise reduction was begun. While the local concentrations of simulation systems are not directly relatable to the experimental samples used in OT, visible trends between systems should mirror observed experimental trends.

Thin film thicknesses, or heights, are reported in table 5.1. Estimated critical thin film height h_c is reported by two definitions. (*Top*) height of thin film is calculated as distance between oil slabs. (*Bottom*) height of thin film is calculated as height/thickness of water slab and ions only, surfactant layer being omitted. In addition, as interfaces were not flat, averaging was employed for both definitions of film height.

From the simulation results in table 5.1*(Top)*, it is noted that thin film thickness prior to breakage appear to be very stochastic. Parameters of both surfactant coverage and ion concentration are not showing any effect regarding the height of the thin film. Although, since the simulation scheme subjects each reduction step to a relatively short 10 ns calculation time, longer runs would potentially give slightly different results for film thickness.

Table 5.1: Critical thin film thickness calculated for different NaCl concentration and surfactant surface concentration. *Top:* height of water layer only, and *Bottom:* height of thin film with both water and surfactant included. Table lifted from **Paper IV**.

SDS mol/mm^2 \ NaCl mol/L	0.018	0.073	0.165	0.293	0.658	1.170
0.936	1.620	1.578	1.677	1.674	1.762	1.655
3.744	1.648	1.639	1.599	1.717	1.692	1.695
8.424	1.658	1.767	1.668	1.687	1.697	1.796
14.976	1.603	1.705	1.654	1.808	1.890	1.681
33.696	1.740	1.615	1.715	1.756	1.699	1.891
59.984	1.756	1.812	1.818	1.786	1.828	1.950
SDS mol/mm^2 \ NaCl mol/L	0.018	0.073	0.165	0.293	0.658	1.170
0.936	1.622	1.579	1.687	1.690	1.807	1.748
3.744	1.692	1.684	1.643	1.777	1.781	1.833
8.424	1.751	1.861	1.770	1.796	1.835	1.982
14.976	1.777	1.880	1.837	1.999	2.109	1.948
33.696	2.142	2.019	2.127	2.175	2.147	2.387
59.984	2.501	2.557	2.572	2.547	2.617	2.787

Results of table 5.1*(Bottom)* is plotted in figure 5.8, and here some trends in the film thickness do appear. It is here shown that the critical film height increases in proportion to surfactants present, and somewhat with the increased ionic presence. With increased surfactant coverage on interfaces, the amount of water necessary to stabilize the thin film is reduced.

It must be reminded that the process of step-wise reduction has its flaws, in that as water molecules are removed, local ion concentration increases. While initially incremental, and more so for the lowest concentration systems, as more and more water is removed, ionic strength in our thin film increases rapidly. It is still worth mentioning that in real thin film systems the local concentration depends on bulk concentration, and with the EDL of DLVO theory the concentration at interface is significantly higher than in bulk.

Regardless, ion mobility in thin film is observed to reduce with the presence of

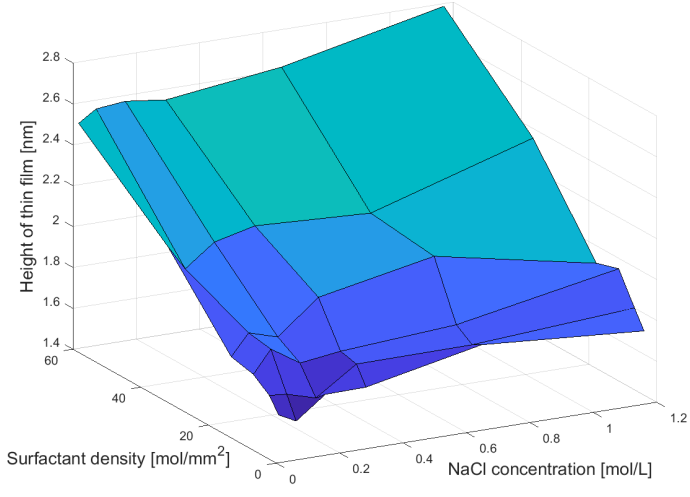


Figure 5.8: Surface plot of data shown in **Table 5.1**(bottom). The calculated height of the thin film (z-axis) is presented as a function of the concentration of NaCl in the solution as well as the interface density of surfactant on x- and y-axes respectively. Figure lifted from **Paper IV**.

ionic surfactant coverage. Additionally, in systems of initial ionic concentrations 0.293 mol/L and up, the formation of small ionic clusters was observed, as shown in figure 5.9. Here, right side column reports ion density for the system during reduction of water, going down by rows, and small clusters of ions in crystalline structure are visible prior to any deformation of water (left column) or surfactant (middle column) layers. Driven by the increased local ion concentration, these nano-crystals appeared prior to any dimpling effects on water slab or subsequent film breakage.

Inspecting the 1-dimensional ion density along thin film height during system reduction, we see how the ions position themselves in the water region close to the interfaces, with counterions further into water region. Figure 5.10 shows the densities of a Na^+ and Cl^- in a system of 0.293 mol/L NaCl and 14.976 mol/mm^2 SDS, at initial configuration and during reduction to film breakage. On reduction, the bi-modal Na^+ density in pushed into a similar Gaussian distribution as Cl^-, with overlap, in effect breaking the EDL that was observed earlier. This would destabilize the thin film, causing coalescence. It could also promote the formation of the observed ion clusters, as both ions are forced into the same volumes along film height, which could further interfere with the surfactant behavior, also

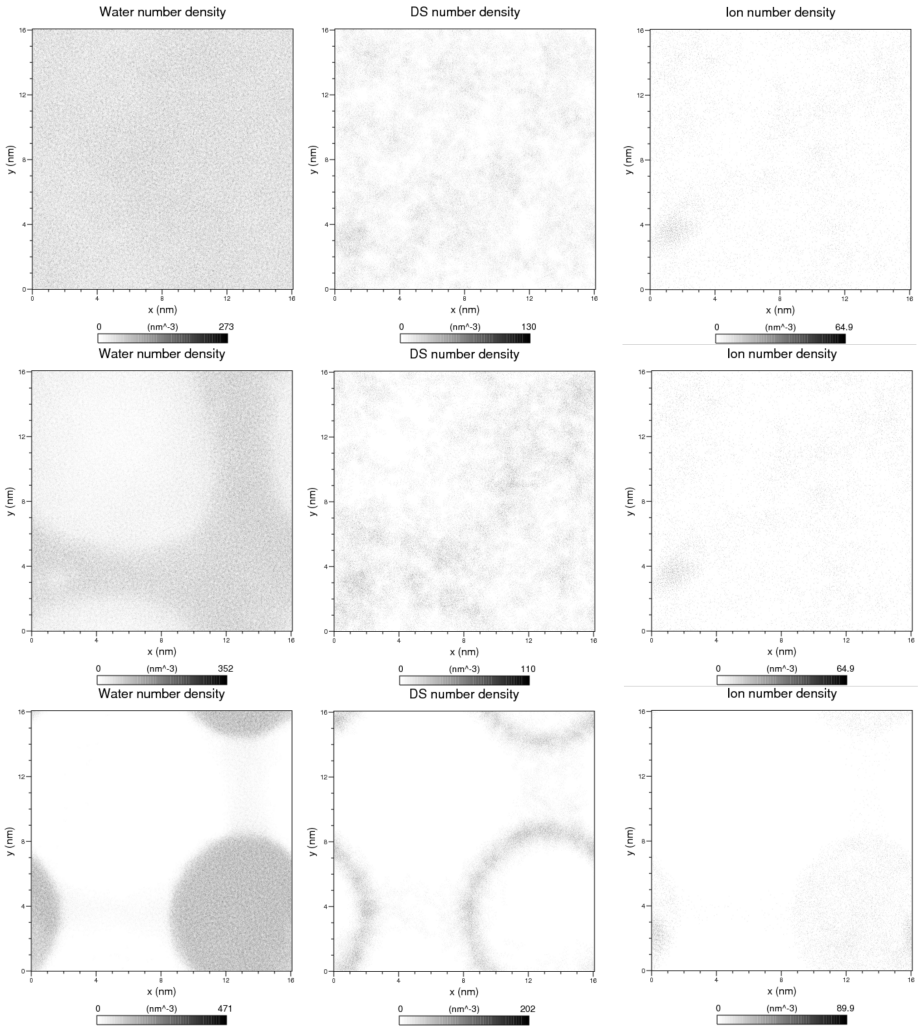


Figure 5.9: Snapshots of MD simulation (Concentrations: SDS=14.976 mol/mm^2, NaCl=0.658 mol/L), showing the 2D density of ions (*left column*), water (*middle column*), and SDS (*right column*), for simulation trajectories during step-wise water molecule removal. *Top row:* # Water molecules (current height) = 3200, oil phases are still separated. *Middle row:* # Water = 3100, system is beginning to show dimpling. *Bottom row:* # Water = 3000, the thin film is broken and the phases merge. Figure lifted from **Paper IV**.

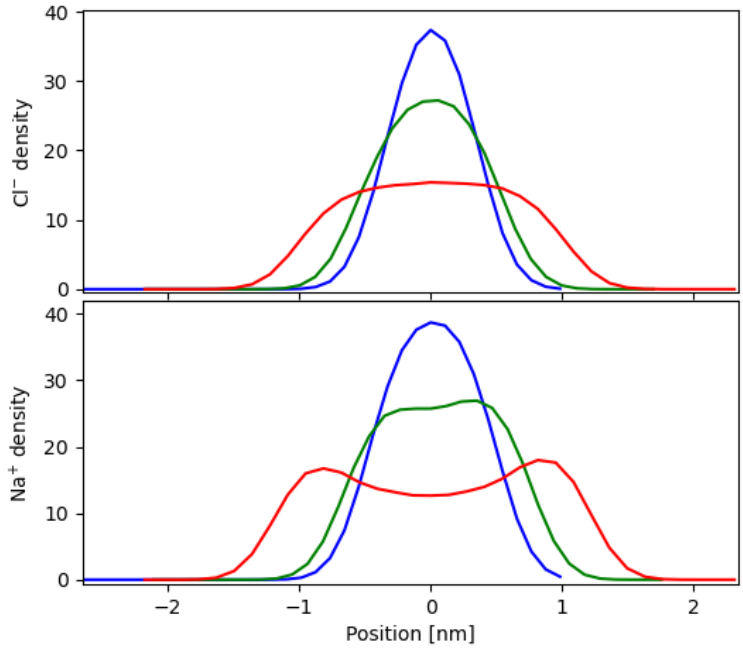


Figure 5.10: Local concentration distribution of ions along the direction normal to the interface at different thin film thickness. *Top:* Cl^- ions and *bottom:* Na^+ ions, ionic concentration 0.293 mol/L, SDS coverage 14.976 mol/mm^2. *Red line:* film height 3,939 nm, *Green line:*, film height 3,159 nm, *Blue line:* critical height of water film, 1.999 nm. Figure lifted from **Paper IV**.

destabilizing the film.

The effects observed in our simulation scheme are consistent with depletion force measurements of high NaCl concentration samples in figure 5.7. The reduced mobility of ions due to ionic surfactant headgroups on interface and in thin film on establishment would counteract the charge difference build up, effectively reducing the measured depletion force once sufficient surfactant and ions are present.

Nevertheless, our simulation scheme is a somewhat poor representation of a realistic system. The observed ion mobility reduction, nucleation rate of ion nanoclusters and subsequent interaction with surfactants requires dedicated simulation campaigns, as our simplified scheme was not set up to include such effects. Thus, a proper rare event simulation scheme would provide a more accurate study of the film breakage, allowing for the estimation a proper probability of film rupture for each film thickness step.

5.5 Improvements to rare event simulations

After its transition to a Python based library (Lervik et al. 2017a), further improvements made to the RETIS algorithm and overall PyRETIS package have been reported in papers **I and II**.

In the following sections, work performed by Dahlen et al. and that of **Paper I** will be briefly presented, before later shifting focus to **Paper II**.

5.5.1 Teaching theory of rare event simulations

With an overall more accessible, and easier-to-learn programming language, the software was opened up to a wider user-base. As mentioned before, the work on PyRETIS 2, its structure and added visualization options, provided a way of creating simplified systems and examples for use in teaching molecular modelling with focus on rare events (Dahlen et al. 2020). Interactive plots were created, based in Python and the Matplotlib library (Hunter 2007a) for front-end control, with PyRETIS MD engine back-end iterating system progression with specified mechanics. These examples illustrated various parts and subsystems of the algorithm of PyRETIS. Some examples include: showing how discrete dynamics were influenced by choice of time-step or equations of motion, or how new trajectories were generated from shooting points in already existing paths, or how to carefully choose order parameters and interface positions along the reaction pathway.

Although the work referred to here was not crucial to the overall project of this thesis, it was a nice addition showing how improvements made to the software itself opened up to other ideas, and helped guide the minds of future potential researchers and users.

5.5.2 PyRETIS 2 version features

The older version of PyRETIS 1 provided built-in interfacing with the external engines of GROMACS (Abraham et al. 2015, Van Der Spoel et al. 2005) and CP2K (Hutter et al. 2014). With the upgrade to v2, the engines were expanded with LAMMPS (Plimpton 1995) and OpenMM (Eastman et al. 2017). Furthermore, existing interfacing with GROMACS was updated, with implementation of MDTraj (McGibbon et al. 2015) to translate external trajectory files of mentioned engines, simplifying the interfacing between PyRETIS and its external engines.

Detailed description of the external engine interfacing, with user guides and examples are available on the PyRETIS webpage (Lervik et al. 2017b).

Two new advanced Monte Carlo moves have been implemented in PyRETIS 2: the stone-skipping and web-throwing moves. Where paths generated in the RETIS al-

gorithm could potentially be very highly correlated due to the use of time-reversal and swapping moves, these new moves effectively reduce correlation between generated paths. Although their execution requires additional computational resources, making them more costly than previous moves, the increase in sampling efficiency will increase. Thus, using them makes accurate estimations of crossing probabilities and rate constants possible with fewer trajectories, still staying within the desired error range (Riccardi et al. 2017).

Other improvements done in PyRETIS 2 is the implementation of a restart feature. In the case of critical errors during simulation, external engine trajectory calculations would stop, writing unfinished trajectories that potentially do not fulfill any criteria of the TIS/RETIS scheme, eq. 2.4. Thus, a feature was added that loads these unfinished trajectories, updates the internal crossing data in PyRETIS, and resumes the unfinished trajectory calculation, before proceeding with the full simulation. Another feature of this restart/load functionality is that completed trajectories can be used as seeds for the initialization of the RETIS scheme, reducing the time spent brute-forcing initial interface crossings.

The improved interfacing capability with GROMACS, plus overall stability improvements and easy restarts in the case of critical errors, smooth the way for further exploring the MD systems of **Paper IV** to a more proper rare event simulation. In a sense, the ground work of initial system set-up and external interfacing is already done, with the implementation of some custom order parameters and RETIS interfaces outstanding.

5.6 Visualization of path sampling data

The final addition to the overall PyRETIS package to be mentioned here is the visualization package introduced with **Paper II**. While briefly touched upon in **Paper I**, the implementation of a separate visualization software solution did merit its own publication, namely **Paper II**. Here, it discusses the components of this software and gives direct examples of how they improves the understanding of mechanics of simulated systems, and presents and example study of the proton transfer in a water trimer system, employing the beforementioned software.

5.6.1 Organizing and condensing simulation results

Path sampling simulations usually generate huge amounts of data, with large systems of n particles in 3 dimensions represented by $3\times$n floating point values. Onto this is an added $3\times$n values of particle momentum (and also for forces if included), further multiplied by a number of MD steps in the order of 10^2 to 10^3 or even beyond. All of this constitutes one single trajectory or path, without mentioning the different path ensembles used in TIS/RETIS, and the overall length of the simula-

tion, i.e. the number of simulation *cycles*.

PyRETIS outputs the system data in folders for each path ensemble. Each folder contains data files where frames of each trajectory are stored as *descriptors* (mostly for system energy) and *collective variables* (CV) representing the atomic positions by functions (distances, angles, etc.). This dimensional reduction of the data produced is significant in reducing the storage requirements during simulations, and for post-processing analysis. Additionally, full trajectory files may still be stored by an external engine.

To aid in the post-processing analysis of simulations, PyVisA was designed to read files of full trajectories and of descriptors and CVs. PyVisA consists of two distict modules, a pre-processing data-parser and condenser, and a visualizer with analysis tools in a custom graphical user interface (GUI). PyVisA parses the PyRETIS file structure for output files, and the pre-prossesing module reads all data and composes a data table of it (order parameter/CVs, energy descriptors). The data is structured and stored according to its respective affiliation (native trajectory, simulation cycle number, path ensemble, sampling status (accepted/rejected, reactive/unreactive) in a pandas(McKinney 2012) dataframe. Any discrepancy due to critical errors (critical crashes) during large simulations are unfortunately not unusual, and are therefore detected by the module, to be removed from the dataframe, as they would present as unfinished paths not fulfilling eq. 2.4. PyVisA does not contribute to the rate calculation, and these removals of very small sections of data influence the path sampling results to a insignificant degree. Furthermore, PyVisA is also able to re-calculate order parameters and collective variables from target frames/trajectories, both in command line and GUI, by adopting the functions required from PyRETIS.

Simulations are usually performed on larger computer clusters, with the analysis of the output performed on a local, less powerful machine. This potentially makes data transfer during analysis a bottleneck to post-processing effectiveness.

By operating the condenser module in command line, pre-processing of data and compiling into a binary file (.pickle or .hdf5) and compressing it (.zip) can be performed on the external machine, simplifying the data transfer and later reading for analysis greatly.

As a small test for the compression efficiency, table 5.2 reports the file sizes of three PyRETIS example file structures of varying size, showing how while structuring short runs increases storage requirements, long simulations benefit greatly from it, increasing with system complexity and file lines. Although the difference between a compressed file system and a compressed binary file is not that large, the

Table 5.2: File dimensions comparison with and without the usage of the PyVisA compressor tool for three different simulation test cases. Table lifted from **Paper II**.

Case	Test 1	Water Trimer	Test2
Number of file lines	20 702	2 505 275	18 722 085
Base filesystem	708 KB	263.3 MB	1.4 GB
Zip of base filesystem (no PyVisA)	261 KB	54.6 MB	399 MB
Pickle of dataset	407 KB	69.4 MB	317 MB
Pickle.zip dataset	257 KB	55.1 MB	213 MB
HDF5 of dataset	2.5 MB	71.6 MB	319 MB
HDF5.zip dataset	253 KB	55.1 MB	213 MB

effectiveness of analysis still increases as the binary data structures have already been checked for simulation errors, and are faster for the visualizer GUI to load.

The compiling of binary data structures and further compressing eases the process of data transfer to a local machine for analysis, and also makes the analysis possible to perform while simulations are still running. This ability could reveal problems during initial sampling like meta-stable states or improper interface positioning.

5.6.2 Visualizing path sampling data

The aforementioned visualizer GUI is a module that provides users with access to various features of visualizing the trajectory data produced by a PyRETIS simulation. The GUI consists of a matplotlib.pyplot (Hunter 2007b) canvas as plotting base, within a PyQt5 window (Summerfield 2007). Within the GUI, the different operations are contained in respective tabs, or panels: data-set selection, plotting layout, trajectory visualization, and analysis.

Figure 5.11 shows the appearance of the GUIs' panels. The data-selection panel (**A**) allows users to select subsets of data to use for visualization, via check-boxes and drop-down menus in the tab. The selection options are: plot type, axis descriptors, ensemble folder, cycle range selection, path generating moves, path status (Accepted, Rejected), and path types (reactive/unreactive). Additional check-boxes are also available, with a more specific performance on the data chosen by previous selection:

- 'Reg. line': calculates and displays the equation obtained by the linear regression of 2D data plot.
- 'Data shift': attempts to transpose the center of observation and apply periodic boundary condition, for periodic collective variables (angles, dihedrals).

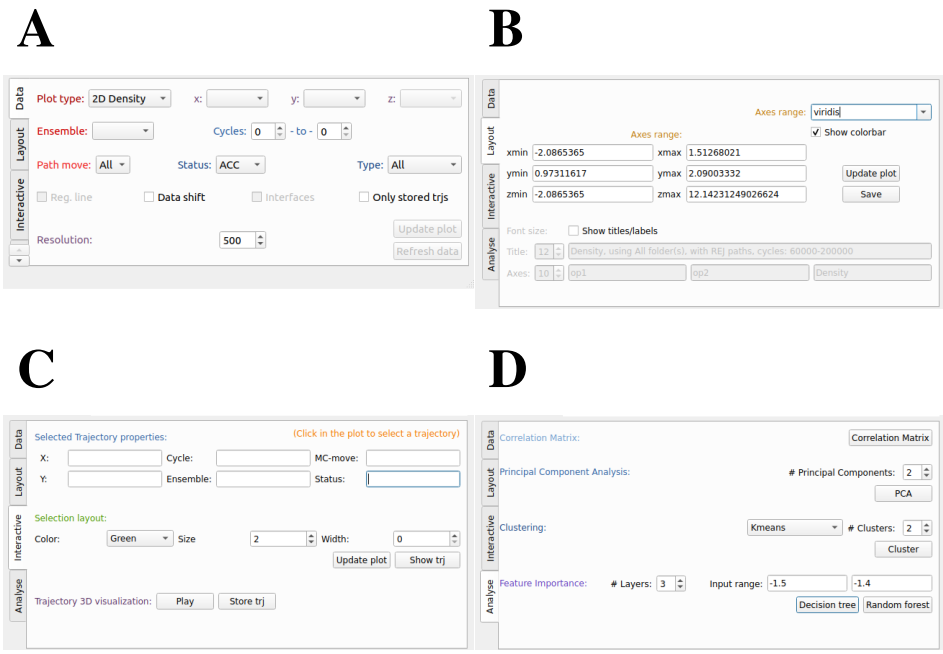


Figure 5.11: Top left (**A**): Visualizer GUI settings tabs for data loading; Top right (**B**): Visualizer GUI settings tabs for plot layout; Bottom left (**C**): Visualizer GUI settings tabs for trajectory visualization; Bottom right (**D**): Visualizer GUI settings tabs for analysis. Figure lifted from **Paper II**.

- 'Interfaces': displays the ensemble interfaces, as lines or planes in 2D and 3D plots, respectively.
- 'only stored trjs': displays only paths for which the full trajectory is stored in memory, following PyRETIS file structure for storage.

Finally, two buttons are also included in the panel: 'Update plot' and 'Refresh data'. 'Update plot' will generate a new visualization (plot) based on the selection options chosen, 'Refresh data' will only re-import data from source files, and is designed to visualize the output on the same machine as simulations are in progress.

The plotting layout (figure 5.11**B**) selects cosmetic changes to the plot, updating the canvas with colormap, titles and axis labels, and includes a rudimentary zoom-feature by axis limit selection. Additionally, in contains one button for updating with new modifications that do not require any re-evalutation of data selection, and another one for saving the plotted figure in '.png' format.

The third panel (figure 5.11**C**) focuses on single trajectories that compose the chosen dataset. The user may select a certain trajectory by clicking the interactive plot and visualize its properties (ensemble, status, move generated) in the panel. Should the trajectory also be stored in full detail (as mentioned previously) a button 'Play' executes a separate Avogadro(Hanwell et al. 2012) application fed with the selected trajectory for full atomistic visualization. Alternatively, 'Save Trj' button stores the selected trajectory to a .pbd file for later visualization. This is useful since PyRETIS shooting moves generate two trajectories, forward and backwards in time, and this way they are combined into one.

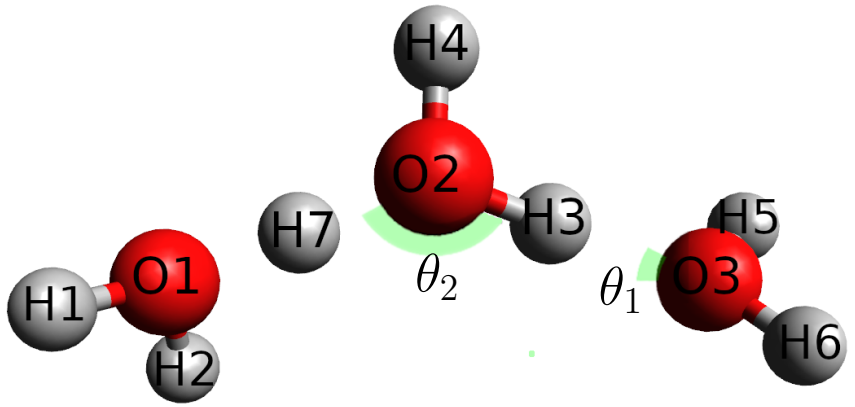
The final panel (figure 5.11**D**), analysis, contains some machine learning-based statistical methods of the python package *sklearn* (Buitinck et al. 2013, Pedregosa et al. 2011): correlation matrix, clustering, principal component analysis (PCA)(Wold et al. 1987) and decision trees (Ho 1995a).

Further, minor features of the PyVisA GUI window are detailed in **Paper II**, including a step–by–step guide for using the tool. Additional features within 'File menu' of the GUI include, but are not limited to: file handling, re-calculating, -compiling and -compressing, compression levels and file extension, and the mentioned re-calculation of collective variables from data structure. The re-calculation could also prove quite useful as it permits the addition of new collective variables after simulation is completed.

5.6.3 Visualizing results of a simple path sampling study

Path sampling study of a protonated water trimer system, and the proton transfer reaction described by Geissler et al. (1999; 2000), provided decent system size for testing PyVisA and an easily understandable reaction pathway. Detailed description of the system, the order parameters and collective variables, interface positioning, and general simulation specifications are all listed in the materials section of **Paper II**. Atomic system, labeled, with key angles for order parameter and the two distances serving as additional collective variables are all shown in figure 5.12:

Figure 5.14 shows plots generated by PyVisa using the data from our water trimer system. Top plot (**A**) shows a frame density plot for reactive trajectories in the $[14^+]$ ensemble, with x/y dimensions as function of the main order parameter, Θ, and CV for the distance between the travelling hydrogen ion and its initial neighbour oxygen, OP_2, respectively. Density in 5.14(**A**) signifies the time the system spent in certain regions, quantified by the colorbar next to plot, counting the number of frames with the values of Θ and OP_2. Bottom plot **B** shows potential energy contour surface plotted against the same data set of **A**.



Collective variable	Definition	Unit
(Order parameter) 1.	$\theta_1 - \theta_2$	rad
2.	O_1-H_7	Å
3.	O_2-H_7	Å

Figure 5.12: The protonated water-trimer system and a description of the collective variables. Figure lifted from **Paper II**.

While top figure shows how reactive trajectories keep an even O–H distance of $\sim$ 1.3 Åup to and past an area of high density at Θ=0 and OP_2=1.3, the bottom figures show how this is due to this region being a saddle point in potential energy required to pass. Both figures later show an extension of the O–H bond, signifying that after the change of angles, the change of center water molecules, the hydrogen ion leaves the old center in favor of the new, and the system is again symmetrical. So the results follow the proposed sequence of events required for proton transfer, illustrated by simulation frames in figure 5.13, a repositioning of water molecules that shifts the center of the trimer.

The implementation of PyVisA to the PyRETIS software is quite recent, yet its tools have already contributed to other work (Daub et al. 2020, Riccardi et al. 2020a) with its easy plotting functionality. Specific examples, where plot interactivity linked with external visualization of complete paths, are not that easy to illustrate; the option is, however, believed to be a nice feature for the convenience of the user, regardless.

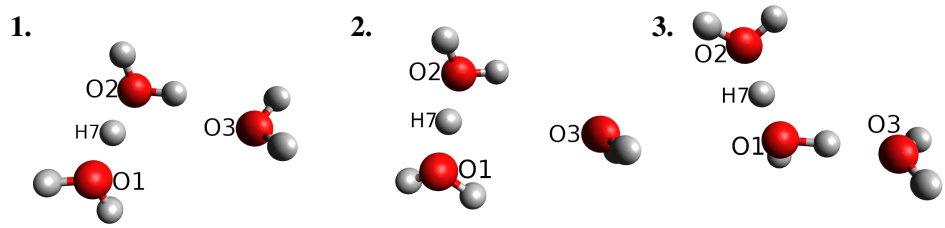


Figure 5.13: Illustration of a reactive pathway generated during the simulation, with snapshots taken along the way. **1.** shows the system at the initial state, $\Theta = -1.5$. At snapshot **2.**, the system is around the saddle point of the transition, with a value of $\Theta \approx 0.0$, and the transferring proton equidistant from the two neighbouring oxygen atoms. **3.** is where the system has reached the defined end state, $\Theta = 1.5$. The length of the path between states A (1.), and B (3.), was ~ 500 time steps long ($\tilde{2}50$ fs). Figure lifted from **Paper II**.

5.6.4 Determining descriptor influence

During the simulation, the system was regularly visiting two regions around OP_2 = 1.1 Åand 1.25 Å. Using PyVisA, a density plot of all ensembles was created, showing the potential energy of the system against O–H distance OP_2. This plot, shown in figure 5.15, consistently shows that the proton transfer occurs with a 1.25 ÅO–H distance.

Whether this O–H distance correlates with the reaction of proton transfer, a decision tree (*chemistree*) (Roet et al. 2021) of classifiers was employed in order to quantify the relevance of different factors of the simulation.

In figure 5.16, the generated decision tree is visualized. Top figure (**A**) shows the three levels of the tree, and bottom (**B**) shows the importance measure of the tree obtained from a random forest of decision trees(Ho 1995b).

Simulation frames of the transition region, in the range of Θ={-0.1, 0.1}, were selected and fed to the classifiers. Each descriptor (energy, CV, etc.) had its importance determined in figure 5.16**B**, which showed that the total system energy in fact was the most important factor for discriminating between reactive and unreactive trajectories reaching the transition region. The decision trees indicated that the highest probability for any frame within the given to belong in a reactive path, occurred when total energy $\leq$ 3304.793 kcal/mol and kinetic energy $\geq$ 5.988 kcal/mol. Furthermore, the random forest also revealed that distances OP_2 and OP_3 were mostly irrelevant to the path reactivity, which agrees with the results from figure 5.14.

With the relatively recent implementation in **Paper II** and its detailed method development (Roet et al. 2021), the presented results here are limited to a quite simple

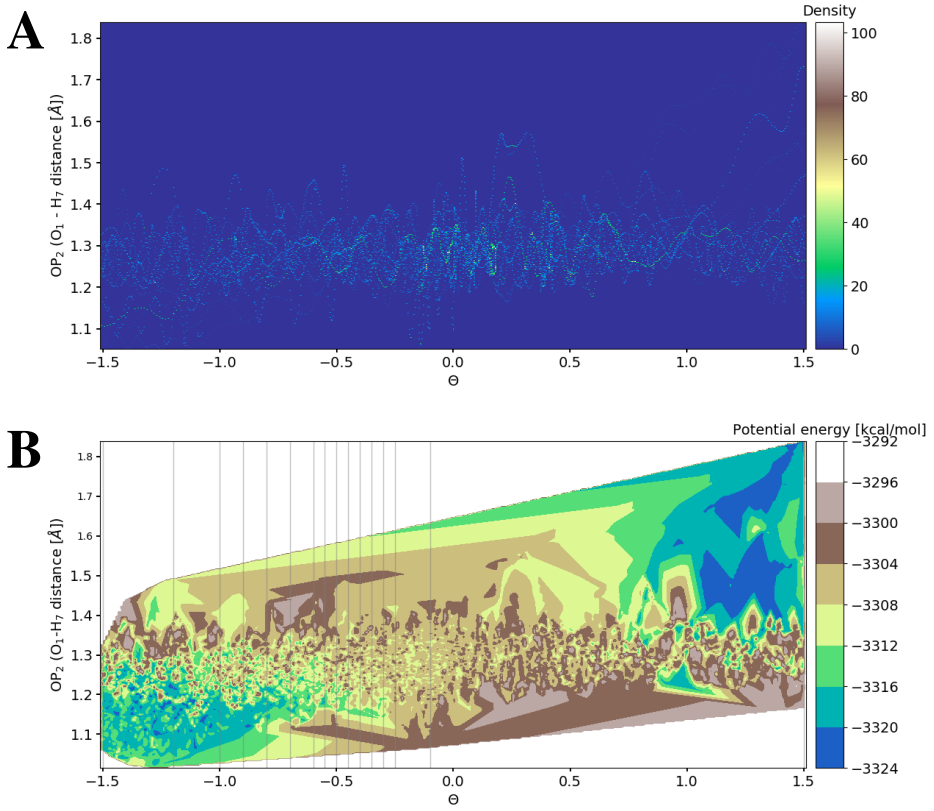


Figure 5.14: **A:** Frame density of reactive trajectories plotted along Θ and OP_2 (O_1–H_7 distance). **B:** Potential energy surface in a scale from blue (lower energy) to brawn (higher energy). Vertical lines indicate interface planes for ensembles. Both plots are composed by trajectories collected in the $[14^+]$ ensemble. Figure lifted from **Paper II**

system, and do not appear very groundbreaking themselves. Still, decision trees may yet contribute more mechanistic insight into other systems studied, in due time.

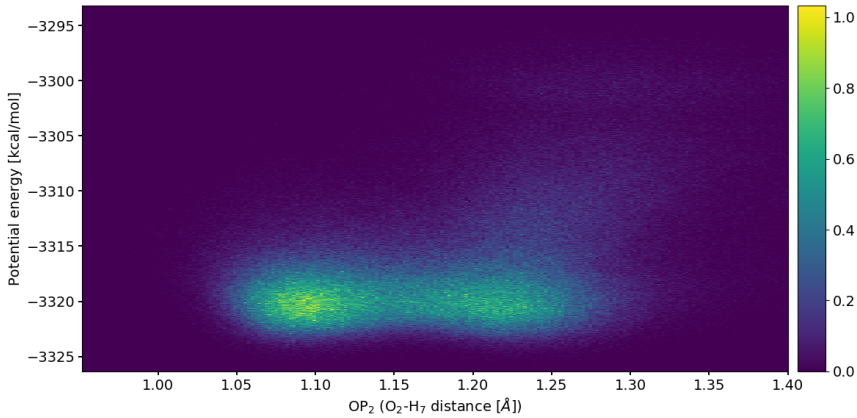


Figure 5.15: Figure reporting, for all ensembles, the density plot of the potential energy as a function of OP_2 (which is mostly symmetrical to OP_3). Figure lifted from **Paper II**.

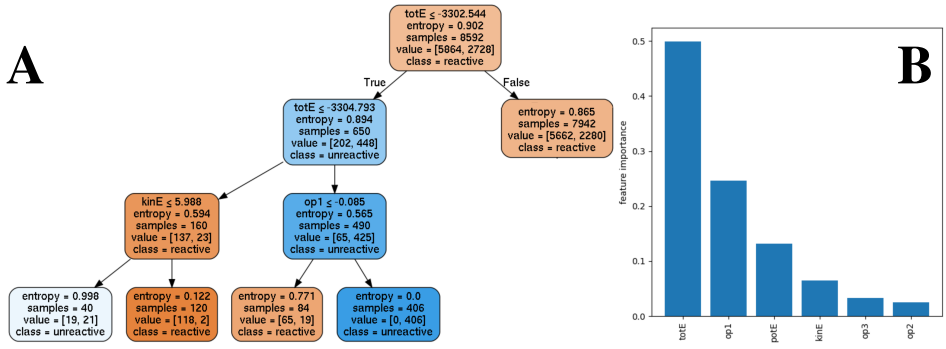


Figure 5.16: Classifiers: panel A reports a decision tree while panel B lists the descriptors (features) with their relative importance in determining path reactivity. The histogram is obtained by a random forest of decision trees. Figure lifted from **Paper II**.

Chapter 6

Concluding remarks

Oil–in–water emulsions for use in optical tweezer-based studies were prepared and measured. Pristine emulsion droplets with ionic presence, and stabilized droplets were created, using dodecane as a model oil and SDS as surfactant molecule. This similarity in structure between oil and surfactant eased the process of setting up and modelling a thin film breakage with MD simulations. For the production of emulsions within the size range of optical trapping, several techniques of homogenization were used and evaluated based on their overall effectiveness. The well established method of microfluidic droplet generation proved to be too unreliable for the smaller droplets desired for trapping. Both lack of surfactant to stabilize the freshly split droplets at nozzle, causing immediate coalescence, and overall tiny sample amount produced, made this method too costly to attempt even when stabilizing agents were introduced. The least work-intensive approach of homogenization showed that the lower time consumption outweighed any concerns arising from the increased distribution of droplet sizes.

Whilst testing size ranges on the optical tweezers, limitations to the trapping and manipulation were quantified. From the high approach velocity of droplets, a significant increase in drag force was observed, to the point where force output data had to be corrected for trap drift. With both increased velocity and size, risk of trap escape also increased. Identifying proper contact points of droplets were subject to errors related to the diffraction limit, making the estimation of traps displacement challenging, and making droplet pairs prone to flocc during retracted segments, halting current measurements.

Using retract–extend measurement schemes, the depletion force and work of separation between droplet pairs was determined for increased bulk ionic strengths,

approach velocity, droplet diameters, surfactant coverage and interplay between ionic surfactants and solvated ions in bulk and in thin film. Separation work was shown to increase with bulk ionic strength, droplet size and approach velocity of droplets. Droplets stabilized with SDS were shown to have similar behavior to depletion force as pristine ones, indicating that ionic headgroups of surfactant at interface did not impede the formation of a charge difference when retracting droplets. However, the interplay between solvated ions and surfactant molecules plateaued depletion force for higher SDS surface coverage. MD simulations of a model system revealed as a possible reason for this a reduction in ion mobility during early phases of thin film establishment. This effect would keep the local ionic strength in thin film higher by screening the ionic headgroups during retraction, resulting in a lower charge difference between thin film and bulk.

Coalescence times were similarly examined, finding an inverse relation between increased approach velocity and coalescence (film drainage) time. Large variations in data sets limit firmer conclusions, and higher influence on coalescence times was seen in relation to increased droplet sizes, rather than approach velocity. A theoretical lower bound for coalescence times was not observed, likely due to the limiting control factor of the tweezers and drift from drag force.

Both studies on coalescence, pristine and stabilized, showed a strong stochastic nature of the coalescence event, leading to the broad range of results from the parameters studied. The addition of low concentration of ionic surfactant headgroups to droplet interfaces did not change the expected effects of solvated ions to the coalescence times that was observed for pristine droplets. The stabilizing effect of SDS also became dominating at the highest coverage tested, combined with lower ionic concentrations. There was a lack of coalescence observed for intermediate SDS sample series, however, and MD simulations of film breakage provided more information.

Estimations done on the MD systems showed a correlation between amount of surfactant and thin film height, but a slight reduction in the water amount of the film. Furthermore, ionic strength in the film also tend to increase the critical height, although to a lesser extent than surfactant coverage. Plots of ionic density during the process of water reduction from systems reveal that ions do contribute to the stability, presenting a double layer on each interface that must be neutralized prior to any film breakage. An effect of the step-wise reduction was also the observed formation of small ionic clusters, or nano-crystals, as the local concentration increased for each removal of water molecules. These crystals appear to influence the thin film, and their presence promote coalescence in systems of intermediate SDS coverage.

With all observations, both experimental and simulated, two things are clear. Firstly, the power of optical tweezers on macroemulsions, their control and precision, is unparalleled. In the size range 1–10 μm there is a lack of studies on droplet interaction and coalescence, as most techniques are limited to the use of larger emulsion droplets, while this range remains mostly untouched. The tweezers are, however, not perfect tools, i.e. they embody harsh restrictions to measurements that materialize when pushing on parameters of drag and diffraction limit.

In order to further expand the knowledge of the underlying mechanics of film breakage, more dedicated rare event simulations should be performed. The relatively simple method used provided us with some interesting results, but the stochastic process of film rupture is not well explored with only deterministic simulations. Additionally, the behavior of surfactant molecules and ions in thin film require longer simulation runs, providing more statistics to sample the reaction pathway more accurately, and to calculate a rate/probability of film breakage at different heights. Thus, the current simulation scheme has paved the way for more sophisticated methods, using the already set-up GROMACS systems with PyRETIS.

Improvements to the PyRETIS algorithm, such as with new MC moves, better integration of external engines and more analysis tools in PyVisA, all would be of great benefit to the proposed study. The RETIS algorithm have been shown to more efficiently explore possible pathways of reactions, converging much faster on the optimal solution A sample study of proton transfer in a protonated water trimer was used as a basis for explaining the features of PyVisA, although it has already shown great potential in facilitating analysis of simulations in other work. The ease of use allows users with less experience in the field to create plots and figures to publication standards without investing much time in manual coding. Additionally, its self-standing structure opens up for usage with other software solutions, even though it was developed for PyRETIS outputs.

With the large oil–water/ion/surfactant–oil system, more collective variables monitoring the mobility, coverage and cluster formation described previously, could be introduced. These, and other descriptors could then be crosschecked for correlating behavior with PyVisA's tools, both visualization by plotting and with the new decision trees included. Given the computational resources, external simulations can be run with visualization performed locally, the compressed binary file being a quick way of transferring and accessing the data for visualization. With this, storage of complete trajectories could also be feasible, and would be directly accessible from the GUI.

References

Aarøen, O., Kiær, H., and Riccardi, E. (2020). Pyvisa: Visualization and analysis of path sampling trajectories. *Journal of Computational Chemistry*, 42(6):Pages 435–446.

Aarøen, O., Riccardi, E., and Sletmoen, M. (2021a). Exploring the effects of approach velocity on depletion force and coalescence in oil-in-water emulsions. *RSC Advances*, 11(15):8730–8740.

Aarøen, O., Riccardi, E., Sletmoen, M., and van Erp, T. S. (2021b). Multidisciplinary study of effects on thin film breakage in oil–in–water emulsions. *Preprint*.

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25.

Anderegg, L., Cheuk, L. W., Bao, Y., Burchesky, S., Ketterle, W., Ni, K.-K., and Doyle, J. M. (2019). An optical tweezer array of ultracold molecules. *Science*, 365(6458):1156–1158.

Ashkin, A. (1970). Acceleration and trapping of particles by radiation pressure. *Physical review letters*, 24(4):156.

Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E., and Chu, S. (1986). Observation of a single-beam gradient force optical trap for dielectric particles. *Optics letters*, 11(5):288–290.

Bauer, W.-A. C., Kotar, J., Cicuta, P., Woodward, R. T., Weaver, J. V., and Huck, W. T. (2011). Microfluidic production of monodisperse functional o/w droplets and study of their reversible ph dependent aggregation behavior. *Soft Matter*, 7(9):4214–4220.

Bolhuis, P. G., Chandler, D., Dellago, C., and Geissler, P. L. (2002). Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry*, 53(1):291–318.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

Bzdek, B. R., Collard, L., Sprittles, J. E., Hudson, A. J., and Reid, J. P. (2016). Dynamic measurements and simulations of airborne picolitre-droplet coalescence in holographic optical tweezers. *The Journal of chemical physics*, 145(5):054502.

Chen, A., Li, S.-W., Jing, D., and Xu, J.-H. (2019). Interactions between colliding oil drops coated with non-ionic surfactant determined using optical tweezers. *Chemical Engineering Science*, 193:276–281.

Chen, A., Li, S.-W., Sang, F.-N., Zeng, H.-B., and Xu, J.-H. (2018). Interactions between micro-scale oil droplets in aqueous surfactant solution determined using optical tweezers. *Journal of colloid and interface science*, 532:128–135.

Chen, A., Wang, F., Zhou, Y., and Xu, J.-h. (2020). In situ measurements of interactions between switchable surface-active colloid particles using optical tweezers. *Langmuir*, 36(17):4664–4670.

Chime, S. A. and Onyishi, I. V. (2013). Lipid-based drug delivery systems (ldds): Recent advances and applications of lipids in drug delivery. *African Journal of Pharmacy and Pharmacology*, 7(48):3034–3059.

Dahlen, O., Lervik, A., Aarøen, O., Cabriolu, R., Lyng, R., and van Erp, T. S. (2020). Teaching complex molecular simulation algorithms: Using self-evaluation to tailor web-based exercises at an individual level. *Computer Applications in Engineering Education*, 28(4):779–791.

Daub, C. D., Riccardi, E., Hänninen, V., and Halonen, L. (2020). Path sampling for atmospheric reactions: formic acid catalysed conversion of so3+ h2o to h2so4. *PeerJ Physical Chemistry*, 2:e7.

Dellago, C., Bolhuis, P., and Geissler, P. L. (2002). Transition path sampling. *Advances in chemical physics*, 123:1–78.

Deng, Y., Bechhoefer, J., and Forde, N. R. (2007). Brownian motion in a modulated optical trap. *Journal of Optics A: Pure and Applied Optics*, 9(8):S256.

Dudek, M., Fernandes, D., Herø, E. H., and Øye, G. (2020). Microfluidic method for determining drop-drop coalescence and contact times in flow. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 586:124265.

Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., et al. (2017). Openmm 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*, 13(7):e1005659.

Gebauer, F., Villwock, J., Kraume, M., and Bart, H.-J. (2016). Detailed analysis of single drop coalescence—influence of ions on film drainage and coalescence time. *Chemical Engineering Research and Design*, 115:282–291.

Geissler, P. L., Dellago, C., and Chandler, D. (1999). Chemical dynamics of the protonated water trimer analyzed by transition path sampling. *Physical Chemistry Chemical Physics*, 1(6):1317–1322.

Geissler, P. L., Dellago, C., Chandler, D., Hutter, J., and Parrinello, M. (2000). Ab initio analysis of proton transfer dynamics in $[(H_2O)_3H]^+$: comparison of density functional theory and wavefunction-based methods. *Chemical Physics Letters*, 321(3):225–230.

Goldberg, O. (1980). Köhler illumination. *The microscope*, 27(34):15–21.

Goodarzi, F. and Zendehboudi, S. (2019). A comprehensive review on emulsions and emulsion stability in chemical and energy industries. *The Canadian Journal of Chemical Engineering*, 97(1):281–309.

Guo, X., Rong, Z., and Ying, X. (2006). Calculation of hydrophile–lipophile balance for polyethoxylated surfactants by group contribution method. *Journal of Colloid and Interface Science*, 298(1):441–450.

Haddrell, A. E., Miles, R. E., Bzdek, B. R., Reid, J. P., Hopkins, R. J., and Walker, J. S. (2017). Coalescence sampling and analysis of aerosols using aerosol optical tweezers. *Analytical chemistry*, 89(4):2345–2352.

Hadjialirezaei, S., Picco, G., Beatson, R., Burchell, J., Stokke, B. T., and Sletmoen, M. (2017). Interactions between the breast cancer-associated muc1 mucins and c-type lectin characterized by optical tweezers. *PLoS One*, 12(4):e0175323.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., and Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4(1):17.

Hargreaves, A. L., Gregson, F., Kirby, A. K., Engelskirchen, S., and Bain, C. D. (2015). Microemulsion droplets in optical traps. *Journal of Molecular Liquids*, 210:9–19.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Hiemenz, P. C. and Rajagopalan, R. (2016). *Principles of Colloid and Surface Chemistry, revised and expanded*. CRC press.

Ho, T. K. (1995a). Proceedings of the 3rd international conference on document analysis and recognition. *Random Decision Forests*, pages 278–282.

Ho, T. K. (1995b). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE.

Huang, R., Chavez, I., Taute, K. M., Lukić, B., Jeney, S., Raizen, M. G., and Florin, E.-L. (2011). Direct observation of the full transition from ballistic to diffusive brownian motion in a liquid. *Nature Physics*, 7(7):576–580.

Hunter, J. D. (2007a). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95.

Hunter, J. D. (2007b). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95.

Hutter, J., Iannuzzi, M., Schiffmann, F., and VandeVondele, J. (2014). cp2k: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(1):15–25.

Jang, S. S. and Goddard, W. A. (2006). Structures and properties of newton black films characterized using molecular dynamics simulations. *The Journal of Physical Chemistry B*, 110(15):7992–8001.

Ji, S. and Walz, J. Y. (2015). Depletion forces and flocculation with surfactants, polymers and particles—synergistic effects. *Current Opinion in Colloid & Interface Science*, 20(1):39–45.

JPK/Bruker (2021). Characterizing quantiative measurements of force and displacement with optical tweezers on the Nanotracker™. https://www.jpk.com/app-technotes-img/Optical-Tweezers/pdf/jpk-nt-tech-calibration.1.pdf. Accessed: 2021-06-06.

Kamal, M. S. (2016). A review of gemini surfactants: potential application in enhanced oil recovery. *Journal of Surfactants and Detergents*, 19(2):223–236.

Kamp, J. and Kraume, M. (2016). From single drop coalescence to droplet swarms–scale-up considering the influence of collision velocity and drop size on coalescence probability. *Chemical Engineering Science*, 156:162–177.

Khan, B. A., Akhtar, N., Khan, H. M. S., Waseem, K., Mahmood, T., Rasul, A., Iqbal, M., and Khan, H. (2011). Basics of pharmaceutical emulsions: A review. *African Journal of Pharmacy and Pharmacology*, 5(25):2715–2725.

Kirkpatrick, R. and Lockett, M. (1974). The influence of approach velocity on bubble coalescence. *Chemical Engineering Science*, 29(12):2363–2373.

Klaseboer, E., Chevaillier, J. P., Gourdon, C., and Masbernat, O. (2000). Film drainage between colliding drops at constant approach velocity: experiments and modeling. *Journal of colloid and interface science*, 229(1):274–285.

Kovalchuk, K., Riccardi, E., and Grimes, B. A. (2014). Multiscale modeling of mass transfer and adsorption in liquid–liquid dispersions. 1. molecular dynamics simulations and interfacial tension prediction for a mixed monolayer of mono-and tetracarboxylic acids. *Industrial & Engineering Chemistry Research*, 53(29):11691–11703.

Kühne, T. D., Iannuzzi, M., Del Ben, M., Rybkin, V. V., Seewald, P., Stein, F., Laino, T., Khaliullin, R. Z., Schütt, O., Schiffmann, F., et al. (2020). Cp2k: An electronic structure and molecular dynamics software package-quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics*, 152(19):194103.

Kundu, P. and Mishra, I. M. (2018). Treatment and reclamation of hydrocarbon-bearing oily wastewater as a hazardous pollutant by different processes and technologies: a state-of-the-art review. *Reviews in Chemical Engineering*, 35(1):73–108.

Langevin, D. (2019). Coalescence in foams and emulsions: Similarities and differences. *Current Opinion in Colloid & Interface Science*, 44:23–31.

Lervik, A., Riccardi, E., and van Erp, T. S. (2017a). PyRETIS: A well-done, medium-sized python library for rare events. *Journal of Computational Chemistry*, 38(28):2439–2451.

Lervik, A., Riccardi, E., and van Erp, T. S. (2017b). The PyRETIS manual. http://www.pyretis.org. Accessed: 07-04-2017.

Lyklema, J., Van Leeuwen, H., and Minor, M. (1999). Dlvo-theory, a dynamic re-interpretation. *Advances in colloid and interface science*, 83(1-3):33–69.

Mao, Y., Cates, M., and Lekkerkerker, H. (1995). Depletion force in colloidal systems. *Physica A: Statistical Mechanics and its Applications*, 222(1-4):10–24.

Mao, Y., Cates, M., and Lekkerkerker, H. (1997). Theory of the depletion force due to rodlike polymers. *The Journal of chemical physics*, 106(9):3721–3729.

McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., Schwantes, C. R., Wang, L.-P., Lane, T. J., and Pande, V. S. (2015). Mdtraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophysical journal*, 109(8):1528–1532.

McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. " O'Reilly Media, Inc.".

Mitsunobu, M., Kobayashi, S., Takeyasu, N., and Kaneta, T. (2017). Temperature-induced coalescence of droplets manipulated by optical trapping in an oil-in-water emulsion. *Analytical Sciences*, 33(6):709–713.

Moffitt, J. R., Chemla, Y. R., Smith, S. B., and Bustamante, C. (2008). Recent advances in optical tweezers. *Annual review of biochemistry*, 77:205–228.

Negin, C., Ali, S., and Xie, Q. (2017). Most common surfactants employed in chemical enhanced oil recovery. *Petroleum*, 3(2):197–211.

Neuman, K. C. and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature methods*, 5(6):491–505.

Nilsen-Nygaard, J., Sletmoen, M., and Draget, K. I. (2014). Stability and interaction forces of oil-in-water emulsions as observed by optical tweezers–a proof-of-concept study. *RSC Advances*, 4(94):52220–52229.

Otazo, M. R., Ward, R., Gillies, G., Osborne, R. S., Golding, M., and Williams, M. A. (2019). Aggregation and coalescence of partially crystalline emulsion drops investigated using optical tweezers. *Soft matter*, 15(31):6383–6391.

Ozan, S. C. and Jakobsen, H. A. (2019a). On the effect of the approach velocity on the coalescence of fluid particles. *International Journal of Multiphase Flow*, 119:223–236.

Ozan, S. C. and Jakobsen, H. A. (2019b). On the role of the surface rheology in film drainage between fluid particles. *International Journal of Multiphase Flow*, 120:103103.

Ozan, S. C. and Jakobsen, H. A. (2020). Effect of surface viscoelasticity on the film drainage and the interfacial mobility. *International Journal of Multiphase Flow*, 130:103377.

Pan, L., Chen, Y., Chen, D., Dong, Y., Zhang, Z., and Long, Y. (2018). Oil removal in tight-emulsified petroleum waste water by flocculation. In *IOP Conf. Ser. Mater. Sci. Eng*, volume 392, page 042005.

Papaleo, E., Riccardi, E., Vreede, J., van Erp, T., Cabriolu, R., and Lervik, A. (2019). Transition path sampling simulations via PyRETIS: Theory and application of rare events methods to compute transition and reaction rates. https://www.lorentzcenter.nl/transition-path-sampling-simulations-via-pyretis-theory-and-application-of-rare-events-methods-to-compute-transition-and-reaction-rates.html.

Pasquali, R. C., Sacco, N., and Bregni, C. (2009). The studies on hydrophilic-lipophilic balance (hlb): Sixty years after william c. griffin's pioneer work (1949-2009). *Lat Am J Pharm*, 28(2):313–317.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pettit, R. M., Ge, W., Kumar, P., Luntz-Martin, D. R., Schultz, J. T., Neukirch, L. P., Bhattacharya, M., and Vamivakas, A. N. (2019). An optical tweezer phonon laser. *Nature Photonics*, 13(6):402–405.

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19.

Polimeno, P., Magazzu, A., Iati, M. A., Patti, F., Saija, R., Boschi, C. D. E., Donato, M. G., Gucciardi, P. G., Jones, P. H., Volpe, G., et al. (2018). Optical tweezers and their applications. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 218:131–150.

Power, R., Reid, J., Anand, S., McGloin, D., Almohamedi, A., Mistry, N., and Hudson, A. (2012). Observation of the binary coalescence and equilibration of micrometer-sized droplets of aqueous aerosol in a single-beam gradient-force optical trap. *The Journal of Physical Chemistry A*, 116(35):8873–8884.

Pralle, A., Prummer, M., Florin, E.-L., Stelzer, E., and Hörber, J. (1999). Three-dimensional high-resolution particle tracking for optical tweezers by forward scattered light. *Microscopy research and technique*, 44(5):378–386.

Rekvig, L. and Frenkel, D. (2007). Molecular simulations of droplet coalescence in oil/water/surfactant systems. *The Journal of chemical physics*, 127(13):134701.

Riccardi, E., Dahlen, O., and van Erp, T. S. (2017). Fast decorrelating Monte Carlo moves for efficient path sampling. *The journal of physical chemistry letters*, 8(18):4456–4460.

Riccardi, E., Krämer, A., van Erp, T. S., and Ghysels, A. (2020a). Permeation rates of oxygen through a lipid bilayer using replica exchange transition interface sampling. *The Journal of Physical Chemistry B*.

Riccardi, E., Lervik, A., Roet, S., Aarøen, O., and van Erp, T. S. (2020b). PyRETIS 2: An improbability drive for rare events. *Journal of Computational Chemistry*, 41(4):370–377.

Riccardi, E. and Tichelkamp, T. (2019). Calcium ion effects on the water/oil interface in the presence of anionic surfactants. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 573:246–254.

Rios-Reyes, L. and Scheeres, D. (2005). Generalized model for solar sails. *Journal of Spacecraft and Rockets*, 42(1):182–185.

Roet, S., Daub, C. D., and Riccardi, E. (2021). Chemistrees: data driven identification of reaction pathways via machine learning.

Sengupta, P. and Chatterjee, B. (2017). Potential and future scope of nanoemulgel formulation for topical delivery of lipophilic drugs. *International journal of pharmaceutics*, 526(1-2):353–365.

Summerfield, M. (2007). *Rapid GUI programming with Python and Qt: the definitive guide to PyQt programming*. Pearson Education.

Sun, X., Zhang, Y., Chen, G., and Gai, Z. (2017). Application of nanoparticles in enhanced oil recovery: a critical review of recent progress. *Energies*, 10(3):345.

Swenson, D. W., Prinz, J.-H., Noe, F., Chodera, J. D., and Bolhuis, P. G. (2018). OpenPathSampling: A Python Framework for Path Sampling Simulations. 1. Basics. *Journal of chemical theory and computation*, 15(2):813–836.

Thammawongsa, N., Zainol, F. D., Mitatha, S., Ali, J., and Yupapin, P. P. (2012). Nanorobot controlled by optical tweezer spin for microsurgical use. *IEEE transactions on nanotechnology*, 12(1):29–34.

Tulpar, A., Tilton, R. D., and Walz, J. Y. (2007). Synergistic effects of polymers and surfactants on depletion forces. *Langmuir*, 23(8):4351–4357.

Umar, A. A., Saaid, I. B. M., Sulaimon, A. A., and Pilus, R. B. M. (2018). A review of petroleum emulsions and recent progress on water-in-crude oil emulsions stabilized by natural surfactants and solids. *Journal of Petroleum Science and Engineering*, 165:673–690.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005). Gromacs: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–1718.

van Erp, T. S. (2006). Efficiency analysis of reaction rate calculation methods using analytical models I: The two-dimensional sharp barrier. *The Journal of chemical physics*, 125(17):174106.

van Erp, T. S. (2007). Reaction rate calculation by parallel path swapping. *Physical review letters*, 98(26):268301.

van Erp, T. S. (2008). Efficient path sampling on multiple reaction channels. *Computer physics communications*, 179(1):34–40.

Van Erp, T. S. (2012). Dynamical rare event simulation techniques for equilibrium and nonequilibrium systems. *Advances in Chemical Physics*, 151:27.

van Erp, T. S., Moroni, D., and Bolhuis, P. G. (2003). A novel path sampling method for the calculation of rate constants. *The Journal of Chemical hysics*, 118(17):7762–7774.

Vegas Mendoza, S. M., Avella Moreno, E., Guerrero Fajardo, C. A., and Fierro Medina, R. (2019). Liquid–liquid continuous extraction and fractional distillation for the removal of organic compounds from the wastewater of the oil industry. *Water*, 11(7):1452.

Wang, X., Chen, S., Kong, M., Wang, Z., Costa, K. D., Li, R. A., and Sun, D. (2011). Enhanced cell sorting and manipulation with combined optical tweezer and microfluidic chip technologies. *Lab on a Chip*, 11(21):3656–3662.

Ward, A. D., Berry, M. G., Mellor, C. D., and Bain, C. D. (2006). Optical sculpture: controlled deformation of emulsion droplets with ultralow interfacial tensions using optical tweezers. *Chemical communications*, (43):4515–4517.

Wigner, E. (1938). The transition state method. *Transactions of the Faraday Society*, 34:29–41.

Windhab, E. J., Dressler, M., Feigl, K., Fischer, P., and Megias-Alguacil, D. (2005). Emulsion processing—from single-drop deformation to design of complex processes and products. *Chemical Engineering Science*, 60(8-9):2101–2113.

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.

Woods, D. A., Mellor, C. D., Taylor, J. M., Bain, C. D., and Ward, A. D. (2011). Nanofluidic networks created and controlled by light. *Soft Matter*, 7(6):2517–2520.

Zwier, M. C., Adelman, J. L., Kaus, J. W., Pratt, A. J., Wong, K. F., Rego, N. B., Suárez, E., Lettieri, S., Wang, D. W., Grabe, M., Zuckerman, D., and Chong, L. (2015). WESTPA: An interop rable, highly scalable software package for weighted ensemble simulation and analysis. *Journal of chemical theory and computation*, 11(2):800–809.

Appendices

List of Figures

List of Tables

Paper I

Journal of COMPUTATIONAL CHEMISTRY

PyRETIS 2: An Improbability Drive for Rare Events

Enrico Riccardi,*[a] Anders Lervik,[a] Sander Roet,[a] Ola Aarøen,[b] and Titus S. van Erp[a,c]

The algorithmic development in the field of path sampling has made tremendous progress in recent years. Although the original transition path sampling method was mostly used as a qualitative tool to sample reaction paths, the more recent family of interface-based path sampling methods has paved the way for more quantitative rate calculation studies. Of the exact methods, the replica exchange transition interface sampling (RETIS) method is the most efficient, but rather difficult to implement. This has been the main motivation to develop the open-source Python-based computer library PyRETIS that was released in 2017. PyRETIS is designed to be easily interfaced with any molecular dynamics (MD) package using either classical or ab initio MD. In this study, we report on the principles and the software enhancements that are now included in PyRETIS 2, as well as the recent developments on the user interface, improvements of the efficiency via the implementation of new shooting moves, easier initialization procedures, analysis methods, and supported interfaced software. © 2019 The Authors. *Journal of Computational Chemistry* published by Wiley Periodicals, Inc.

DOI: 10.1002/jcc.26112

Introduction

Simulation of long time scales has been a major challenge in molecular simulations. Although the increase of system scale is straightforwardly parallelizable, extending simulation time is not. This is a severe problem as molecular dynamics (MD) typically requires femtoseconds time steps. Even if the computational evaluation of such a step is usually achieved within a fraction of second, it would take centuries of CPU time to compute 1 s of real time. This makes it nearly impossible to study processes like chemical reactions, phase transitions, and conformational changes with MD. These processes typically occur so infrequently that years of computer time are needed to observe even a single event.

The vast majority of methods, which have been developed for overcoming this problem, either alter the potential energy surface or the dynamics of the system (see, e.g., Refs. [1,2]). The use of Monte Carlo (MC) in path space is an approach that does not disturb the underlying physical dynamics, but generates unlikely molecular events like an *improbability drive*.[3,4] This approach is the essence of transition path sampling (TPS)[5–7] in which repetitively a new path is being generated from an old path via MC moves that obey detailed balance. The most important MC move is the so-called shooting move in which first a random time slice (comprising the phase point at a certain MD step) of the old path is selected and then stochastically modified. For instance, random disturbances could be applied to the velocities of that point. After that, this point is first propagated backward and then forward in time using a standard symplectic and time-reversible integrator for MD, Langevin, or Brownian dynamics. Moreover, as paths consist of time slices (phase points at discrete time steps), velocity Verlet,[8] and other reversible integration schemes should be preferred above leapfrog[9] as the latter provides velocities that are shifted in time by half a time step.[10] After the completion of these time integrations, the forward and backward trajectories are glued together resulting in a new continuous path that follows the natural dynamics of the system. This path is finally accepted or rejected based on a Metropolis–Hastings rule.[11,12]

Transition interface sampling (TIS)[13] introduced several fundamental key elements that made efficient quantitative path sampling possible. Firstly, TIS introduced the statistical path ensemble with flexible path length, reducing the redundant exploration of the stable regions. Secondly, a series of path ensembles were defined based on a set of hyperplanes (interfaces). These hyperplanes are generally defined by a value of the order parameter (reaction coordinate/progress coordinate) which is a function of the coordinates (and possibly velocities) of the system.

The first interface, called λ_0 or λ_A, defines the region of the initial state. The last interface, called λ_n or λ_B, defines the region of the final state. The first interface is placed such that a straightforward MD simulation starting from the reactant side would cross this interface sufficiently frequently, which enables the determination of the flux through this plane by straight forward MD. The last interface is placed sufficiently far across the barrier so that each trajectory from λ_0 to λ_n will not easily return to the

[a] *E. Riccardi, A. Lervik, S. Roet, T. S. van Erp*
Department of Chemistry, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway
E-mail: enrico.riccardi@ntnu.no
[b] *O. Aarøen*
Department of Biotechnology and Food Science, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway
[c] *T. S. van Erp*
Center for Molecular Modeling (CMM), Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium

initial state A and, therefore, can be considered as a successful transition from the initial state to the final state. As long as the positioning of λ_0 and λ_n is reasonable and respects the above criteria, the final result will not largely be affected by their exact placement. The other interfaces are defined in between λ_0 and λ_n and their positions are solely based on efficiency arguments.

Once the interfaces are defined, the TIS rate equation is then given by

$$k_{AB} = f_A \mathcal{P}_A(\lambda_B|\lambda_A)$$
$$\mathcal{P}_A(\lambda_B|\lambda_A) = \prod_{i=0}^{n-1} \mathcal{P}_A(\lambda_{i+1}|\lambda_i) \quad (1)$$

where f_A is the flux through λ_A and $\mathcal{P}_A(\lambda_B|\lambda_A)$ is the very small probability that a crossing with λ_A will lead to a crossing with λ_B without recrossing λ_A. As this probability is generally extremely small, it cannot be calculated directly. However, it can be computed by the exact factorization in the second expression in eq. (1). One should realize that $\mathcal{P}_A(\lambda_{i+1}|\lambda_i)$ is not just the probability to go from λ_i to λ_{i+1}, but rather a history-dependent conditional probability.

Replica exchange TIS (RETIS) samples all path ensembles in parallel and applies replica exchange moves between those.[14] In addition, it replaces the MD simulation for the flux calculation with the ensemble [0$^-$], which consists of paths that start and end at λ_0 but explore the reactant well region rather than the reaction barrier. For these features, RETIS is considerably more complex to implement in computer codes. This has been the main driving force to develop the open source PyRETIS[15] library.

A year after the official disclosure of PyRETIS, the Open Path Sampling (OPS) library was released.[16,17] The aims of PyRETIS and OPS are similar, but the libraries have been written with slightly different user communities in mind. OPS is more generic and allows the expert user to design different path ensembles. PyRETIS, on the other hand, has a stronger focus on the RETIS algorithm and a stronger emphasis on user-friendly accessibility (i.e., toward the nonexpert user). There are presently active collaborations between the two developer groups, which potentially could lead to a partial, or even full, merger of the two libraries in the future.

In this article, we discuss the code developments made in the release of PyRETIS 2. PyRETIS 1 was interfaced with GROMACS[18] and CP2K,[19] for respectively, performing classical and ab initio MD. In PyRETIS 2, we improved the GROMACS interface and added interfaces with OpenMM[20,21] and LAMMPS.[22] Several structural improvements have been made to improve the readability and reliability of the code. The major ones are shortly described in this study. To improve the efficiency, the new MC moves in path space, Stone Skipping and Web Throwing developed by Riccardi et al.[23] have been implemented. An easier initialization procedure has also been introduced such that trajectories, or simple snapshots, can now directly be read by PyRETIS 2 and used to initialize the RETIS simulation. In terms of outputs, a graphical user interface (GUI) to quickly inspect trajectories and density plots as functions of different descriptors (collective variables) has been constructed and added to the library.

Algorithmic Improvements

Advanced path generating MC moves

To optimize the sampling efficiency, PyRETIS 2 allows a direct and intuitive selection of path sampling strategies. Two of the most promising and recent sampling methods, namely, *Stone Skipping and Web Throwing*,[23] have now also been included in the code. Stone skipping and web throwing are two advanced MC-based moves that reduce correlations between successively generated paths. This implies that fewer generated trajectories are required to estimate crossing probabilities and rate constants within a desired error range. Therefore, despite that the execution of a single MC step is more costly than standard shooting, the sampling efficiency can increase considerably (for the case study of Ref. [23], the increase in efficiency was more than an order magnitude).

We expect that these new moves will become the default choice as they are intrinsically faster than standard shooting, though it will require some adaptations with how PyRETIS presently handles external engines, before this will be paid off in practice. The essential aspect of the new MC moves is that they launch a sequence of short subpaths via a shooting protocol that shoots from the ensemble-specific interface, that is, from a time slice just before or after the interface. The shooting move for creating a subpath propagates in one time direction only, though requires to cross the interface in one single time step backward in time or forward in time, depending on whether the selected shooting point is a point just beyond or before the interface. If this single-step crossing condition is not fulfilled, new random velocities should be generated until the condition is met. While this single-step crossing condition can be verified in principle for a single MD step without doing an actual force calculation,[23] the single-step crossing test has to be carried out explicitly whenever a (PyRETIS) step consist of several MD steps generated by the external engine. This makes the repeated generation of random velocities, followed by the testing of the one-step crossing condition, potentially, a very expensive element of this MC move. If the dynamics is sufficiently stochastic, then this one-step crossing test can be avoided by maintaining the same two time slices before and after the interface and create subpaths via one-way shooting protocol from the point that is after the interface, that is, without changing the velocities. For deterministic dynamics and dynamics that is only moderately stochastic (e.g., underdamped Langevin dynamics), the new MC moves might not yet outperform standard shooting until a new approach for handling the external engines has been implemented.

The [0$^-$] ensemble with additional confining interface

The [0$^-$] ensemble was introduced by van Erp[14] to replace the MD simulation for computing the flux in eq. (1) and to allow for replica exchange moves between all path ensembles. Paths in the [0$^-$] ensemble start at λ_0, like all other paths in the other ensembles, but move away from the barrier exploring the reactant well. The path is terminated once it recrosses λ_0. As a result, the time slices of a valid path in the [0$^-$] path ensemble

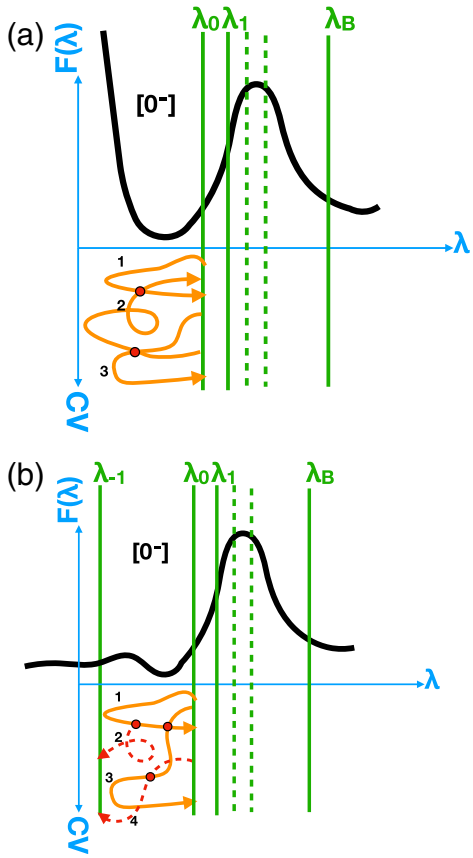


Figure 1. Illustration of the difference between standard RETIS interface positioning and RETIS with an extra λ_{-1} interface. Upper parts of a) and b) show the free energy as a function of λ with an example of how interfaces could be positioned. Bottom parts of a) and b) show examples of possible trajectories depicted in the (λ, CV) plane generated via the shooting move in the [0$^-$] ensemble, where CV is an additional collective variable different from λ. a) The standard situation of systems for which RETIS was originally designed. The free energy provides a natural boundary at the left side, which implies that the λ coordinate cannot get much smaller than λ_0 and respective path generation trials (indicated by the numbers 1, 2, and 3) starting from the shooting points (red circles) will relatively quickly end at λ_0 in the backward and forward time direction, yielding new acceptable paths. This situation is, for instance, typical for bond breaking, nucleation, and protein folding. b) The situation for which the additional λ_{-1} interface was introduced and is typical in, for example, permeation studies. In this case, the free energy does not provide a natural boundary at the left. Trajectories can in principle continue to move toward the left, allowing endlessly long trajectories. PyRETIS 2, however, allows the user to define a λ_{-1} interface, such that the shooting moves hitting the λ_{-1} interface are directly rejected. This is the case for trials 2 and 4 which hit the λ_{-1} interface in the backward and forward time direction, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

have order parameter values $<\lambda_0$ except for the start and end points.

This assumes that the reactant well has a natural confinement at the left side of the barrier, either because the potential energy increases when the order parameter is decreased below the equilibrium position of the reactant state, or because of periodic boundaries. For instance, if the order parameter is (minus) the distance between two molecules that can react if they approach, the periodic boundaries will ensure that their separation is bounded. However, if the box dimensions are large or if another undesired stable state exists at the left side of the reactant well, it might be needed to terminate the path early, as a completed path starting and ending at λ_0 could be exceedingly long. This can be done in PyRETIS 2 by setting an extra boundary, a "minus one" interface λ_{-1} with $\lambda_{-1} < \lambda_0$, that ensures that the reactant state ensemble is restricted between λ_{-1} and λ_0. In the present implementation, a path is rejected when it hits this left boundary. Please note that the use of a left boundary on this ensemble results in an incorrect flux calculation. A more sophisticated approach will be implemented in a future release in which the flux calculation will account for the presence of the extra boundary. At the present, the usage of the left boundary requires a reevaluation of flux term by means of another simulation without the left boundary. An illustrative scheme of the λ_{-1} is reported in Figure 1.

Interface with External Engines

The PyRETIS 1 program provided interfaces with GROMACS and CP2K. In PyRETIS 2, we have extended this with LAMMPS and OpenMM. In addition, some fundamental changes related to the GROMACS interface has been established in PyRETIS 2. These new developments are shortly discussed below. Working examples for each external engine can be found in the website (section: *Getting started*). The external engines communicate with PyRETIS at a certain frequency, defined by the *subcycles* number in the PyRETIS input file in the engine section. At each *subcycles* number, the external engine provides the order parameter or the information for its computation. Its magnitude depends on the selected engine and on the system under investigation (trade between speed, accuracy, and storage requirements).

GROMACS

The GROMACS engine has been updated for PyRETIS 2. The strategy for the GROMACS engine in PyRETIS 1 relied on repeatedly starting and stopping the execution of GROMACS. That is, PyRETIS 1 executes GROMACS for a few MD steps, defined by the number of *subcycles*, obtains the order parameter and uses this to determine if the GROMACS run should be ended or continued. PyRETIS 2 still not only supports this, but also provides a potentially more efficient strategy. It will execute GROMACS and obtain the order parameter while GROMACS is running and will use this to determine when it is time to end the GROMACS run. This increases the efficiency of running GROMACS with PyRETIS as it reduces the number stop/restart calls to the GROMACS engine. To select the old approach, the engine class to select in the input file is *gromacs*, while the latter method is called by the *gromacs2* keyword. The new approach exploits the functionality of the MDTraj[24] library for efficiently reading GROMACS trajectory files.

LAMMPS

PyRETIS 2 also includes an interface for LAMMPS.[22] For this engine, PyRETIS 2 will only provide information about the initial configuration and the stopping conditions (i.e., the maximum number of steps to perform and the location of the relevant interfaces). This enables the execution of LAMMPS to be fully handled by LAMMPS itself. This also requires that the order parameter is defined in a separate LAMMPS input file, which LAMMPS can use to calculate it. After the completion of a LAMMPS MD run, PyRETIS 2 will read the calculated order parameter and energies, and store the generated trajectory. Currently, PyRETIS 2 only supports microcanonical (NVE) dynamics when using LAMMPS. Note that temperature effects can still be studied as the temperature is linked to the MC sampling. That is, paths describe dynamics at constant energy, but the energies of different paths will fluctuate due to the shooting move. The approach can be defined as a canonical (NVT) sampling of NVE paths, and it is a popular approach in path sampling since it removes the dynamics from unphysical modifications of the equations of motion by a thermostat, while still sampling the canonical distribution. It reflects a system that is weakly coupled to a heat bath such that its effect is not noticeable on the time scale of the path length.

OpenMM

PyRETIS 2 has added an interface with OpenMM.[21] In this version, OpenMM can only be used with PyRETIS as a library. That means that if an OpenMM Simulation class is initialized, this object can then be used to initiate the OpenMMEngine class inside PyRETIS. This OpenMMEngine class can then be used for all the PyRETIS internals. An automated setup from restructured text input, like the way PyRETIS handles the connection with the other engines, is not yet supported and will be added in a later version of PyRETIS. As OpenMM also supports running on GPUs, special care was taken to not create new OpenMM contexts, but instead update the coordinates and velocities by an *in-place* operation. This minimizes the communication and prevents unnecessary compilation time for running on GPUs. The current implementation is only suitable for simulation in the canonical ensemble. Support for the isothermal–isobaric (NPT) will be added to a later version of PyRETIS.

Library Structure

Ensemble structure

The paths generated by Path Sampling are grouped into ensembles. Each of them focuses on a different region of path space. Each one can rely on different MC rules to generate new paths. Dedicated setups, tailored to the region to explore, can thus improve the sampling by enabling the application of the most suitable techniques. These possible techniques are for instance the use of Stone Skipping Web Throwing moves (see "Advanced path generating MC moves" section) or the different ways to disturb the velocities when performing a shooting move. To use this feature, a user should simply declare the ensemble to modify and specify the dedicated input as shown in the example in Figure 2.

```
Simulation
----------
task = retis
steps = 10000
interfaces = [ -0.9, -0.6, -0.3, 0.0, 1.0]

Ensemble
--------
interface = -0.9
shooting_move = sh

Ensemble
--------
interface = -0.6
shooting_move = ss
n_jumps = 16

Ensemble
--------
interface = -0.3
shooting_move = wt
n_jumps = 4
interface_sour = -0.5
```

Figure 2. New input structures to insert specific input for an ensemble. Each ensemble section refers to an interface specified in within. In the example, for the three ensembles selected, different shooting moves, with different parameters, have been selected. Note that the default shooting move is "sh," the first ensemble section is, therefore, not required.

In case that no special treatment for an ensemble is indicated, the main settings will be applied, preserving the same functionality as the previous version of PyRETIS.

Defining the order parameter and additional collective variables

The input file to PyRETIS has been updated to directly support several collective variables. This implies that a set of additional collective variables can be listed in the input file and those will be calculated along with the main order parameter. These collective variables do not affect the RETIS algorithm but can provide valuable information for the analysis of the path ensembles. The additional collective variables are hence descriptors to be used in postprocessing to elucidate mechanisms occurring in the investigated transition. There are, therefore, no constraints on the number or type of collective variables. The new input scheme to include additional collective variables is illustrated in Figure 3.

Paths storage and restart reproducibility

The storage of generated paths has been updated for PyRETIS 2. The trajectories can be saved with any frequency in a compressed format. The respective order parameter and energies (as a function of time) are saved in separate files with arbitrary frequencies too. This setup allows independent visual inspections of the generated trajectories and the restart of a new path

```
Orderparameter
--------------
class = Position
dim = x
index = 0
periodic = False

Collective-variable
-------------------
class = Velocity
dim = y
index = 16

Collective-variable
-------------------
class = Angle
index = 1, 3, 7

Collective-variable
-------------------
class = Custom
module = orderp.py
index = 11, 13, 17, 22
```

Figure 3. New input structures to include multiple collective variables along with the main order parameter. The order of the collective variables will be maintained in the generated output by PyRETIS in the orderp.txt files. In this example, the order parameter is the *x* position of atom 0, the first collective variable is the y component of the atom 16 velocity. The second collective variable is the angle between the atoms with indexes 1, 3, and 7. While these descriptors are computed with internal functions, the latter order parameter, called Custom, is an example of an externally computed descriptor (located in the module orderp.py).

sampling simulation from a previous successful trajectory (in case that the latest data got corrupted, e.g., by a hardware failure). Furthermore, by selecting a large number of descriptors (order parameter and collective variables), it is now possible to limit the size of stored data, while still having a detailed system description that can be quickly handled by the visualization tool introduced in the present release.

Random number generators

A new treatment of random number generators has been implemented in PyRETIS 2 to allow an exact reproducibility of simulations even if they have been performed in parallel. The random number generator is called at many places in the RETIS algorithm: In each cycle, a random number is drawn to select (1) the RETIS move (to select between the options swap/time-reversal/shooting or other types of path MC moves), (2) the relative shooting (if applicable), (3) the selection of the frame index, (4) the selection of the new velocities (velocities may be kept unchanged in the case of stochastic dynamics), and (5) the random forces whenever stochastic dynamics is applied. In PyRETIS 2, each task has a dedicated and unique random number generator. The feature permits a more accessible reproduction of the simulation results and a precise continuation of simulations even in parallel jobs.

MDTraj

PyRETIS 2 includes MDTraj[24] as an interpreter for external files. That is, external trajectories (from GROMACS, CP2K, LAMMPS) can be read with this Python library that was developed to deal efficiently with massive trajectories. The aim is to gain, in later releases of PyRETIS, increased independence from the external engine. As the minimal output that PyRETIS needs to receive from the external engine is only an ordered list of order parameters describing a trajectory, a universal interpreter to read external files in the various formats would simplify and uniform the interface for the various external engines. In PyRETIS 2, MDTraj is used in the load function to extract the desired frames. The package allows the use and development of arbitrary external functions to compute the order parameters and additional collective variables.

Input/Output Schemes

Load functionality and initialization

A new functionality, the *load* feature, has been added to PyRETIS 2 that simplifies one of the most user demanding tasks of the path sampling algorithm, the initialization of the simulation. The feature permits a direct initialization of path sampling simulations using configuration frames that could be generated by any type of fast simulation method and software. The new load function reads frames and trajectories supplied to PyRETIS 2 without the need for any further descriptor. PyRETIS 2 will compute the order-parameter, the additional collective variables and eventually the energy of the provided frames and trajectories. The input information will be automatically rearranged to satisfy the various ensemble definitions, when possible. PyRETIS 2 will then start the exploration of the path space from these initial frames indicating, with the "ld" flag in the output "pathensemble.txt" files, that the latest accepted path is a repetition of the initial path loaded. Once the initialization is completed this "ld" flag should no longer be present in the newly produced output lines in the "pathensemble.txt" files.

The strategy allows the inclusions of frames along the transition of interest that can be constructed by for instance constrained dynamics, nudge elastic band[25] or metadynamics[2] using any type of software. It should be underlined that the load function only simplifies the initialization procedure and will not influence the final converged result. It is, therefore, possible to provide to the load function a hypothetical path that has no physical meaning. The load initialization procedure also has a preprocessing feature to limit the overall simulation memory requirements. That is, for a massive trajectory, that explores the transition region only in a relatively short period, PyRETIS 2 automatically omits the frames that are not part of the ensemble of interest.

The load function can also be invoked after a restart. The frames from the restart file are then used as frames for the *load* method. The difference compared to a normal restart, which is a continuation, is that the load permits a broad change of the simulation parameters: from the order parameter selection to the interface positioning. It should be reminded here that, unlike a normal restart, the use of load with concomitant changes in the simulation parameters, should not be viewed as a continuation, and the previous sampling should be discarded in the final analysis.

Visualization

As specified by the user, different output files are created and also the frequency of writing the data can be set. Data sets can be created by optionally writing all-atom coordinates, velocities, and a list of descriptors (collective variables) for each trajectory at each time slice. Writing of all coordinates and velocities gives rise to huge data sets which asks for significant memory requirements for storage, and complicates the interpretation and data analysis. Therefore, we advise the user to reflect on what may be potentially important collective variables and then define a large set of descriptors *prior* to performing production runs. This simplification reduces the dimensionality of the data on which the interpretation of the reaction mechanisms shall be based. Still, even with this reduced dimensionality it can be nontrivial to filter out the essential information.

To facilitate this task, PyRETIS 2 includes a new visualization tool, named PyVisA. The software permits an almost immediate visualization of the initial, partial and final simulation results by allowing the automated generation of plots of various simulation collective variables. A user can promptly plot energies, the order parameter, collective variables, cycle number, and path lengths as a function of each other in different type of plots, for different ensembles, for selected cycle ranges, for accepted or rejected paths. A GUI has been constructed facilitate this visualization and to easily navigate through the PyRETIS 2 outputs.

The visualization of the various descriptors, for example, the collective variable density plots, in the GUI allows the user to interactively inspect different parameter combinations, potentially revealing additional information about the system dynamics. In the initialization stage, a user can better position the various interfaces, select the most appropriate MC move (e.g., standard shooting, Stone Skipping, Web Throwing, etc.), determine metastable states and even evaluate the efficacy of the selected order parameter in comparison with other collective variables. The descriptors include order parameters, collective variables, energy, number of simulation steps for path, RETIS cycles. The user can select the range of cycles to visualize, their type (accepted/rejected) and the plot type (e.g., 2D scatter, 3D scatter, local density). Figures 5 and 4 show two reports that can be obtained by the visualization tool's GUI. To execute the visualization and analysis tool (named "PyVisA"), the flag -pyvisa shall be added to the pyretisanalyse command. Further details, instructions, and examples can be found in Ref. [27], where the tool structure and features are detailed.

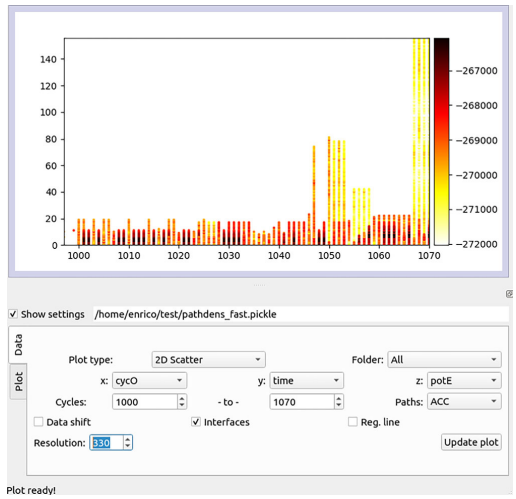


Figure 4. Visualization ("Plot type: 2D Scatter") of a set of accepted trajectories ("Paths: ACC") for a given range of cycles ("Cycles: 1000 to 1070") for all ensembles ("Folder: All"), plotted according to the cycle number ("x: cycO"), trajectory length ("y: time") and the potential energy ("z: potE"). With these selections, an user can gain a statistical insight in the progress of the sampling for a restricted range of cycles. From this illustrative plot, it can be noticed that the paths generated in the surrounding of cycle 1070 are longer and with lower potential energy than the previous for the given range. A nonuniform path length distribution can be symptom of a nonconverged sampling. The paths in the latter cycles seems to have identified a region with lower potential energy, a different pathway might have been, therefore, identified. [Color figure can be viewed at wileyonlinelibrary.com]

The visualization tool, furthermore, provides a structured computational framework that will permit a general implementation of advanced analysis approaches. Predictor methods[28] or machine learning methods[29] to evaluate the quality of the selected order parameter in the description of the sampled event are the two most immediate examples.

User Support

Compatibility and installation

PyRETIS 2 supports Python 3.6 and 3.7 and is distributed via pip and conda (via the conda-forge channel) for the main releases. The visualization package, PyVisA, can be found in the development versions of PyRETIS (PyRETIS v2.develop and PyREITS 3.beta) and in the forthcoming releases beyond Version 2.4. The development versions can be installed by downloading the PyRETIS source code from gitlab via the command "python setup.py install" from the main directory. Further details on its installation and usage can be found in Ref. [27].

Test examples

Along with a long list of unit tests, the PyRETIS development is also tested versus a series of main test simulations that are automatically executed daily to assert the code and its

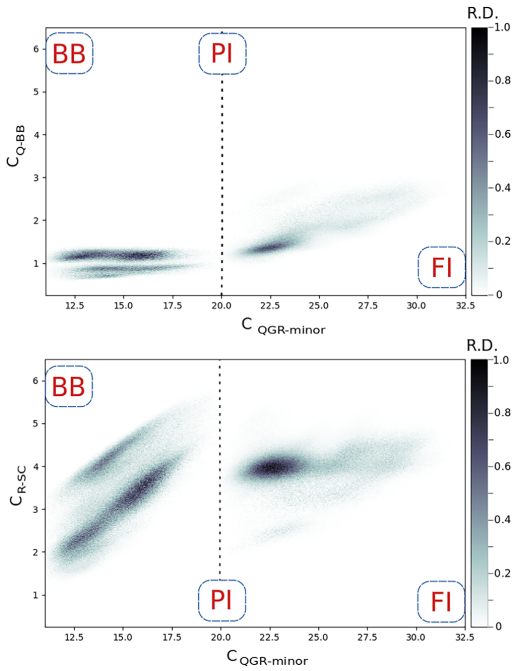


Figure 5. (Figure taken from Ref. [26]) Path frame density plots of RETIS trajectories for the insertion transition of H-NS to DNA. $C_{QCR-minor}$ is the main order parameter and it is obtained by a contact map between H-NS (QGR motif) and the DNA minor groove region. The descriptors C_{R-SC} and C_{Q-BB} are obtained by the contact map of specific H-NS region (R or Q motif) and DNA region (side chain and back bone). A complete description of the descriptors on the axis and of the meta-stable states can be found in the work of Riccardi et al.[26] Darker color represents a higher probability. The plots showed that for the Q–BB interaction to happen, the protein has to be in contact with DNA first, while the R–SC interaction seems to anticipate the H-NS–DNA interaction. In essence, these descriptive plots showed part of the mechanisms of H-NS adsorption. [Color figure can be viewed at wileyonlinelibrary.com]

dependencies status. These test simulations are available in the development versions of PyRETIS (which can be installed via git as described in the section on code availability). The test simulations have also a pedagogical purpose: users can gain experience and familiarity with the PyRETIS input scheme. These simulations are grouped by the external engine selected and by the functionality under test. Hereby, we are constantly increasing the number of test simulations. At the current stage, test simulations show the usage of different engines, initiation functions (e.g., "kick," "load," and "restart"), different simulation schemes (e.g., MD, TIS, and RETIS) and different selection of shooting moves (e.g., Stone Skipping and Web Throwing).

Web interface

By the release of PyRETIS 1, we created a website (www.pyretis.org) to give support to PyRETIS users. With PyRETIS 2, new examples and guides have been included in the *Examples* section and the *User Guide* section to further facilitate and guide the usage of PyRETIS. In particular, the example section contains working examples for each of the external engines supported.

Cases of study performed with PyRETIS are constantly uploaded on the PyRETIS web page under the "Main studies performed by PyRETIS" section, and the files used to initiate and control the simulations are shared to facilitate the reproducibility of our investigations, with respect to FAIR data policies.[30]

In particular, our recent investigation of water autoionization[31] and the Histone-like Nucleoid Structuring protein (H-NS) binding to DNA[26] are listed on the website. They constitute two successful applications of the RETIS algorithm, where the sampling software has been interfaced with CP2K[19] in the first work, and GROMACS[18] in the latter. The *User Guide* section contains a new set of entries to facilitate the usage of the code. The guide comprises (1) instructions for the installation of PyRETIS in a user and a developer mode, (2) information on how to use and set up external engines and order parameters, (3) help for common errors, and (4) instructions on how to report bugs.

Future Work

Despite the considerable efforts put in the code development, there are various expansions that would be desirable to increase usability, efficiency, and compatibility with other sampling software. We aim to automatize some of the parameter selections such as the interface positions, the number of jumps in the stone skipping and web throwing moves and the SOUR interface[23] in web throwing, the relative shooting weights, and the frequency of selection of the various MC moves.

An interface with VASP has been initiated and partially completed. It will be released after completion and after performing sufficient testing. In parallel, we are considering the implementation of a "translation" platform that might enable a single scheme to deal with the input and output of multiple engines. Besides the already implemented MDTraj,[24] Python packages such as MDanalysis[32,33] and Atomic Simulation Environment (ASE)[34] can facilitate the realization of such a platform.

In the next PyRETIS release, the visualization tool will become an integral part of the code. It will provide a multidimensional and structured analysis framework. Advanced analysis approaches, such as the predictive power analysis method[28] and machine learning based methods to evaluate the quality of different collective variables,[29] will be readily implemented. We are therefore interested in direct support and collaboration with potential new developers that are interested to apply and expand PyRETIS.

Software Availability

PyRETIS 2 is free (released under a LGPLv2.1+ license) and can be obtained as described previously and at http://www.pyretis.org/user/install.html. The source code, the visualization tool and the development version are accessible at: https://gitlab.com/pyretis/pyretis.

Acknowledgments

The authors thank the Research Council of Norway for funding (project nos: 250875 and 267669), NOTUR for providing HPC facilities (project no.: NN9254K), the Olav Thon foundation for the support in the development of interactive visualization tools for teaching and the Peder Sather Center for the support in the development of the LAMMPS-PyRETIS interface. The authors also thank the Lorentz Center for supporting our workshop on PyRETIS and Path Sampling in Leiden, March 11–15, 2019. Sudi Jawahery, Anastasia Maslechko, Raffaela Cabriolu, An Ghysels, Jocelyne Vreede, Christopher Daub, and Mahmoud Moqadam are thanked for their feedbacks and useful discussions.

Keywords: PyRETIS · rare event · path sampling · python · kinetics

How to cite this article: E. Riccardi, A. Lervik, S. Roet, O. Aarøen, T. S. van Erp. *J. Comput. Chem* **2020**, *41*, 370–377. DOI: 10.1002/jcc.26112

[1] F. G. Wang, D. P. Landau, *Phys. Rev. Lett.* **2001**, *86*, 2050.
[2] A. Laio, F. L. Gervasio, *Rep. Prog. Phys.* **2008**, *71*, 126601.
[3] D. Castelvecchi, *Sci. News* **2007**, *171*, 372.
[4] D. Adams, The Hitch Hiker's Guide to the Galaxy: A Trilogy in Five Parts, Hitchhiker's Guide to the Galaxy Series, Penguin Random House, **1995**.
[5] C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler, *J. Chem. Phys.* **1998**, *108*, 1964.
[6] P. G. Bolhuis, C. Dellago, D. Chandler, *Faraday Discuss.* **1998**, *110*, 421.
[7] C. Dellago, P. G. Bolhuis, D. Chandler, *J. Chem. Phys.* **1999**, *110*, 6617.
[8] L. Verlet, *Phys. Rev.* **1967**, *159*, 98.
[9] O. Buneman, *J. Comput. Phys.* **1967**, *1*, 517.
[10] B. J. Leimkuhler, S. Reich, R. D. Skeel, Integration Methods for Molecular Dynamics. In Mathematical Approaches to Biomolecular Structure and Dynamics, New York: Springer, **1996**, p. 161.
[11] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087.
[12] W. K. Hastings, *Biometrika* **1970**, *57*, 97.
[13] T. S. van Erp, D. Moroni, P. G. Bolhuis, *J. Chem. Phys.* **2003**, *118*, 7762.
[14] T. S. van Erp, *Phys. Rev. Lett.* **2007**, *98*, 268301.
[15] A. Lervik, E. Riccardi, T. S. van Erp, *J. Comput. Chem.* **2017**, *38*, 2439.
[16] D. W. H. Swenson, J.-H. Prinz, F. Noe, J. D. Chodera, P. G. Bolhuis, *J. Chem. Theory Comput.* **2019**, *15*, 813.
[17] D. W. H. Swenson, J.-H. Prinz, F. Noe, J. D. Chodera, P. G. Bolhuis, *J. Chem. Theory Comput.* **2019**, *15*, 837.
[18] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19.
[19] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *WIREs Comput. Mol. Sci.* **2014**, *4*, 15.
[20] P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, V. S. Pande, *J. Chem. Theory Comput.* **2013**, *9*, 461.
[21] P. Eastman, J. Swails, J. D. Chodera, R. T. McGibbon, Y. Zhao, K. A. Beauchamp, L.-P. Wang, A. C. Simmonett, M. P. Harrigan, C. D. Stern, R. P. Wiewiora, B. R. Brooks, V. S. Pande, *PLoS Comput. Biol.* **2017**, *13*, 1.
[22] S. Plimpton, *J. Comput. Phys.* **1995**, *117*, 1.
[23] E. Riccardi, O. Dahlen, T. S. van Erp, *J. Phys. Chem. Lett.* **2017**, *8*, 4456.
[24] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, V. S. Pande, *Biophys. J.* **2015**, *109*, 1528.
[25] G. Henkelman, B. P. Uberuaga, H. Jónsson, *J. Chem. Phys.* **2000**, *113*, 9901.
[26] E. Riccardi, E. C. van Mastbergen, W. W. Navarre, J. Vreede, *PLoS Comput. Biol.* **2019**, *15*, e1006845.
[27] O. Aarøen and E. Riccardi. [Submitted in October 2019 to PeerJ Physical Chemistry]. Pyvisa: Visualization and Analysis of Path Sampling Trajectories.
[28] T. S. van Erp, M. Moqadam, E. Riccardi, A. Lervik, *J. Chem. Theory Comput.* **2016**, *12*, 5398.
[29] M. Moqadam, A. Lervik, E. Riccardi, V. Venkatraman, B. K. Alsberg, T. S. van Erp, *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E4569.
[30] E. Riccardi, S. Pantano, R. Potestio, *Interface Focus* **2019**, *9*, 20190005.
[31] M. Moqadam, E. Riccardi, T. T. Trinh, A. Lervik, T. S. van Erp, *Phys. Chem. Chem. Phys.* **2017**, *19*, 13361.
[32] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein. MD Analysis: A Toolkit for the Analysis of Molecular Dynamics Simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327. https://dor.org/10.1002/jcc.21787.
[33] R. J. Gowers, M. Linke, J. Barnoud, T. J. Reddy, M. N. Melo, S. L. Seyler, J. Domański, D. L. Dotson, S. Buchoux, I. M. Kenney, O. Beckstein, MDanalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. In Proceedings of the 15th Python in Science Conference, Vol. *98*, SciPy, Austin, TX. **2016**.
[34] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, *J. Phys. Condens. Matter* **2017**, *29*, 273002.

Received: 8 July 2019
Revised: 25 October 2019
Accepted: 29 October 2019
Published online on 19 November 2019

Paper II

Received: 27 August 2020 | Revised: 3 November 2020 | Accepted: 26 November 2020

DOI: 10.1002/jcc.26467

FULL PAPER

Journal of COMPUTATIONAL CHEMISTRY WILEY

PyVisA: Visualization and Analysis of path sampling trajectories

Ola Aarøen[1] | Henrik Kiær[2] | Enrico Riccardi[2]

[1]Department of Biotechnology and Food Science, Norwegian University of Science and Technology, Trondheim, Norway

[2]Department of Chemistry, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence
Enrico Riccardi, Department of Chemistry, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway.
Email: enrico.riccardi@ntnu.no

Funding information
Norges Forskningsråd, Grant/Award Number: 267669

Abstract

Rare event methods applied to molecular simulations are growing in popularity, accessible and customizable software solutions have thus been developed and released. One of the most recent is PyRETIS, an open Python library for performing path sampling simulations. Here, we introduce PyVisA, a postprocessing package for path sampling simulations, which includes visualization and analysis tools for interpreting path sampling outputs. PyVisA integrates PyRETIS functionalities and aims to facilitate the determination of: (a) the correlation of the order parameter with other descriptors; (b) the presence of latent variables; and (c) intermediate meta-stable states. To illustrate some of the main PyVisA features, we investigate the proton transfer reaction in a protonated water trimer simulated via a simple polarizable model (Stillinger–David).

KEYWORDS

kinetics, path sampling, PyRETIS, python, PyVisA, rare event, trimer, water

1 | INTRODUCTION

Rare event simulations are becoming a common method of sampling transitions. Thanks to the recent advancements,[1–3] molecular simulations can now capture events happening in a time scale of seconds.[4–6] Between rare event methods, transition interface sampling[7–10] appears to be a particularly promising approach since it does not require any bias potential or further approximations. In particular, a recent development, named "replica exchange transition interface sampling," (RETIS) has further increased method sampling efficiency.[11,12]

Since a detailed description of path sampling can be found in the work of van Erp,[11] only a brief introduction will be provided here.

Transition interface sampling approaches compute the transition rate as the product of a *flux* and a *crossing probability* term. To compute the two quantities, an order parameter has to be defined: a one dimensional descriptor that quantifies the progress along the transition of interest. The initial and final states of the transition are thus marked by two values of the order parameter at which interfaces can be arbitrarily positioned: λ_0 being the first interface, and λ_B the last interface.

The frequency at which the system leaves the initial state by crossing the interface λ_0 determines the *flux* term. The crossing probability is the likelihood of a given system having passed λ_0, to reach λ_B. Between the two interfaces, a set of intermediate ones, λ_i, can be positioned to increase the sampling efficiency. Each interface defines an ensemble containing trajectories that have visited the respective region of the transition.

Determination of the transition rate is one of the main targets of path sampling. In systems with a large number of degrees of freedom, transitions can be hindered by intermediate metastable states and/or intermediate events. The definition of a sufficiently descriptive order parameter is not always trivial since the free energy surface for the transition is not known *a piori* for most studies. Better insight into the transition allows a better set up of the path sampling simulation and of the output interpretation. For instance, the sampling can be focused on the most difficult regions to explore and more advanced path sampling strategies[12] can be adopted. Furthermore, a transition can be the result of a complex sequence of substeps that, in some cases, could provide more valuable quantitative information regarding the transition than that provided by its overall rate.

Methods aimed at improving the analysis of transitions are thus commendable since significant information has to be extracted from a large quantity of data. Our recent work on water autoionization[5] and DNA interaction with H-NS[6] has highlighted this necessity: massive simulation campaigns requiring extremely involved postprocessing analysis.

To facilitate the usage of the sampling technique, we have recently released a Python[13] open library named PyRETIS,[14,15] which

includes the most recent transition interface sampling method developments.[12]

While different path sampling software are present, very few tools exist to facilitate the visualization and analysis of trajectory data. WESTPA[16](The weighted ensemble simulation toolkit with parallelization and analysis) and OPS[17](OpenPathSampling) are two packages for performing path sampling with similarities to PyRETIS. While they contain tools to read and process the outputs they generate, only a limited set of visualization and analysis features are offered.

Previously, PyRETIS postprocessing features were also limited and mostly geared towards the computation of the rate. To handle the various descriptors that can be generated during simulations (i.e. collective variables and energy terms), we have developed a Python library named PyVisA.

The package can extract, visualize, analyze, and store the subset data of interest, providing output ranging from raw data to full Python image objects. The constructed framework allows a direct data analysis on path sampling results,[18] and it is not bound to PyRETIS[14,15] internal object representation. The library includes a discrete set of analysis methods and provides a base computational structure upon which more complex analysis can be built (e.g. predictive power methods[19] and advanced machine learning techniques to recognize the most relevant latent variables, reaction mechanisms, and energy minimum).

To demonstrate the capabilities of PyVisA, we have included a study on the proton transfer reactions in a water trimer in the second part of this manuscript. With this aim in view, we have chosen a relatively simple model that, without electronic configuration calculations, permits bond breakage. For this study, we adopted the Stillinger-David (SD) model,[20] which considers a water molecule to be two fixed electric charges for the hydrogen atoms and a polarizable charge for the oxygen atom. The model has proven to be effective at illustrating both the formation of hydrogen bonds[21] in water, as well as the charge molecular (re)-distribution of reacting molecules.[22]

A description of the reaction mechanism will be given and the reaction rate reported. It should be noted that our work does not aim to improve the current physical understanding of the reaction but instead has chiefly a modeling integration objective. Our results integrates the findings of the work of Geissler et al.[22] For an improved physical description of the proton transfer mechanisms, path sampling shall be used with more recent models such as improved SD,[23] density functional theory,[24] Møller-Plesset perturbation theory,[25] vibrational motion[26,27] and finally, the Neural Network Potentials at Coupled Cluster Accuracy.[28]

2 | PYVISA

PyVisA is a Python open library designed for postprocessing and analysis of path sampling results. It has been primarily developed to integrate PyRETIS functionalities.[14,15] The information flow between PyRETIS and PyVisA is depicted in the flowchart of Figure 1.

To allow a robust and efficient analysis of data and to avoid GUI compatibility limitations with clusters, two modules have been designed: an input processing module (compressor) and a GUI module (visualizer). The first reads the inputs, checks their integrity, and saves the data in a compressed and structured file. The second loads the data and launches a GUI which allows quick visualization of arbitrarily

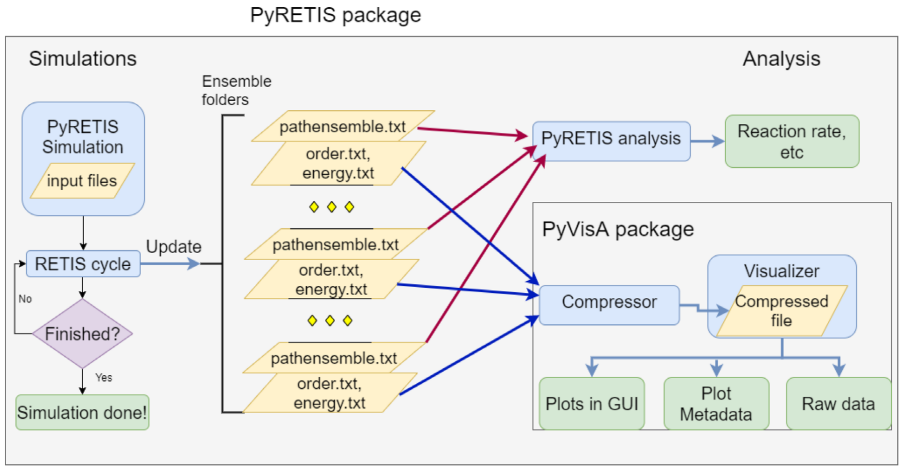


FIGURE 1 Flowchart of the PyRETIS and PyVisA software. PyRETIS generates and continuously writes trajectory data to output files in ensemble folders (000, 001, etc.). The PyVisA compressor script reads the descriptors files, checks their integrity, and compiles a compressed file. The visualizer retrieves the subdata set selected, plots it in a dedicated panel of the GUI, and allows further data analysis. Since PyVisA is mostly independent from the main simulations' flow and software, it can operate on the fly

definable subsets. A set of analysis on the data set of interest can then be operated directly via the GUI and the results saved in several different formats, in order to facilitate reporting and further analysis.

2.1 | PyVisA Preprocessing module

Path sampling simulations can require a massive quantity of computations consequently producing a massive quantity of data.

The calculation of a transition rate requires a minimal set of descriptors for each generated trajectory to be stored. To describe reaction mechanisms, further information (i.e. collective variables and energy descriptors) shall be additionally considered.

PyRETIS outputs different files in different folders for each ensemble. Even for relatively small systems, data storage can become excessively demanding: each trajectory is composed of a sequence of frames, each of them consisting of $3 \times N_{atoms}$ entries, if only the atom positions are considered, and up to $9 \times N_{atoms}$, if velocities and forces are also saved. For large systems, full trajectories are commonly stored at a relatively low frame frequency and mostly for visualization purposes. By using functions that map the system's atomic position to a set of descriptors (e.g. distance, angles, energies), each frame becomes a relatively short list of descriptors. Such a dimensional reduction of output data is significant and reduces the requirements for postprocessing operations. PyRETIS outputs full trajectories and the various descriptors to files at arbitrary output frequencies. As described in Figure 1, PyRETIS saves in text format the output files containing the path description (pathensemble.txt) and the various descriptors (order.txt and energy.txt) for each ensemble. Full trajectories are stored in the same format as the one given by the external engine. PyVisA can read stored full trajectories and descriptor files. If full trajectories are provided as an input, PyVisA can recompute the relative descriptors (i.e. compute new collective variables). It should be noted that since full trajectories are not commonly stored as often as descriptors, due to storage limitations, the resulting output may contain less information. This basic data structure (text file table and full trajectory in the external engine format) is a simple PyVisA input requirement which allows a rather simple development of interfaces with other sampling software.

In intensive simulation campaigns, abrupt interruptions of simulations are common. Large file dimensions can lead to storage and handling issues. PyRETIS output can result in being incomplete or corrupted and/or the descriptor files mismatching. The PyVisA compressor tool automatically handles several different types of file inconsistencies, discarding the incomplete and/or corrupted data. Since PyVisA does not process the outputs to compute the transition rate, it can be expected that removing small sections of data will have a very modest influence on the path sampling results (the source files remain unaltered). The preprocessing module composes a data table of descriptors (order parameter, collective variables and energy) for each trajectory. Each trajectory is stored with the respective labels (cycle number, ensemble number, generating shooting move, status [ACC or REJ], reactive/unreactive path) in a pandas[29] dataframe.

Since visualization and analysis of simulation output are often operated in physically distinct machines (e.g. computation on clusters and analysis on simple laptops/workstations), data transfer can be a factor limiting postprocessing effectiveness. By operating PyVisA from the command line, a compressed binary .hdf5 file can be generated (.hdf5.zip).

To test PyVisA compression efficiency, we report in Table 1 three different test cases. In each case, the input data is well-formatted, so no instances have been disregarded. It can be noted that, while for shorter runs PyVisA data structuring actually increases the storage requirements, for long simulations the observation is reversed, outperforming the zip compression tool included in standard Linux distributions. For the water trimer test case (second column), PyVisA generated a single structured binary file of 55 MB from 32 text files and 263 MB in total.

As a command line operation, PyVisA can recalculate order parameters and collective variables of frames/trajectories/simulations. By reading the input file (.rst), which should contain the information to handle the calculation of the order parameter, (e.g. the name and input parameter for an external module). PyVisA can recompute collective variables by adding the flag -recalculate to PyVisA executable. When using the flag -data <source>, where <source> can be either a file (frame or trajectory) or a folder, the order parameter of the contained frames/trajectories will be recomputed. When <source> is a main simulation folder, PyVisA will seek for the full trajectories consistently with PyRETIS storage scheme and it will regenerate the relative "order.txt" for each of the stored trajectory. The recalculation of collective variables can also be launched in the GUI, as later described.

2.2 | PyVisA GUI module

The central component of PyVisA is a GUI window that gives users access to various features. It has four panels: data-set selection, plotting layout, trajectory visualization, and analysis. Figure 2 shows the appearance of the GUIs' panels.

TABLE 1 File dimensions comparison with and without the usage of the PyVisA compressor tool for three different simulation test cases

Case	Test 1	Water trimer	Test 2
Number of file lines	20,702	2,505,275	18,722,085
Base filesystem	708 KB	263.3 MB	1.4 GB
Zip of base filesystem (no PyVisA)	261 KB	54.6 MB	399 MB
HDF5 of dataset	2.5 MB	71.6 MB	319 MB
HDF5.zip dataset	253 KB	55.1 MB	213 MB

The first panel permits a quick selection of subsets of data via different check-boxes and drop-down menus. The available options for data selection and visualization are:

- plot type,
- axis descriptors,
- ensemble folder,
- cycle range selection,
- path generating moves,
- path status (Accepted, Rejected),
- path types (reactive/unreactive).

Four check boxes are also present:

- regression line: to calculate and display the equation obtained by the linear regression of the data (only for 2d plot).
- data shift: to transpose the center of observation and apply periodic boundary condition (e.g periodic collective variables).
- interfaces: to display the ensemble interfaces in 2D plots or planes in 3D plots (input taken from .rst file).
- only stored trajs: to display only the paths for which the full trajectory is stored in the memory (it assumes folder naming and location according to PyRETIS data storage).

In the same panel, two buttons are also located: Update Plot and Refresh Data. When triggering Update Plot, a new visualization will appear based on the selected criteria. "Refresh Data," on the other hand, will reimport the data from the source files. The latter function is designed to visualize newly generated output while simulations are in progress.

The panel contains a matplotlib.pyplot[30] canvas, as a base for plotting within a PyQt5 window.[31] Figure 2(A) shows the appearance of the data panel.

A second panel manages the plotting layout. In addition to plot appearance considerations and due to the large variety of possible visualizations produced by PyVisA, plot color selection permits a clearer interpretation: colors are often used to represent a third dimension. The desired setting is applied by clicking on the "Update Plot" button. The generated figure can then be printed in a .png format in the working directory, by clicking on "Save" button. Figure 2(B) shows the outlook of the GUI for the plotting selection.

A third panel in the PyVisA GUI, reported in Figure 2(C), permits further data visualization, focusing on the single trajectories that compose the dataset. By clicking on the plot, a user can select a certain trajectory, visualize its properties (e.g., shooting move, status), and emphasize it with dedicated settings present in the panel. In case the trajectory has also been stored in full detail, Avogadro[32] application fed with the selected trajectory can be initiated by clicking on the "Play" button. By selecting the "Save Trj" button, instead, PyVisA stores the full continuous trajectory in a new .pdb file to facilitate further visualizations. This function is convenient as the trajectories produced by RETIS shooting moves are composed of two segments: one forward in time and one backward in time.

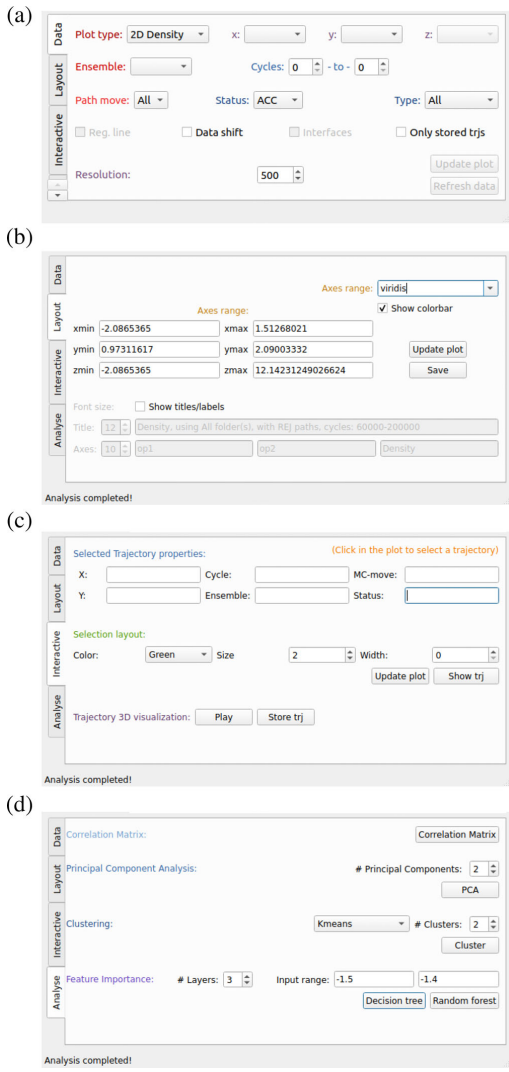


FIGURE 2 Top panel (A): Visualizer GUI settings tabs for data loading; middle panel (B): Visualizer GUI settings tabs for plot layout; bottom panel (C): Visualizer GUI settings tabs for trajectory visualization; bottom panel (D): Visualizer GUI settings tabs for analysis

The fourth and latest panel, labeled Analysis, contains a set of statistical methods that can be operated on the selected data sets. The methods are mostly machine learning-based approaches provided by the Python package *scikit-learn*:[33,34] correlation matrix, clustering, principal component analysis (PCA)[35] and decision trees[36] .

The correlation matrix is generated by computing the Pearson's correlation coefficients[29] between order parameters, collective

variables, and energy terms. Clustering can be performed via a set of methods: k-means,[37] agglomerative hierarchical clustering,[38] Gaussian mixture,[39] spectral clustering.[40] A user can select the number of clusters to be generated. The principal components analysis also allows the selection of the number of principal components to be found. The information entropy-based decision trees can be generated as a single analysis or as a random forest of decision trees. In the latter case, a report on the relative importance of each feature will be provided. A range of order parameter values can also be provided as input for the analysis: all frames with an order parameter value included in the given range will then be considered. The decision tree is generated according to the relative path property (reactive, unreactive).

2.3 | File menu

To increase the accessibility of the GUI, a drop-down "File" menu has been included. The menu conveniently reproposes some functionalities from the various panels with further options. The first entry of the menu is a load data function. It can be used to import raw output from PyRETIS, a PyVisA compressor generated file (.hdf5 or .hdf5.zip), and even a single folder or full trajectory for which the descriptors have to be recalculated.

With the second menu's entry, the whole of the simulation data can be (re) saved in a binary file, in either standard (.hdf5) or compressed (.hdf5.zip) format. It should be noted that the full trajectories are not included in the compressed file and that they should be considered separately.

A third entry allows the storage of the subset of data selected in the GUI. The plot can be exported in raw formats, either .txt or JSON. The plot object can be saved as a .hdf5 file, while the figure can be simply stored as an image .png. To further provide intuitive data interpretation, PyVisA can also generate a script to reproduce the selected image from the main compressed data file (option ".py script"). The name of the output file is automatically generated by PyVisA, and it is composed of the selection criteria used to display the selected dataset.

A fourth entry in the menu, "Recalculate Data", contains a convenient automatic function. When selected, PyVisA will analyze all the stored paths present in the running directory according to the PyRETIS storage scheme. For each of these trajectories, the order parameter and collective variables will be recomputed generating new data files to display. While the procedure can be extremely useful when adding new descriptors in the result analysis, it should be noted that since few full trajectories are commonly stored in memory, the procedure can result in less data than the initial selection.

The *Visualizer* tool, therefore, allows the immediate selection between very large sets of data and a quick preparation of plots. Different analyses can also, thereafter, be executed. Physical mechanisms can be thus captured with a reduced burden of data management. Data can be readily extracted enabling further postprocessing analysis with different computational platforms and improved visualizations.

2.4 | Availability and usage

To respond to different user demands, PyVisA has been released in two versions. The basic version included in PyRETIS 2 contains only the first two panels of the GUI (data selection and layout) and has limited functionalities. It has an easier installation procedure and requires only the PyQt5 library on top of PyRETIS requirements. This version should cover most initial user needs, and it is included in the pyretisanalyse PyRETIS 2 executable. The installation procedure and usage are detailed at the PyRETIS website www.pyretis.org.

To access the full PyVisA version described in the present paper, the develop version of PyRETIS should be installed. It can be freely downloaded from the git repository gitlab.com/pyretis/pyretis, develop branch, and installed with the relative dependencies listed in the file "requirement-dev.txt."

To launch PyVisA, it is sufficient to enter the command:

pyvisa

The GUI interface will be launched. To point to source data and/or access to specific PyVisA functionalities, the main executable can be used with different flags. The mains are:

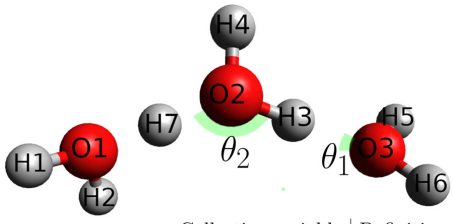
- -i < input-file>
- -cmp: to execute only the compressor tool (no GUI required)
- -recalculate: to recalculate the collective variable (no GUI required)
- -data <input-data>: to specify the folders where the source data is stored
- -oo Read only the order parameter files. This can be useful when the input data is very large.

where <input-file> can be either the input file used for PyRETIS simulation or the PyVisA compressed file (.hdf5 or relative .hdf5.zip file). <input-data> can be used to point towards different source data. If a folder in which simulations have been performed is selected, PyVisA assumes the folder naming and structure of PyRETIS to retrieve the input information (the default value for <input-data> is the execution directory). If the -recalculate flag is used with -data flag, PyVisA will compute the collective variables and write a order.txt file from the frames and/or trajectories contained in <input-data>.

The current PyRETIS develop branch will become the forthcoming PyRETIS 3 release. Detailed updates and examples will be consistently added to the PyRETIS website www.pyretis.org.

3 | PROTON TRANSFER IN THE PROTONATED WATER TRIMER

To demonstrate the various PyVisA features, we will report herein a detailed description of the path sampling simulation results of the proton transfer reaction in a protonated water-trimer system,[22,24] depicted in Figure 3. The polarization model for water[20] describes the atomic and dipole interactions as it allows for bond breakage and provides good approximation of the structure of water molecules because of a self-consistently calculated dipole moment.



Collective variable	Definition	Unit
(Order parameter) 1.	$\theta_1 - \theta_2$	rad
2.	O_1-H_7	Å
3.	O_2-H_7	Å

FIGURE 3 The protonated water-trimer system with a description of the collective variables

3.1 | Model and methods

According to the Stillinger–David potential, a single water molecule includes three electric charges for the three atoms, where only the oxygen is polarizable. The potential energy of interaction is composed by the pair-interaction contribution and the dipole polarizability:

$$\Phi_{total} = \Phi_1 + \Phi_2 \tag{1}$$

The first term is the potential energy of interaction and is obtained by the sum of Coulombic forces between each pair of the atoms in the system.

$$\Phi_1 = \sum_{i<j=1}^{N_H} \phi_{HH}(r_{ij}) + \sum_{i=1}^{N_H}\sum_{j=1}^{N_O} \phi_{OH}(r_{ij}) + \sum_{i<j=1}^{N_O} \phi_{OO}(r_{ij}) \tag{2}$$

In the model, hydrogen is considered as a lone proton with a point-charge. Its interaction potential, $\phi_{HH}(r_{ij})$, has a simple Coulombic repulsion, while oxygen atoms have a full shell of electrons that deviates from the purely Coulombic form. The interatomic potential $\phi_{OH}(r_{ij})$ describes the covalent bond between the oxygen and hydrogen atoms, while the interaction potential $\phi_{OO}(r_{ij})$ considers the electronic cloud overlap repulsion.

The second term of the potential in Equation (1) describes the dipole moment from the electric field of each particle due to the presence of other particles, with a relatively simple formula:

$$\Phi_2 = \sum_{i=1}^{N_o} \sum_{\substack{j=1, \\ j\neq i}}^{N_{tot}} \frac{1}{2} \frac{e_j\left(\vec{\mu}_i \cdot \vec{r}_{ij}\right)}{r_{ij}^3} \left(1 - L_O(r_{ij})\right) \tag{3}$$

which is composed by the sum of the Coulomb contribution of all oxygen, by a screening function $(1 - L_O[r])$, and the polarizable dipole moment on each of the oxygen particles $\vec{\mu}_i$ ({xyz}-format), calculated from the electric field of that given particle due to neighboring charges from all other particles in the system. Further details on the model can be found in the work of Stillinger and David.[20] It is worth noting that each oxygen atom in this model is regarded as a doubly charged anion O^{2-}, with a polarizability, α, of 1.444Å^3 chosen to match water vapor-phase measurements.[41]

In regards to the molecular dynamic simulation set up, the velocity Verlet scheme was chosen to integrate the equation of motion with a timestep of 0.48887 fs in NVE ensemble, while the RETIS shooting moves have been within a NPT ensemble with a kinetic energy randomly given according to a Boltzmann distribution centered at 300 K. System energy, order parameters and collective variables were calculated every 3rd MD step of the simulation and stored every 1000 cycles (one RETIS cycle consists of a Monte Carlo move for each ensemble). Path sampling simulation was initialized with a series of Monte Carlo moves from an equilibrated initial state (Figure 3). The transition was sampled with 4×10^6 RETIS cycles. Shooting, time-reversal and swapping moves were used with a ratio of 25–25–50, respectively. The simulations were performed on in-house software based on Fortran90, named TISMOL: the library which inspired the open source code PyRETIS. (TISMOL and PyRETIS have a similar output format).

The main order parameter, Θ (OP_1), was selected to represent the difference between $\angle(O_2 - O_1 - O_3)$ and $\angle(O_1 - O_3 - O_2)$ of the trimer, as depicted in Figure 3. Two collective variables, OP_2 and OP_3, report the distance of the proton (H_7) from the two neighboring water molecules (O_1 and O_2). The interfaces for the RETIS algorithm were located at Θ values of: [−1.5, −1.2, −1.0, −0.9, −0.8, −0.7, −0.6, −0.65, −0.6, −0.55, −0.5, −0.45, −0.40, −0.35, −0.30, −0.25, −0.1, 1.5].

3.2 | Results and discussion

The present study of the proton transfer in a water trimer reports the mechanisms of the reaction qualitatively and quantitatively. In Figure 4, three snapshots of a reactive trajectory are provided to facilitate the visualization of the transition: at the initial stage, on the transition barrier and at the reactant state.

In Table 2, we report the computed flux, crossing probability and rate for the proton transfer reaction.

In the work of Geissler et al.,[22] a rate is computed for a large set of initial energies. As a consequence, the authors report reaction ranges with differences of 25 orders of magnitude. In a subsequent investigation, Geissler et al.[24] predicted a reaction rate in the order of 10^8 s^{-1} by using DFT and transition state theory (TST) at ambient temperature. DFT has known limitations,[42] and TST is mostly regarded as a qualitative method to investigate reaction mechanisms.[43] It should be noted that the rare event approach we used in this study (replica exchange transition interface sampling[10]) is based on Boltzmann distribution of kinetic energies around a given temperature (new paths are generated according to a NPT ensemble). Hereby, the present results integrate the previous finding quantitatively while

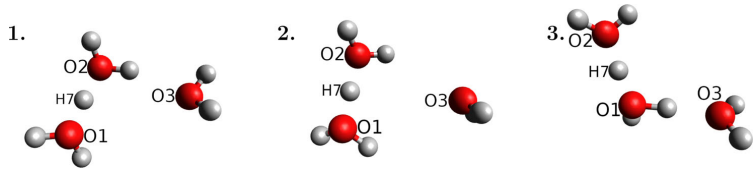


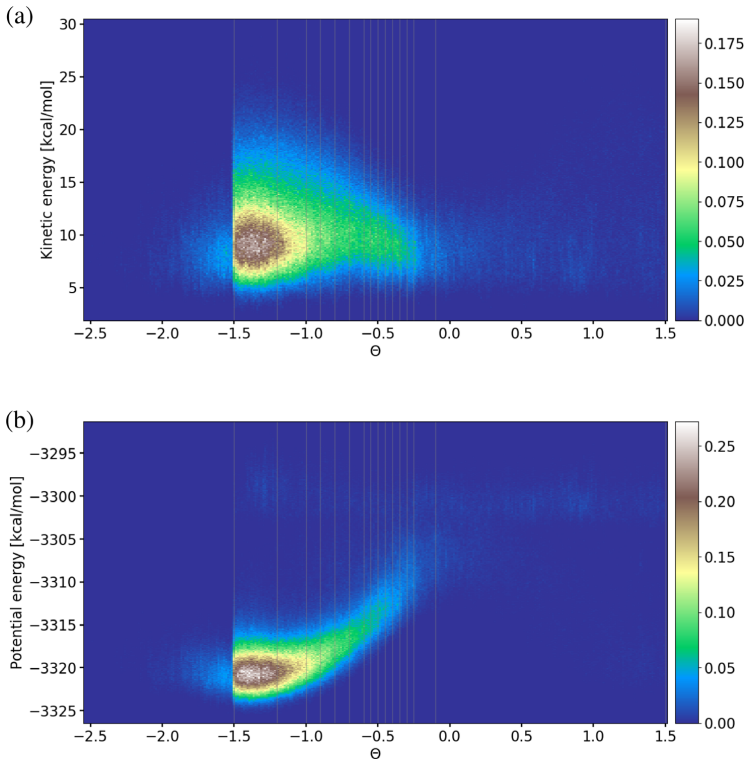
FIGURE 4 Illustration of a reactive pathway generated during the simulation, with snapshots taken along the way. Frame 1 shows the system at the initial state, $\Theta = -1.5$. At frame 2, the system is around the saddle point of the transition, with a value of $\Theta \approx 0.0$, and the transferring proton equidistant from the two neighboring oxygen atoms. At frame 3, the system has reached the defined end state, $\Theta = 1.5$. The length of the path between states *A* (frame 1) and *B* (frmae 3) is $\sim$ 500 time steps long ($2 \sim 50$ fs)

TABLE 2 Table of results from the RETIS simulation using the Stillinger–David potential

	Symbol	Value	[unit]	Error (%)
Flux out of region *A*	f_A	3.2293×10^{12}	s^{-1}	0.25309
Crossing probability	$P(\lambda_{\mathcal{B}}\|\lambda_A)$	1.42986×10^{-9}	—	6.80686
Rate of reaction	k_{AB}	4.61742×10^{3}	s^{-1}	6.81156

Note: Crossing probability—The probability of reaching state B from state A. Flux—The frequency of crossing the interface used to define state A. The reaction rate is obtained by multiplying the crossing probability by the flux. The relative (%) error is calculated via block averaging.

FIGURE 5 Frame density for all ensembles plotted as a function of the kinetic energy, panel (A), and potential energy, panel (B), versus the main order parameter



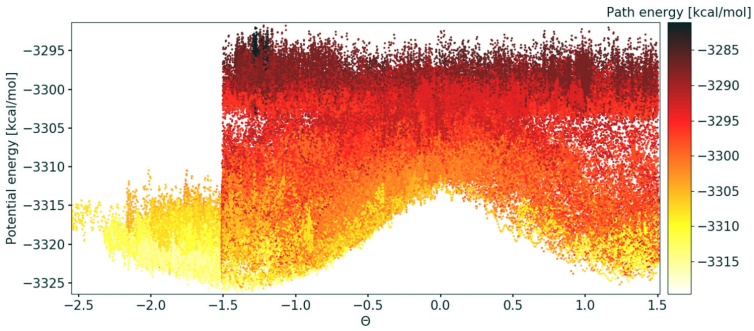


FIGURE 6 Scatter plot: The potential energy for all ensembles is plotted as a function of the main order parameter. Colors indicate the system energy, which is constant within a path

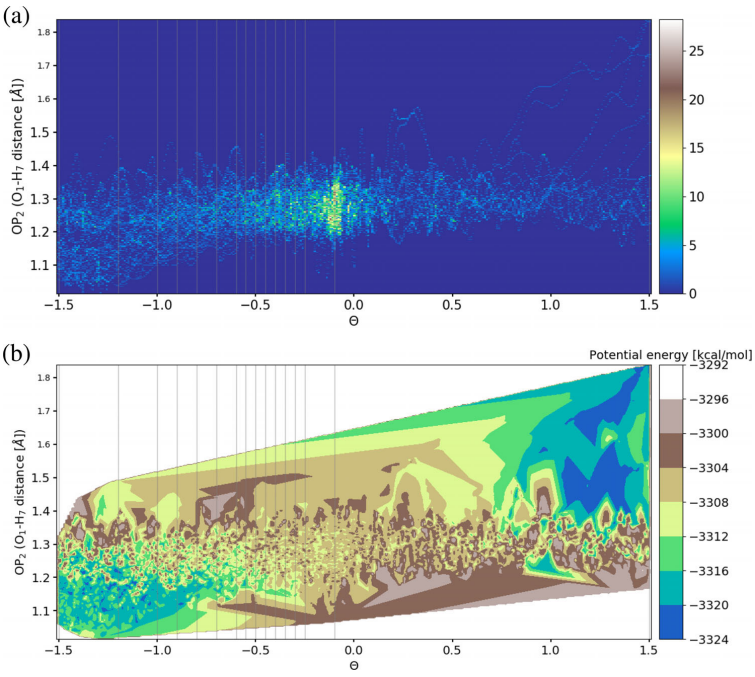


FIGURE 7 (A) Frame density of reactive trajectories plotted along Θ and OP_2 (O_1—H_7 distance). (B) Potential energy surface in a scale from blue (lower energy) to brown (higher energy). Vertical lines indicate interface planes for ensembles. Both plots are composed by trajectories collected in the [14^+] ensemble

simultaneously accomplishing the main aim of the work, which is to illustrate PyVisA features. We emphasize here that, for a more accurate description of the proton transfer mechanism, RETIS with more advanced potential models should be considered (i.e. Neural Network Potentials at Coupled Cluster Accuracy.[28]).

3.2.1 | Projected energy landscapes

Figure 5(A),(B) report the potential and kinetic energy densities. In both plots the initial stable state is clearly identified in the region θ = (−1.5, −1.25). Figure 5(A) reports a progressive reduction of the variance of the kinetic energy as a function of the main order parameter. While in the initial region there is a wider distribution of kinetic energy between paths, its average and the shape of its distribution remains consistent along the transition. The potential energy frame density plot, depicted in Figure 5(B), shows instead a clear increase of potential energy around the transition region. It should be noted that the frame density is much lower when a path is "descending" from a system's potential energy maximum.

In Figure 6, the potential energy is plotted as a function of the main order parameter. The color for each path is assigned according to its energy, which is constant within a path. On a scatter plot, each data point is a circle whose radius can be controlled by the resolution setting. In case of overlap, the latter printed points will cover previous points. From Figure 6, it appears that two transitions seem possible.

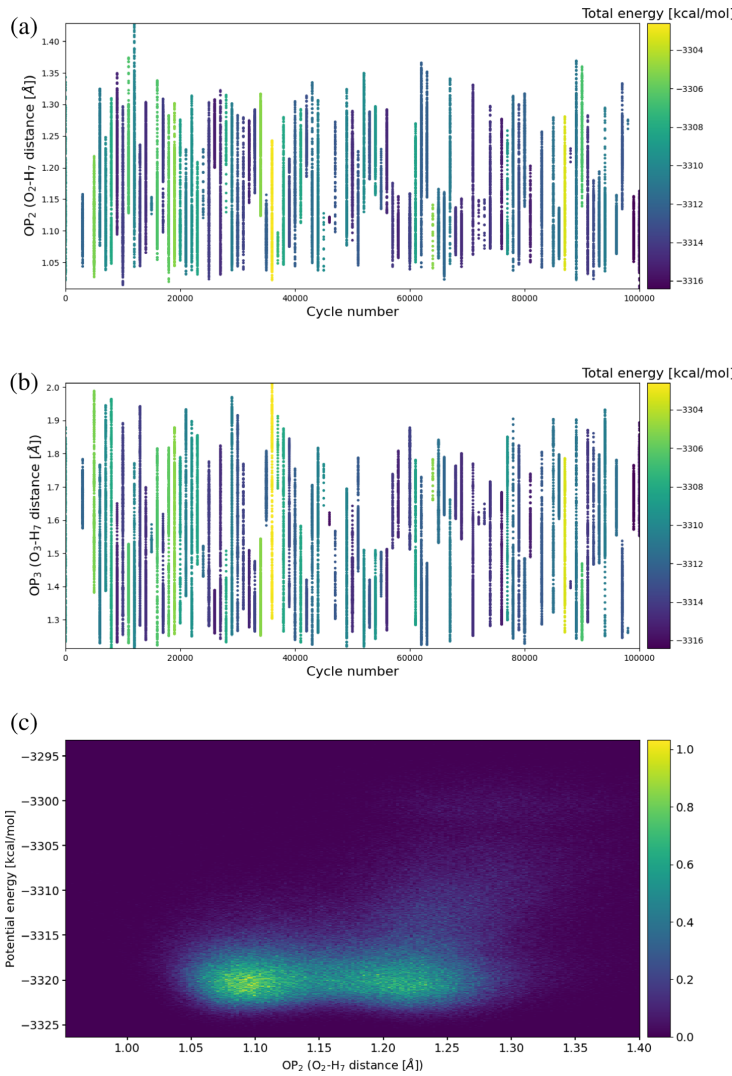


FIGURE 8 Panel a and B reports the energy plots (kinetic plus potential) as function of cycle number and O–H distance for the [0^{-}] ensemble. Plot a refers to OP_2 and plot B to OP_3 O—H distance. Panel C reports, for all the ensembles, the density plot of the potential energy as a function of OP_2 (which is mostly symmetrical to OP_3)

For paths with higher energy, a transition on a nearly flat potential energy surface appears, while the paths with a lower energy experience a potential energy landscape with a maximum of the potential energy around the transition region of roughly 14 kcal/mol higher than the initial state, which is consistent with the previous study.[25]

3.2.2 | Collective variables

To achieve a comprehensive description of the reaction mechanism, we considered different collective variables to study the transition.

Figure 7 reports the frame density for reactive trajectories and the potential energy for all trajectories as a function of the main order parameter and the OP_2 in the ensemble [14^{+}]. The selection of reactive trajectories allows a better visualization of the transition region. Figure 7(A) shows a slightly higher frame density in the area around Θ = 0.0 and OP_2 = 1.3. In the same region, as shown in Figure 5, a potential energy maximum is located. The projection of the system potential energy along the transition and the OP_2 collective variable, displayed in Figure 7(B), also confirm the presence of the maximum around Θ = 0.0. In the plot, the areas with low potential energy are reported in blue, while the regions with higher potential energy are light brown.

The two stable states, at OP_1 = −1.5, −1.0 and OP_1 = 1.0, 1.5 have similar energies, while the potential energy maximum, around OP_1 = 0.0, is roughly 16–20 kcal/mol higher. With NVE dynamics and for a system with few degrees of freedom, slower dynamics in the surrounding of the maximum of the potential energy can be expected, which results in the higher frame density registered in Figure 7(A) around Θ = 0.0. Consistently, in the adjacent regions at lower potential energy, the kinetic energy increases and the system quickly evolves towards either the reactant or the product state. It can be noted that maintaining the initial configuration while increasing only the distance of the proton requires a much higher energy than passing through the symmetrical saddle point first: the proton swaps, therefore, between water molecules on which it is bonded as a result of the change in the oxygen configuration. For path at higher energy, the oxygen configuration can be held during the proton transition.

The dynamics of minor transitions occurring in the system can also be monitored via collective variables. It should be noted that, when computing their relative occurrence, the weighted ensemble approach should be adopted.[44]

In Figure 8(A),(B), the O—H distances considered in OP_2 and OP_3 are reported as a function of the cycle number for the [0^-] ensemble. The two collective variables oscillate, nearly symmetrically, around two discretely separated steady states, located around 1.1 Å and 1.25 Å. For the symmetry of the system, it can be noted that the transitions of OP_2 are analogous, but reversed, compared to the transition shown by OP_3. The paths with higher energy are, most commonly, the ones capable of visiting both intermediate states.

Figure 8(C) shows that two regions are similarly visited around OP_2 = 1.1 Å and 1.25 Å. It can also be noted that the proton transfer occurs with a O–H distance around 1.25 Å, consistently as noted in Figure 7.

Flipping vibrations have been previously identified with higher resolution methods (DFT).[45] As also shown by our results, the distribution of the hydrogen position is bi-modal. This has been claimed to

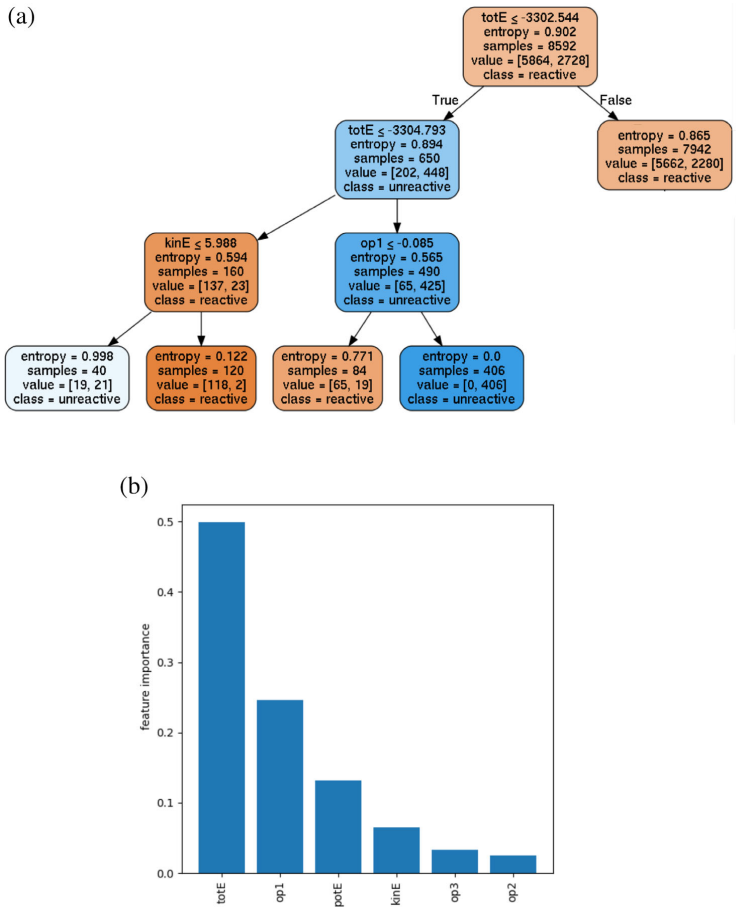


FIGURE 9 Classifiers: Panel (A) reports a decision tree while panel (B) lists the descriptors (features) with their relative importance in determining path reactivity. The histogram is obtained by a random forest of decision trees

generate large deviations in the rate computation when considering reduced dimensionality approaches.[46] This vibration can be mapped, as an independent degree of freedom, in the initial and final stable states, as well as during the reaction since it can induce a Fermi resonance[26,27] that might significantly alter the reaction dynamic. To assert if the hydrogen distances correlate with the reaction, classifiers can be employed.

With a decision tree, and a random forest of decision trees[36] the sequence of features and their relative importance to each can be quantified. By selecting the frames around the transition region, OP_1 = (−0.1, 0.1) to be fed to the classifiers, the relative importance of each descriptor for the reaction can be assert. Figure 9, panel A and panel B, reports the generated decision tree with three levels and the importance measure obtained from a random forest of decision trees, respectively. At the transition, the total energy appears to be the discriminant between reactive and unreactive trajectories. The decision tree indicates that a frame with a total system energy lower than 3305.793 kcal/mol located prior to OP_1 = −0.085, is part of an unreactive path. At the opposite end, the highest probability for a frame to be part of a reactive path, given that it is included in the OP_1 = (−0.1, 0.1) region, is when it has a total energy lower than 3304.793 kcal/mol and a kinetic energy higher than 5.988 kcal/mol. The random forest confirms that total energy is the most important feature for frames included in the OP_1 = (−0.1, 0.1) range to determine if the path is reactive. In the transition region, it appears that the OP_2 and OP_3 distances have little or no influence on determining the path reactivity. While in the present work a simple feature discrimination between paths is reported, we strongly believe that decision trees can provide further mechanistic insight into transitions if properly generated and customized to the system of interest. A detailed method development and a generalization on how classifiers can be used to predict the importance of features not considered a priori is reported in the work of Roet et al.[47]

4 | CONCLUSION

A transition is well described when the minimum free energy path is located on the free energy surface between an initial and a final state. PyVisA has been designed to simplify such search, facilitating path sampling set up and its output postprocessing.

PyVisA is a component of the PyRETIS library and incorporates a large set of relatively simple functionalities. While the tool has been developed to manipulate PyRETIS outputs, the self-standing structure facilitates its usage with other software (e.g. OPS[17]) and allows a user to execute various analyses on the fly.

PyVisA has been designed taking into consideration the most common computational set up in clusters, where large amounts of data are produced and stored. The library allows the recalculation of collective variables and data compression from the command line, without requiring the installation of the GUI module and its requisites. The generated compressed binary file (or the PyRETIS output) can be directly loaded by the GUI module. This solution permits a quick visualization of the various descriptors in a local machine even when the main simulations are performed in a cluster.

In the postprocessing of path sampling simulations, the statistical correlation between observables in time and in path space can guide the interpretation of the sequence of events in a transition (commonly described as transition mechanism) by identifying eventual metastable states and related observables in transition regions. The evaluation of different collective parameters can also allow a systematic employment of the predictor analysis,[19] providing reaction probability given a certain system status and a mechanistic description of the transition with the usage of classifiers.[47]

We used the PyRETIS library to compute the rate of the proton transfer reaction in a water trimer. We then used PyVisA to unveil the mechanism of the reaction of the proton transfer. The initial states and the potential energy barrier were mapped and the transition studied by different energetic and space dependent descriptors (O–H distances). Our contribution completes the investigation with the basic SD model on the mechanism and rate of the proton transfer reaction in a small cluster. We have herein reported a methodological advancement well-suited for studying the proton transfer reaction. A deeper understanding and quantification of the proton transfer can be obtained by using more advanced models, such as Neural Network Potentials at Coupled Cluster Accuracy reported by Schran et al.[28]

PyVisA adheres to FAIR and Open Data policies.[48] A compressed open format has been shared providing a quick access to our data for external studies.[49]

4.1 | Software and data availability

PyVisA is an analysis and visualization software developed for path sampling in general and PyRETIS 2 in particular. The software can be freely downloaded from gitlab.com/pyretis/pyretis. It is an open Python library released under a LGPLv2.1+ license. A "light" version of PyVisA is included in PyRETIS 2 (from version 2.3) with no additional requirements. To increase the PyVisA compatibility with clusters, Qt5[50] has been included as an optional requirement, and it must be installed separately for the *Visualizer* / GUI component to function. To install the full version of PyVisA, scikit-learn[33,34] and graphviz[51] shall also be installed. Avogadro[32] is required only to visualize full trajectories. These packages are included in the PyRETIS development requirements (requirement-dev.txt). Please refer to the PyRETIS website (www.pyretis.org) for the latest installation instructions and updates.

The TISMOL simulation setting file and the simulations output files for the water proton transfer study can be freely downloaded at http://www.pyretis.org/source_files/Water_trimer_2020/.

DATA AVAILABILITY STATEMENT

Data openly available in a public repository that does not issue DOIs

ORCID

Enrico Riccardi https://orcid.org/0000-0003-1890-7113

REFERENCES

[1] P. Tiwary, V. Limongelli, M. Salvalaglio, M. Parrinello, *Proc. Natl. Acad. Sci.* **2015**, *112*, E386.
[2] F. Paul, C. Wehmeyer, E. T. Abualrous, H. Wu, M. D. Crabtree, J. Schöneberg, J. Clarke, C. Freund, T. R. Weikl, F. Noé, *Nat. Commun.* **2017**, *8*, 1095.
[3] S. Paul, N. N. Nair, H. Vashisth, *Mol. Simul.* **2019**, *45*, 1273.
[4] M. Moqadam, E. Riccardi, T. T. Trinh, A. Lervik, T. S. van Erp, *Phys. Chem. Chem. Phys.* **2017**, *19*, 13361.
[5] M. Moqadam, A. Lervik, E. Riccardi, V. Venkatraman, B. K. Alsberg, T. S. van Erp, *Proc. Natl. Acad. Sci.* **2018**, *115*, E4569.
[6] E. Riccardi, E. C. van Mastbergen, W. W. Navarre, J. Vreede, *PLoS Comput. Biol.* **2019**, *15*, e1006845.
[7] C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler, *J. Chem. Phys.* **1998**, *108*, 1964.
[8] P. G. Bolhuis, D. Chandler, C. Dellago, P. L. Geissler, *Annu. Rev. Phys. Chem.* **2002**, *53*, 291.
[9] B. Peters, *Mol. Simul.* **2010**, *36*, 1265.
[10] T. S. van Erp, D. Moroni, P. G. Bolhuis, *J. Chem. Phys.* **2003**, *118*, 7762.
[11] T. S. van Erp, *Phys. Rev. Lett.* **2007**, *98*, 268301.
[12] E. Riccardi, O. Dahlen, T. S. van Erp, *J. Phys. Chem. Lett.* **2017**, *8*, 4456.
[13] G. van Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam; **1995**.
[14] A. Lervik, E. Riccardi, T. S. van Erp, *J. Comput. Chem.* **2017**, *38*, 2439.
[15] E. Riccardi, A. Lervik, S. Roet, O. Aarøen, T. S. van Erp, *J. Comput. Chem.* **2019**, *41*, 370.
[16] M. C. Zwier, J. L. Adelman, J. W. Kaus, A. J. Pratt, K. F. Wong, N. B. Rego, E. Suárez, S. Lettieri, D. W. Wang, M. Grabe, D. Zuckerman, L. Chong, *J. Chem. Theory Comput.* **2015**, *11*, 800.
[17] D. W. Swenson, J.-H. Prinz, F. Noe, J. D. Chodera, P. G. Bolhuis, *J. Chem. Theory Comput.* **2018**, *15*, 813.
[18] P. Bolhuis, C. Dellago, *Eur. Phys. J. Special Top.* **2015**, *224*, 2409.
[19] T. S. van Erp, M. Moqadam, E. Riccardi, A. Lervik, *J. Chem. Theory Comput.* **2016**, *12*, 5398.
[20] F. H. Stillinger, C. W. David, *J. Chem. Phys.* **1978**, *69*, 1473.
[21] F. Stillinger, C. David, *J. Chem. Phys.* **1980**, *73*, 3384.
[22] P. Geissler, C. Dellago, D. Chandler, *Phys. Chem. Chem. Phys.* **1999**, *1*, 1317.
[23] I. Zhyganiuk, Generalized stillinger–david potential, arXiv preprint arXiv:1202.4491; **2012**.
[24] P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, M. Parrinello, *Chem. Phys. Lett.* **2000**, *321*, 225.
[25] P. L. Geissler, T. Van Voorhis, C. Dellago, *Chem. Phys. Lett.* **2000**, *324*, 149.
[26] C. H. Duong, O. Gorlova, N. Yang, P. J. Kelleher, M. A. Johnson, A. B. McCoy, Q. Yu, J. M. Bowman, *J. Phys. Chem. Lett.* **2017**, *8*, 3782.
[27] N. R. Samala, N. Agmon, *Chem. Phys.* **2018**, *514*, 164.
[28] C. Schran, J. Behler, D. Marx, *J. Chem. Theory Comput.* **2019**, *16*, 88.
[29] W. McKinney, *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*, O'Reilly Media, Inc., **2012**.
[30] J. D. Hunter, *Comput. Sci. Eng.* **2007**, *9*, 90.
[31] M. Summerfield, *Rapid GUI Programming with Python and Qt: The Definitive Guide to PyQt Programming (paperback)*, Pearson Education, **2007**.
[32] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminf.* **2012**, *4*, 17.
[33] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, G. Varoquaux, ECML PKDD Workshop: Languages for Data Mining and Machine Learning; **2013**, pp. 108–122.
[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825.
[35] S. Wold, K. Esbensen, P. Geladi, *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37.
[36] T. K. Ho, Proceedings of the 3rd International Conference on Document Analysis and Recognition; **1995**, pp. 278–282.
[37] J. MacQueen et al., *Proc. Berkeley Symp. Math. Stat. Prob.* **1967**, *1*, 281.
[38] Z. Zhang, F. Murtagh, S. V. V. Poucke, S. Lin, P. Lan, *Ann. Trans. Med.* **2017**, *5*, 75.
[39] F. Pernkopf, D. Bouchaffra, *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1344.
[40] U. Von Luxburg, *Stat. Comput.* **2007**, *17*, 395.
[41] D. Eisenberg, W. Kauzmann, *The structure and properties of water*, Oxford University Press on Demand, **2005**.
[42] A. J. Cohen, P. Mori-Sánchez, W. Yang, *Science* **2008**, *321*, 792.
[43] R. Cabriolu, K. M. Skjelbred Refsnes, P. G. Bolhuis, T. S. van Erp, *J. Chem. Phys.* **2017**, *147*, 152722.
[44] G. A. Huber, S. Kim, *Biophys. J.* **1996**, *70*, 97.
[45] M. Losada, S. Leutwyler, *J. Chem. Phys.* **2003**, *119*, 304.
[46] J. Rheinecker, T. Xie, J. M. Bowman, *J. Chem. Phys.* **2004**, *120*, 7018.
[47] S. Roet, E. Riccardi, D. Daub, Chemistrees: Using decision trees to gain insight into reaction triggers for the deprotonation of formic acid; **2020**.
[48] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons, *Sci. Data* **2016**, *3*, 160018.
[49] E. Riccardi, S. Pantano, R. Potestio, *Interface Focus* **2019**, *9*, 20190005.
[50] L. Z. Eng, *Qt5 C++ GUI Programming Cookbook*, Packt Publishing Ltd, **2016**. https://books.google.no/books?hl=en&lr=&id=Rc6qDQAAQBAJ&oi=fnd&pg=PP1&dq=Qt5+C%2B%2B+GUI+Programming+Cookbook&ots=U7hdpZx33c&sig=yxez5kW5XAYa9Fu3vSbptcN3VWM&redir_esc=y#v=onepage&q=Qt5%20C%2B%2B%20GUI%20Programming%20Cookbook&f=false.
[51] J. Ellson, E. Gansner, L. Koutsofios, S. C. North, G. Woodhull, *Graphviz—open source graph drawing tools. International Symposium on Graph Drawing*, Springer, Berlin, Heidelberg **2001**, p. 483.

How to cite this article: Aarøen O, Kiær H, Riccardi E. PyVisA: Visualization and Analysis of path sampling trajectories. *J Comput Chem*. 2021;42:435–446. https://doi.org/10.1002/jcc.26467

Paper III

RSC Advances

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Adv.*, 2021, **11**, 8730

Exploring the effects of approach velocity on depletion force and coalescence in oil-in-water emulsions

Ola Aarøen,*[a] Enrico Riccardi[bc] and Marit Sletmoen[a]

Received 25th January 2021
Accepted 18th February 2021

DOI: 10.1039/d1ra00661d

rsc.li/rsc-advances

An emulsion is a thermodynamically unstable system consisting of at least two immiscible liquid phases, one of which is dispersed in the other in the form of droplets of varying size. Most studies on emulsions have focused on the behaviour of emulsion droplets with diameter from ~50 μm and upwards. However, the properties of smaller droplets may be highly relevant in order to understand the behaviour of emulsions, including their performance in numerous applications within the fields of food, industry, and medical science. The relatively long life-time and small size of these droplets compared to other emulsion droplets, make them suited for optical trapping and micromanipulation technologies. Optical tweezers have previously shown potential in the study of stabilized emulsions. Here we employ optical tweezers to examine unstable oil-in-water emulsions to determine the effects of system parameters on depletion force and coalescence times.

Introduction

Emulsions play an important role in a wide range of fields, with applications from oil processing and recovery,[1] to the food industry,[2,3] and pharmaceutical manufacture.[4] Emulsions arise from the forced mix of immiscible liquids in a multiphase system, by breaking one or more liquids into smaller droplets dispersed in a bulk phase. Their types and classifications vary, dependant on composition, use, and stability. Nomenclature generalizes them as O/W for an non-polar liquid (*i.e.* oil) dispersed in a polar (*i.e.* water) phase, or in the opposite case W/O. Recently, a series of more complex emulsions have been developed. One of these is double emulsions, also called multiple emulsions, and the properties of these emulation, including preparation methods[5] and stabilisation[6] have been reviewed. Widely used in the food industry, their thermodynamic instability, resulting in gradual separation into two pure phases, is the main challenge for product shelf-life demands. Different mechanisms for phase-separation exist but they can be categorized into the following main groups:

(1) Sedimentation/creaming due to gravitational forces and difference in phase densities.

(2) Flocculation/aggregation due to low range attractive droplet–droplet interactions.

(3) Diffusion and re-precipitation over interfaces, know as Ostwald ripening.

(4) Coalescence, where droplets or bubbles of the dispersed phases merge, usually aided by mechanism (1) and (2).

Measures can be taken to reduce the rate of separation of emulsions, by addition of stabilizing agents to counteract some of the mechanisms driving separation. These emulsifiers, or surfactants, stabilize the dispersed phase by reducing the interfacial tension over the oil–water interface, and form a layer of adsorbed molecules that prevent aggregation of droplets and subsequent coalescence. Conversely, clarifying agents, or flocculants, reduce emulsion stability by inducing droplet flocculation to separate liquid phases.

Even with no stabilizing agents present, OH^- ions originating from water autoprotolysis or ionic additions to solution adsorbs to the oil–water interface, forming a negatively charged layer.[7] Subsequently, a positively charged layer forms, surrounding the negative, forming a double layer on the interface, as illustrated in Fig. 1 (with Na^+ and Cl^- ions). This electrical double layer will result in a repulsive force between neighbouring droplets, reducing droplet contact and stabilizing the emulsion even in the absence of surfactant. The double layer extent is given by the distance from the interface to a point away from interface, where the charge density is equal to that of the bulk solution. The theoretical extent of this layer, κ^{-1}, from the droplet interface is given by

$$\kappa^{-1} = \sqrt{\frac{\varepsilon_r \varepsilon_0 k_B T}{2 \times 10^3 N_A e^2 I}} \quad (1)$$

where ε_r and ε_0 is the permittivity of water and vacuum, respectively, $k_B T$ the Boltzmann energy at temperature T, N_A is

[a]Department of Biotechnology and Food Science, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway. E-mail: ola.aaroen@ntnu.no

[b]Department of Chemistry, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway

[c]Department of Informatics, UiO, Gaustadalléen 23B, 0373 Oslo, Norway

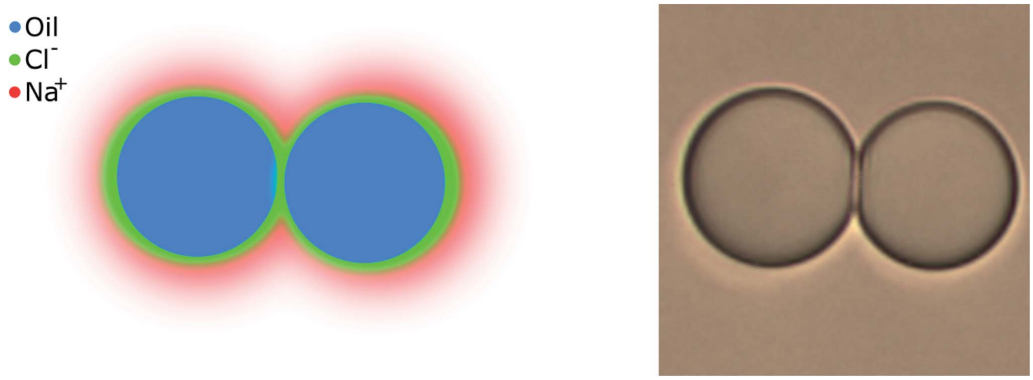


Fig. 1 Left: Illustration of double layer (*not to scale*) surrounding a pair of emulsion droplets in ionic solution. The innermost layer colored in green is characterized by the predominant presence of chlorine ions, whereas in the outer layer colored in red, sodium ions are predominant. Right: Light microscopy image of two optically trapped emulsion droplets brought into proximity by moving one of the optical traps. The image was obtained using the Zeiss Axio Observer inverted optical microscope integrated with the NanoTrackerTM 2.

Avogadro's number and e the electron charge. The ionic strength is defined as $I = \frac{1}{2}\sum_i c_i z_i^2$, with c_i being the concentration of ions and z_i the charge. Increasing the overall ionic strength in bulk solution can reduce the distance κ^{-1}. Table 1 presents the results obtained when using eqn (1) to calculate the double layer extents as a function of the ionic strength of the surrounding solution. In pure water emulsions (at 25 °C) an ionic concentration equal to 10^{-7} mol L^{-1} is assumed due to the autoprotolysis in water. When adding ions (NaCl) to concentrations equal to or exceeding 0.1 mol L^{-1}, this contribution from the autoprotolysis of water to the total ionic strength becomes negligible. Reducing the repulsive force originating from the electrical double layer is expected to lead to increased probability for spontaneous coalesce between emulsion droplets when brought to proximity. Temperature effects would also change droplet coalescence time, as increased thermal energy has been shown to decrease the interfacial tension in bulk crude oil,[8] and reduce coalescence time in micromanipulated droplets,[9] although for the small size ranges of the latter, temperature control has proved to be inaccurate.

This de-stabilization of emulsions have become important especially within the petroleum industry, where wastewater bi-product from drilling contains a variety of petroemulsions. These emulsions consist of a mix of polar crude oil components, polar hydrocarbons ("resins") and polyaromatic hydrocarbons ("asphaltenes"), which can form surfactant layers that are considered to be primarily responsible for their stability and behaviour.[10–12]

Table 1 Concentration of ions in solution and the corresponding Debye length κ^{-1}

Ions	Concentration [mol L^{-1}]	κ^{-1} [nm]
H^+ OH^-	10^{-7}	960.28
Na^+ Cl^-	0.01	3.04
	0.1	0.96
	1.0	0.30

The performance of transport and separation processes involving petroemulsions is affected by the dynamics of coalescence between droplets.[10,11,13] Dynamics of coalescence are controlled by the drainage and eventual rupture of the thin liquid film[13] that develops between approaching droplets. Thin film drainage times and film rupture behavior depend on the intermolecular forces between the molecules present at the interfaces. Solvent molecules interacting with surfactants and interfaces determine the behaviour of the thin film formed between the two interface regions encircling two approaching droplets.

It is preferable both due to environmental concerns, and commercial ones, to separate petroemulsions into the constituent phases prior to wastewater release.[14–16] This has led to the development of (green) enhanced oil recovery ((G)EOR) techniques, which are based on the injection of specific green chemicals (surfactants,[17,18] polymers or particles,[1] gas,[19] *etc.*) that effectively displace oil because of their phase-behavior properties, which decrease the interfacial tension (IFT) between the displacing liquid and the oil.

During the past decade, extensive experimental campaigns and modelling efforts have been undertaken to characterize the rheological, thermodynamic, and structural properties of interfacial layers composed of resin and asphaltene surfactants. Experimentally, crude oils containing a myriad of indigenous resin and asphaltenes have been studied.[10] In computational studies, the interface induces limitation in the molecular simulation predictions. Firstly, common force fields are developed assuming an homogeneous environment (bulk).[20,21] Thereafter, surfactants coming from oil sources are composed by a wide distribution of different amphiphilic molecules. Model oils employing complex synthetic (and synthesizable) surfactants[22] have been employed[23,24] to successfully, but approximately describe the interface region under different ionic conditions.

The current theoretical models for film drainage during coalescence focus primarily on systems employing relatively simple descriptions of surfactant behaviors. These systems are based on London dispersion forces in the case of oil continuous emulsions and Derjaguin, Landau, Verwey and Overbeek (DLVO) theory for water continuous emulsions. Despite the studies trying to unify the macroscopic behavior and the underlying molecular mechanisms,[25,26] a comprehensive understanding is still lacking.

Recent investigations have shown the importance of describing the structural and conformational rearrangements at molecular level.[27,28] These structural rearrangements occur in the inter-phase region as a result of altered conditions (surfactant concentration and type, ionic strength and type) in the thin film generated between two drops. As a result, the current continuous models have a limited validity in predicting the overall system behavior by adopting a simplistic description of the disjoining pressure and film rupture mechanisms. In emulsions, and in particular in petroemulsions, there is a lack of agreement between the mechanisms considered in state-of-the-art models for film drainage during coalescence[13,29] and the mechanisms responsible for the stability behavior of petroemulsions that have been suggested by experimental studies.[10,30]

As explained above and also recently stated by Langevin[31] coalescence in emulsions and foams is far from being understood, despite many years of investigations. In addition to the properties described above, also other inherent properties of the coalescence process may explain the still unsatisfactory understanding of this phenomenon. The coalescence process is characterized by being extremely rapid. Furthermore, the rupture of the thin films separating the droplets or bubbles prior to coalescence is influenced by several properties of the system, including hydrodynamics, surface rheology, surface forces, and thermal fluctuations. The rapidness of the coalescence event as well as its dependence on a high number of factors makes experimental studies challenging. Experimental challenges may also explain why most previous studies have focused on oil droplets in the size range of 20–200 μm. The lack of attention to the smaller droplets is problematic since the emulsion droplets found in many applications have a diameter below 10 μm, and these small droplets are known to show a behavior that deviate from this of the larger droplets, both related to deformation, internal pressure and hydrodynamic effects.[32] Filling this gap and providing information on the properties of the small droplets is therefore important.

During the past decades, new methods have been developed that trap small particles and even single molecules while at the same time applying and measuring forces acting on them. Among these, atomic force microscopes, optical tweezers, and magnetic tweezers have enabled the study of a wide range of molecular processes in which force plays a crucial role.[33] These tools have also been continuously refined in order to offer frequent improvements in spatial and temporal resolution. These tools are therefore today recognized as interesting tools when addressing questions related to a wide variety of processes, including droplet coalescence, as further explained in the following.

The manipulation of microparticles with light was first demonstrated by Ashkin.[34] His experiments revealed that radiation pressure originating from a beam of light would impart momentum to particles and thus push them in the direction of light propagation. He later showed that by introducing an additional force counteracting the force acting in the propagation direction of the light, the particle can be trapped in a small volume in space.[35] This is referred to as an optical trap and can be described as a simple harmonic oscillator potential,[36] within which a particle can diffuse by Brownian motion.[37] An optical trap can be created by the use of a laser and a lens creating a tightly focused laser beam. While optical trapping can be used simply to position particles, the method is also often used to quantified forces on microscopic length scales. Optical tweezers set-ups can be used to exert forces in excess of 100 pN on particles ranging in size from nanometers to micrometers while simultaneously measuring the three-dimensional displacement of the trapped particle with sub-nanometer accuracy and sub-millisecond time resolution. The forces are determined by either a sensitive photodiode or a high-speed camera that reports the displacement of a particle from the center of an optical trap of known spring constant. The spring constant is usually obtained from studying the restriction of the Brownian motion of the particle.[38]

Optical tweezers have been used in a range of research fields,[39] but despite the obvious potential of optical tweezers to provide new insight into the coalescence process of micro-droplets, the number of studies published in this field is still limited. In 2006, Ward and coworkers published a study[40] demonstrating the deformation of micron-sized emulsion droplets having ultralow interfacial tensions, by optical tweezers. A few years later, Bauer and coworkers trapped one emulsion droplet in each of two optical traps and quantified the pH dependent interactions between the individual emulsion droplets.[41] A similar approach was used by Nilsen-Nygård and coworkers[42] who obtained force *versus* distance curves of emulsion droplets of both micro- and macromolecular stabilization. These studies were followed by publications from Chen and coworkers. In their first study they reported the interaction forces between tetradecane droplets with a diameter equal to 5.0 μm in water, as a function of varying SDS and NaCl concentration.[43] In the following paper they turned to oil droplets stabilized by a non-ionic surfactant[44] before focusing on a system of switchable surface-active colloid particles.[45] Otazo and coworkers performed studies with a different aim. They focused on the aggregation and coalescence of partially crystalline emulsion droplets.[46] In the current paper we use a dual trap optical tweezers set-up to investigate the effects of approach velocity on depletion forces and coalescence times[47] in an oil-in-water emulsion.

Experimental

Materials

Emulsions were made using dodecane oil, $C_{12}H_{26}$ (Sigma-Aldrich, St. Louis, USA) in de-ionized water (MilliQ, <18 Ω). The ionic strength of the emulsions was regulated by adding

small volumes of a solution of high ionic strength. This solution was made by dissolving NaCl (Sigma-Aldrich, St. Louis, USA) in de-ionized water.

Preparation of O/W emulsions

The O/W emulsions were made using the following three alternative methods: the first method was based on the use of a microfluidic cell. In order to produce emulsions with droplet sizes in a suitable range, a cross-junction[48] cell with channel width (nozzle) of 15 μm was fabricated using soft lithography. A second method explored was based on the use of an Ultra Turrax homogenizer (IKA, Straufen, Germany). This is a mechanical stirrer, and in the current project it was operated at 3000 rpm for 5 minutes. The last method explored for the fabrication of emulsions included a simplified homogenization method based on mixing the materials in an Eppendorf tube and shaking for 3 minutes on a variable speed vortex mixer (Fisher Scientific, Pittsburg, US) at 2000 rpm.

In order to promote droplet coalescence in the OT sample chamber, the ionic strength of the emulsion solutions were increased prior to the measurements. Immediately after emulsification, small volumes of 1 mol L^{-1} NaCl solution was added to the sample in order to increase the ionic concentrations to pre-defined values (0.05, 0.1 and 0.15 M). These emulsions were transferred to fluid cells, made as described below, which were immediately mounted in the microscope set up.

Fluid cells were made with a circular Borax glass (35 mm, thickness no. 1, VWR, Pennsylvania, US) and a rectangular cover glass (22 × 50 mm, thickness no. 1, VWR, Pennsylvania, US), separated by two layers of double sided tape or melted parafilm. The spacing between the rectangular pieces of tape or parafilm was in the range 0.5–1.0 cm. This lead to sample cells having an internal height of 100–200 μm, and a total sample volume of 20–30 μL. After filling, cells were sealed with nail polish.

Optical tweezers instrumentation

The OT instrument used was the NanoTrackerTM 2 optical tweezers instrument (JPK Instruments, Berlin, Germany) mounted on a Zeiss Axio Observer Inverted optical microscope. The instrument is equipped with a TEM00 laser with 3 W maximum power and option of a dual beam mode with scalable split-ratio. The laser has a Gaussian beam profile and the two traps can be controlled independently of each other. A quadrant photodiode placed at the back-focal plane of the condenser detects the displacement of the trapped beads. The emulsion solutions were diluted in order to obtain a suitable density of droplets in the sample chamber. The probability of an additional droplet being trapped in the optical trap increases with increasing droplet density, and the droplet density should therefore be kept low. However, the probability of trapping a droplet with suitable size increases with droplet density. The droplet density chosen was a compromise between these two restrictions. Prior to all measurements, one droplet was trapped in each of the two optical traps. The trap stiffness of each trap was calibrated from power spectra obtained by tracking the 3D Brownian motion of the beads.

The approach velocity of the optical trap holding an emulsion droplet was controlled. In experimental series aiming at determining the depletion force, the droplets were brought into contact and were then maintained in contact for a duration of 1 second before retracting one of the droplets, as shown in Fig. 1. The forces acting on the trapped droplets were recorded throughout both approach, pause and retract periods. In experimental series aiming at determining coalescence times, the two droplets were left in contact for a duration of 20 seconds prior to retraction.

Determination of depletion forces

Experiments were performed in order to determine the depletion force acting between oil droplets upon their separation. When performing these experiments, the two optical traps were initially positioned at known separation distance, and set to approach each other at fixed velocity until the surfaces of the two trapped droplets were in contact. The traps where then kept at this position for 1 second. Thereafter, the two traps were moved apart, while continuously recording the forces acting on each of the two droplets, including attractive depletion forces. The recorded forces were plotted as a function of droplet separation distance, giving force *versus* distance curves. As part of the data processing, the retract curve was smoothed using a Savitzky–Golay filter (bins: 200, order: 3), and subtracted from the

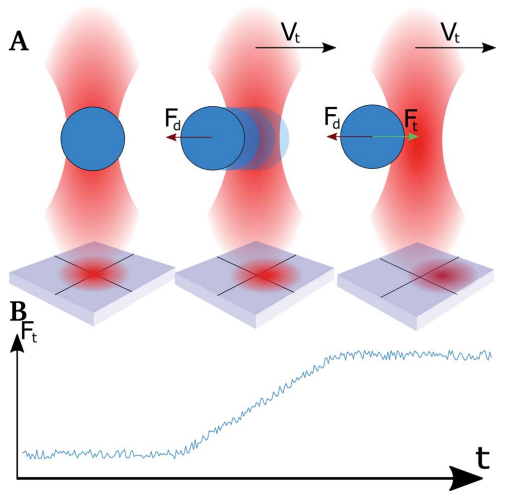


Fig. 2 Schematic illustration of droplet position with respect to the optical trap center and corresponding force output at high droplet velocity (V_t). (A) Left illustration: droplet position in optical trap when idle. Middle illustration: translocation of droplet with respect to the center of the optical trap during the initial phase of high velocity movement of trap. Right illustration: position of the oil droplet with respect to the center of the trap during high velocity movement of the trap when the drag force (F_d) is balanced by the trap force (F_t). An illustration of the QPD sensor readings is included. (B) The expected change in the drag force (F_d) output quantified over time for the events illustrated in (A).

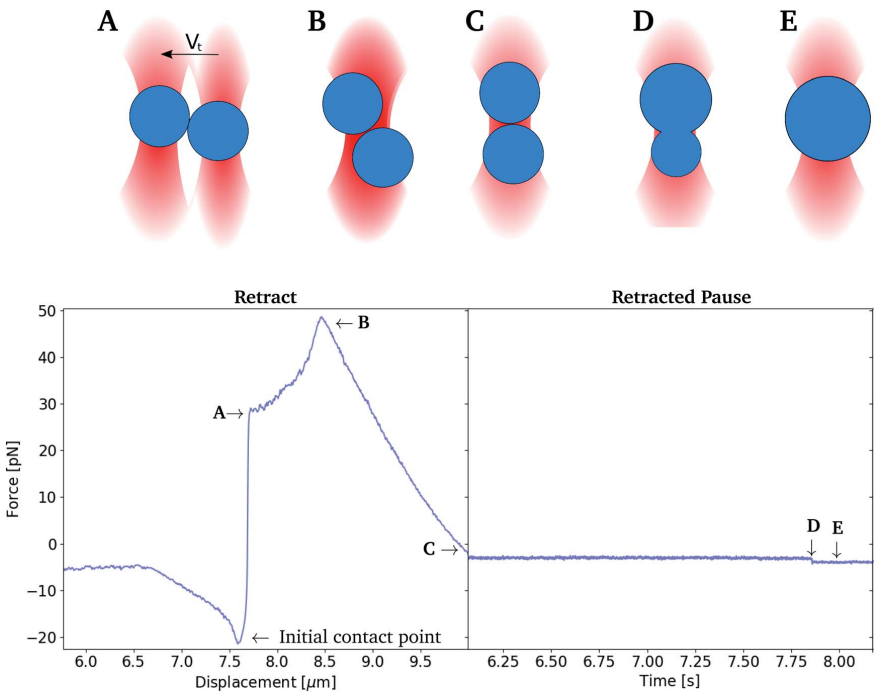


Fig. 3 Top: Schematic illustration of the positions of the oil droplets when bringing them in contact and subsequently forcing them to overlap (retract–overlap trap movements). The traps approach each other at velocity V_t, until the contact point is reached. The contact between the droplets leads to their displacement from trap centers, beginning to show in (A). (A–C) The two optical traps approach each other until they completely overlap. This leads to droplet alignment along beam axis (C), followed by droplet coalescence ((D) t_{coales}). (E) shows the resulting droplet formed after coalescence. Bottom: experimental data showing force recordings obtained during retract and retracted pause segments, with indications of droplet positions (A–E). Only a narrow section of the *X*-axis is presented in order to focus on points of interest. *Left* plot shows force against displacement of traps, initial position with a distance of 10 μm. *Right* plot shows time [s] in overlapped state.

unsmoothed extend curve, giving the difference between the forces acting on the droplets during approach as compared to retract.

Determination of coalescence times

In the current work the coalescence time, t_{coales}, was defined as the time period from the first encounter between the two oil droplets until their rupture. It was determined by bringing two droplets into contact and leaving them in contact for an extended period of time while recording the force acting on the droplets. The two droplets were approached until a point where the two optical traps fully overlapped. This procedure was chosen in order to avoid the consequences of the uncertainty in droplet size, and thus inter-droplet distance and contact point. The t_{coales} was determined with a sampling rate of 50 kHz. When analysing the force–time plots using this procedure (Fig. 3), accurate coalescence times were obtained.

Correction for drag force experienced by moving droplets

When moving the optical traps at high speed, the trapped particles were observed to lag behind the trap center.[49] This lag is caused by a drag force that acts on the droplet when it moves through a viscous medium. The strength of the drag force will depend on the velocity of the droplet as well as its size, and the viscosity of the bulk liquid.[50] At relatively high displacement speeds, in the order of 0.1 mm s^{-1} and exceeding, trapped droplets did not stabilize in a shifted position with respect to the focal point of the optical trapping beam within the time spent approaching neighbouring droplets. The Hookean nature of the traps, and the instantaneous acceleration, made droplets shift out of the trap due to the drag force acting in the opposite direction of the trap movement (Fig. 2A). This resulted in a linear rise in force (Fig. 2B) and not a plateau as expected.[51] By normalizing the retract curve to the linear increase, and shift the curve to the height at the end of the segment, a plateau was obtained which could be used as a basis for the determination of the force difference during retract–extend schemes. The difference in force between the normalized retract and extend curves was used as a measure of the forces acting between the two droplets on extension, including the attractive interaction depletion force.

Results and discussion

Evaluation of alternative emulsion preparation procedures

When studying emulsions using optical tweezers it is essential to obtain emulsions containing droplets of a size that can be trapped by the optical traps. Ideally, the distribution of droplet sizes should also not change between subsequent sample preparations.

The microfluidic method promised droplets with a predefined and reproducible size range as well as a narrow size distribution. However, in the absence of a stabilizing agent, emulsion droplets would coalesce immediately past the cell nozzle. This resulted in droplet populations characterised by a wide spread in diameters, with only a small subpopulation remaining within the wanted range by the time of measurements. Additionally, the microfluidic procedure for producing emulsions was time-consuming. The preparation of 1 ml of sample required several hours. The microfluidics based preparation approach was therefore not used in the later experiments.

Emulsions prepared using an Ultra-Turrax were observed to contain droplets ranging in size from sub micrometer to >100 μm. Compared to the microfluidic set-up, this procedure provided emulsions with higher droplet density. The samples contained droplets within the range suitable for optical trapping (1–30 μm). However, the inherent low stability of the droplets lead to a need for re-homogenization with the stirrer

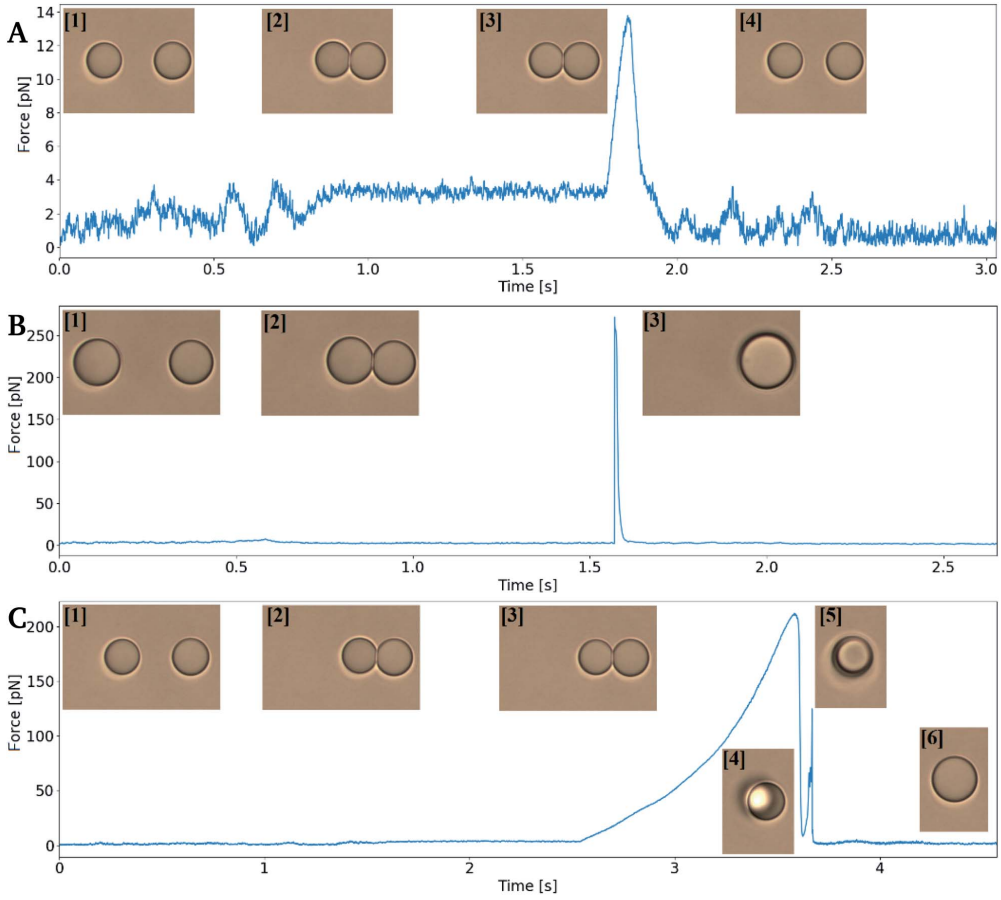


Fig. 4 Trap force *versus* time curves showing droplet behaviour during retract–extend cycles obtained using optical tweezers. Numbered inserts: images taken during the measurements. (A) Non-coalescing run, where droplets were approached ([1] → [2]), paused ([3]) and retracted ([4]) with a depletion force detected but no droplet coalescence. (B) Coalescence occurring during the pause ([2]) segment, with a spike in force resulting from the rapid displacement of the oil droplets prior to coalescence ([3]). (C) Coalescence occurring during the extend segment; droplets were kept in contact ([2] → [3]), then moved apart in the extend segment. The depletion force was sufficiently high to re-position two droplets in one trap ([5]), and finally coalescence is occurs ([6]).

between every experiment. The relatively large volume required (50 ml and upwards) to use the stirrer necessitates access to the chemicals in relatively high quantities compared to what is the requirement in other approaches.

To minimize the time used on sample preparation, while keeping the droplet diameter within the aimed range, a simplified homogenization method was developed. This method is described in the methods section and involved mixing the materials in an Eppendorf tube by vortexing. This approach provided emulsions characterised by a wide range of droplet diameters, ranging from ≈1 μm to well above the limit of optical traps. However, larger droplets would float to the top of the tube, leaving a solution in the mid-tube area containing droplets in the 1–10 μm range. Small volumes of this emulsion were removed by a syringe and used for the OT experiments. This approach did provide emulsions containing droplets of size and stability that was ideal for the intended studies.

Ionic strength effects on depletion force

In order to measure depletion force, two droplets were moved into close contact in order to obtain a contact region between the two interfaces. The droplet contact point was determined visually based on light microscopy based imaging of the droplet pairs. However, the resolution of the light microscope entails an uncertainty in the determination of the contact point, and thus also the depletion force. In case of an un-sufficient contact between the droplets, depletion will not be observed. On the other hand, pressing droplets together with excessive force could result in coalescence (Fig. 4B) or it could cause droplet to stick to each other (flocculate) and thus escape traps on extend, rather than separate (Fig. 4C) to later coalesce.

The depletion force between pairs of droplets was measured using the experimental approach detailed in previous sections. The approach and retract curves were compared by smoothing the retract curve with a moving average and subtracting the extend force curve. This resulted in a flat curve containing a peak reflecting the attractive force acting on the droplets upon their separation. The area under this peak gave the work needed to separate droplet pairs, *i.e.* the depletion work, J.

Initially, the depletion force was determined for the separation of pairs of relatively small droplets in varying bulk solutions containing NaCl. Fig. 5 shows how J increases with increasing ion concentration. Despite a variation in droplet sizes as well as a variation in approach velocity (1–2 μm s^{-1}), the effect of the ionic strength of the solution is visible. More precisely, the increased ion concentration in bulk, leading to increased difference in ion concentration between bulk and thin film, resulted in increased depletion force and work. The small values of the work needed to separate the droplet pairs shown in Fig. 5 are partly explained by the low retract–extend velocities. As previously pointed out in a theory describing the influence of external force on the dissociation of non-covalent bonds, bond survival time depends on how fast the force is applied and the expected survival time specifies the most likely breakage force (strength) at a given loading rate (force/time).[52] In line with this previously developed theory, when moving traps at low speed, charge differences between bulk and thin film have time to equilibrate due to thermodynamically driven movements of charges in solution, reducing the attractive force felt when separating the droplets.

In a subsequent set of experiments, the depletion work associated with droplet pair separation was quantified for higher droplet approach velocity, keeping the ionic strength of the surrounding solution fixed. As for the previous experiments, the droplet pairs consisted of two droplets of similar droplet diameter. Fig. 6 presents the increase in J at increased approach velocity, at a ionic strength of 0.05 mol L^{-1} (A) or 0.15 mol L^{-1} (B), respectively. At the lowest ionic concentration investigated, 0.05 mol L^{-1}, the work J is, at an approach velocity of 4 μm s^{-1}, at the same level (∼1 pN nm) as observed for slower approaches at higher ion concentration. Measurements performed using lower velocity were not included due to the low depletion force (in A), that hindered a reliable determination of the depletion work.

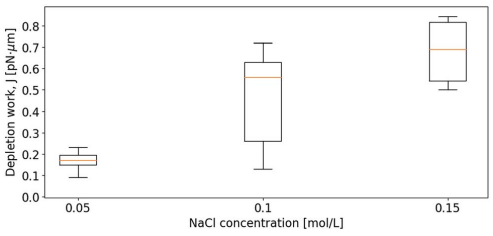


Fig. 5 Depletion work, J, as a function of increasing NaCl concentration in the bulk solution. The approach velocity of the droplets varied in the range 1–2 μm s^{-1} and the droplet diameters ranged from 4 to 7 μm.

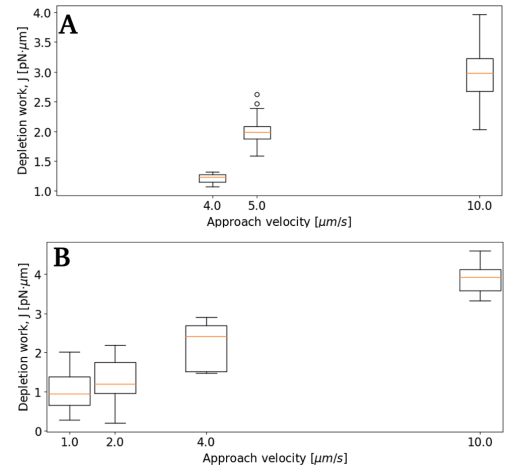


Fig. 6 Boxplots of depletion work, J, as a function of increasing approach velocities. The linear regression lines are obtained from average values calculated based on the data sets. The NaCl concentration in the bulk was 0.05 mol L^{-1} (A) and 0.15 mol L^{-1} (B), and the average droplet size was in the range 6–10 μm.

Effect of droplet velocity on depletion force

Subsequent experimental series focused on the effect of the approach velocity on the depletion work. The experiments were performed using a concentration of NaCl in the solution surrounding the oil droplets equal to 0.1 mol L^{-1}. Care was taken to minimize the size difference between the two trapped droplets. The higher velocities used induced a drag force on the trapped droplets. In order to secure that the final processed curves gave correct estimation of the depletion work, the drag force was accounted for and removed from the original force curve using the procedure described in the Materials and methods section. Fig. 7 presents the distribution of the work J, observed when separating two droplets, as a function of the droplet displacement velocity, for approach velocities up to 200 μm s^{-1}. The droplet diameters were determined for each droplet, and based on this information the droplet pairs were divided into two different size ranges; (A) 3–4 μm and (B) 8–9 μm (Fig. 7). The data presented show that, as for the lower droplet displacement velocities (Fig. 6), the work increases with increasing droplet velocity. However, while the results obtained for the droplets characterized by the largest diameters investigated (B) show a positive correlation between depletion work and droplet velocity throughout the interval of droplet velocities studied, the increase is not linear. At low velocities the results show a steeper rise in the depletion work with increasing approach velocity than what is observed at higher velocities. This behaviour may be fully or in part explained by the drag force and the resulting displacement of the droplet in the optical trap. The drag force complicates the determination of the depletion work and thus makes the force measurements using parameter settings where the drag force was important, unreliable.

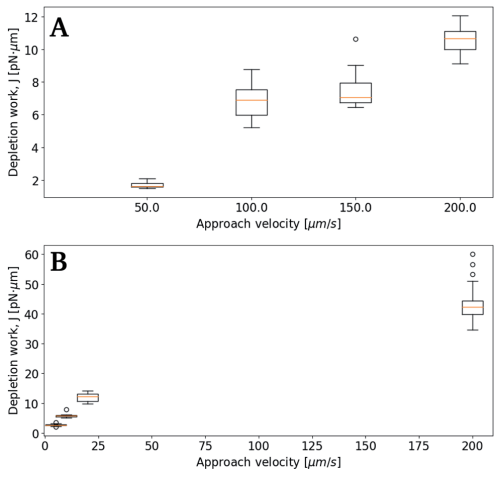


Fig. 7 Boxplots of depletion work J at 0.1 mol L^{-1} NaCl concentration and different approach velocities. Linear regression lines from average values in data sets. (A) Average droplet sizes 3–4 μm, (B) average droplet sizes 8–9 μm.

Effect of droplet size on depletion force

The results obtained in the current paper revealed that at fixed approach velocity of 10 μm s^{-1}, the depletion force and work J increases with increasing droplet size (Fig. 8). This observation is in agreement with previous OT observations on related emulsions.[42] However, droplets with d_{eq} of 12.7 and 14.7 showed similar work of separation as was observed for the droplets with d_{eq} of 10.8 μm. One reason for the unexpected low J recorded may be a fault in achieving proper contact between droplets. In planning the contact point the diameter of the droplets is visually estimated, avoiding coalescence by not over-extending trap movement. Oil droplets of diameter exceeding ~20 μm were not included in the study since, for oil droplets of this size, droplet buoyancy force overcame the trapping force, causing droplets to escape the traps.

Characterisation of factors affecting the coalescence times

The coalescence time was determined both as a function of droplet size and droplet velocity. Droplets were brought to fully overlap along the z-axis in order to make sure they were in contact, as illustrated in Fig. 3. Also, as the goal was to determine the coalescence time, any uncertainty in droplet size that would influence calibration and force measurements were irrelevant. While this scheme gave accurate times of contact and of coalescence, it did however make repeated study of the same droplet pair impossible. With the emulsion being as unstable as chosen, droplets would either coalesce, remain connected unable to separate, or be ejected and lost from the traps.

Attempts to determine coalescence times using the retract–overlap scheme described above were performed at a fixed NaCl concentration of 0.1 mol L^{-1}. When using this ion concentration emulsion droplets could be obtained by vigorous shaking of the sample tube, and a fraction of the droplets constituting the dispersed phase were within size ranges usable on the optical tweezers. In order to obtain correct coalescence times, all force *vs.* time plots obtained from the tweezers were manually examined to determine first contact point, and then the t_{coales} during a retracted pause segment, see Fig. 3. Depending on the differing coalescence dynamics of every sample run, contact, reorganization within traps, trap escape and/or depletion forces exceeding that of the traps, force–time curves would take varying appearances, with recognizable steps.

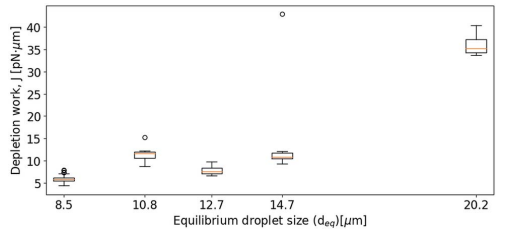


Fig. 8 Boxplots showing depletion work J against equilibrium droplet size, d_{eq}, approaching at 10 μm s^{-1}, in 0.1 mol L^{-1} NaCl concentration.

Fig. 9 shows data collected using the retract–overlap scheme, with coalescence time plotted against average droplet size (A), size of the moving droplet (B) and velocity of the moving droplet (C). The results indicate an increase in coalescence time with increasing droplet size (Fig. 9A and B). This positive correlation becomes even more striking when plotting correlation time as a function of the size of the approaching oil droplet (Fig. 9B). The coalescence time appears to show a negative correlation with droplet velocity (Fig. 9C), but the large error intervals in the data impede firm conclusions.

Despite the numerous studies of droplet coalescence, there is a lack of studies dealing with the effect of the droplet approach velocity on the coalescence time. This lack has previously been explained by the more important challenges connected to controlling the droplet velocity compared to controlling the other parameters influencing the coalescence process (*e.g.* droplet size, geometrical arrangement, temperature, phase properties, density, viscosity, surface tension, additives).[53] Some previous studies have shown that coalescence times should be inversely proportional with velocity,[54] while others have concluded that at low droplet approach velocity the coalescence is rapid and then decreases with increasing velocity.[55] In the current study the highest coalescence times are observed at the lowest droplet velocities investigated (Fig. 9C). However, at the highest droplet velocities investigated, the coalescence time appear to increase slightly. This is in agreement with what is theorized by Ozan and Jakobsen,[56] where coalescence times increases with increasing droplet velocity above a critical approach velocity.

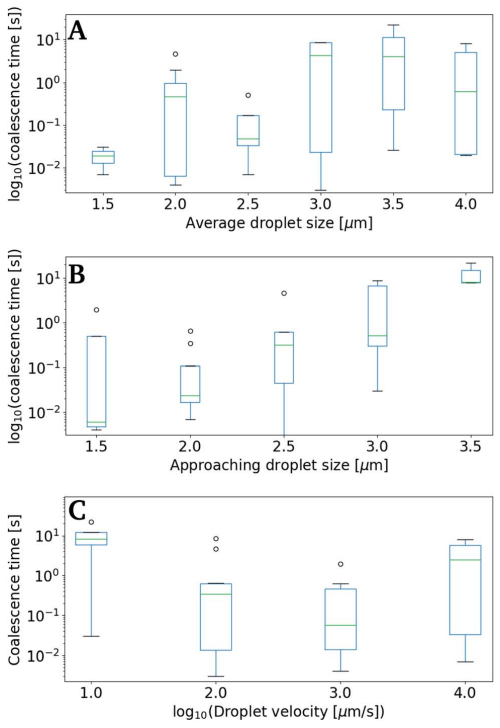


Fig. 9 Coalescence times as a function of either droplet diameter or droplet velocity. The measurements are performed for O/W emulsions containing 0.1 mol L^{-1} NaCl in the polar phase. (A) Droplet coalescence time as a function of equilibrium droplet size d_{eq}, (B) droplet coalescence time as a function of diameter of the droplet present in the approaching trap, and (C) droplet coalescence time, as a function of the approach velocity of the trap (log_{10}).

The diameters of the droplets investigated proved to have a clear effect on the coalescence time, with the coalescence times increasing proportional to both the average size of droplet pairs (Fig. 9A), and the size of the approaching (moving) droplet (Fig. 9B). Although this too follows well with other studies,[57] the domination on our results were unexpected.

As explained for the depletion measurements presented in previous sections, the drag force causes displacement of the droplets in the traps. This causes the droplet contact to be delayed, but this delay can be corrected for in the force–time curves. This delay in contact before establishing a thin film is at most in the order of milliseconds. The coalescence time – which is in order of seconds for some data points – is thus not significantly affected by this phenomenon.

In the current study, measurements were performed on droplets that were significantly smaller than the droplet sizes that are usually investigated using microfluidics,[30] AFM[58] or micropipette[59] based studies. As also pointed out by Chen and coworkers,[43] the measurement and theoretical model of emulsion droplets with diameter below 10 μm are rarely mentioned, even though this is the size scale of real emulsion droplets widely used in various applications. The importance of these small droplets for the overall properties of an emulsion is unclear, and the current study thus contribute to filling this knowledge gap. Due to the small diameters of the droplets investigated, the relative size-difference between two trapped droplets become greater than those of a micro-pipette study or a large micro-fluid cell.[30] The results obtained in the current study are therefore more subject to potential effects of varying droplet sizes.

Compared to flotation or flow-based studies, the experimental approach used in the current study allows for improved control of droplet velocity. The position control of the optical traps offer high precision control of the approach velocity, and we are not limited to the sedimentation velocity of larger droplets, or the pressure of individual pumps in a MF cell. With a force output registered at 50 kHz, the contact point, t_0, and the coalescence point t_{coales}, can be accurately determined based on the force-recordings, and does not depend on the use of a high-speed camera. Additionally, the optical tweezers allow for very precise movement and position control of trapped droplet pairs. The OT instrumentation thus opens for high resolution studies of oil droplets of a size rarely explored by other techniques. Further, the approach used in the current paper (retracting and overlapping of droplets) well emulate the collision, sticking and eventual coalescence occurring during fluidic-based experiments.[30] However, the more forceful separation of phases in rotational gravity separators,[60] or the

cumulative pressure from large number of droplets stacked high in creaming layers, subjects individual droplets to a higher vertical force. This force might possibly deform droplets giving wider and flatter thin films, which promotes the coalescence event.[61]

However, for small droplets, precise size determination is challenging and this may in part explain the large error intervals observed in some of the data presented in the current study. An interesting future improvement to the experimental approach could thus be the inclusion of custom OT/MF cells, which would allow immediate generation and subsequent trapping of emulsion droplets. Results obtained in the current study have shown that the creation of droplets in the desired size ranges for the OT is challenging due to the high interfacial tension in a surfactant-less system. A combined OT/MF approach would thus require the careful choice of surfactant to control the IFT for droplet generation. Additionally, the combined efforts of experimental measurements detailed in this paper and simulations will be important for further progress in the understanding of these emulsions. Ongoing experimental studies are therefore in our group complemented by the use of rare event simulations using the PyRETIS[62,63] and PyVisA[64] packages to aid in understanding the dynamics of thin film breakage.

Conclusions

An optical tweezers-based approach to measure the effects of approach velocity on depletion forces and work of separation, and coalescence times in pristine oil-in-water emulsions was presented. As part of the study different methods for the preparation of oil in water emulsions were evaluated. Even though the microfluidics-based approaches give an exceptional control of droplet sizes generated, subsequent coalescence of the droplets resulted in a wide droplet size distribution in the emulsion. Alternative and less work-intensive approaches for emulsion preparation were found to provide comparable droplet size distributions at a lower time consumption and were therefore used in the subsequent experiments.

Within this paper we explored the limitations on optical manipulation on emulsion droplets. These limitations included the drag force observed at high droplet velocities, and possible flocculation during droplet contact with subsequent trap escape. The identification of the thin film contact point was challenging and depended on estimates of the droplet size and the required trap displacement to contact. Still, the depletion force between pairs of oil droplets as well as the work performed when separating two droplets was determined based on OT retract–extend measurements. The work of separating droplets was found to increase with increasing ionic concentration in the continuous phase, droplet size and approach velocity. Despite large variation in data sets, the coalescence times for droplet pairs appears to decrease with increasing approach velocity. Similarly, the coalescence times for droplet pairs increased with average droplet size. However, the size of the approaching droplet seems to have a more important influence on the coalescence times compared to its approach velocity.

Author contributions

E. R. and M. S. were instrumental in obtaining the funding for the study. O. A., E. R. and M. S. designed the experiments. O. A. and M. S. developed and optimized the protocols for the OT experiments. O. A. prepared the emulsions using the different methods described in the paper. O. A. performed the OT experiments, developed the software to process and analyze the data and performed the data analysis. O. A. and M. S. drafted the manuscript with input from all authors. All authors reviewed and approved the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Swapnil Vilas Bhujbal for preparing the microfluidic cells. This work was supported by the Research Council of Norway through their funding of the PETROMAKS 2 project number 267669.

References

1 X. Sun, Y. Zhang, G. Chen and Z. Gai, *Energies*, 2017, **10**, 345.
2 F. Zhu, *Trends Food Sci. Technol.*, 2019, **85**, 129–137.
3 C. Berton-Carabin and K. Schroën, *Curr. Opin. Food Sci.*, 2019, **27**, 74–81.
4 B. A. Khan, N. Akhtar, H. M. S. Khan, K. Waseem, T. Mahmood, A. Rasul, M. Iqbal and H. Khan, *Afr. J. Pharm. Pharmacol.*, 2011, **5**, 2715–2725.
5 S. Ding, C. A. Serra, T. F. Vandamme, W. Yu and N. Anton, *J. Controlled Release*, 2019, **295**, 31–49.
6 E. Dickinson, *Food Biophys.*, 2011, **6**, 1–11.
7 J. K. Beattie and A. M. Djerdjev, *Angew. Chem., Int. Ed.*, 2004, **43**, 3568–3571.
8 M. Lashkarbolooki and S. Ayatollahi, *Can. J. Chem. Eng.*, 2018, **96**, 1396–1402.
9 M. Mitsunobu, S. Kobayashi, N. Takeyasu and T. Kaneta, *Anal. Sci.*, 2017, **33**, 709–713.
10 I. Kralova, J. Sjöblom, G. Øye, S. Simon, B. A. Grimes and K. Paso, *Adv. Colloid Interface Sci.*, 2011, **169**, 106–127.
11 J. G. Speight, *The chemistry and technology of petroleum*, CRC press, 2014.
12 F. Goodarzi and S. Zendehboudi, *Can. J. Chem. Eng.*, 2019, **97**, 281–309.
13 J. C. Slattery, L. Sagis and E.-S. Oh, *Interfacial transport phenomena*, Springer Science & Business Media, 2007.
14 L. Pan, Y. Chen, D. Chen, Y. Dong, Z. Zhang and Y. Long, *IOP Conf. Ser.: Mater. Sci. Eng.*, 2018, 042005.
15 P. Kundu and I. M. Mishra, *Rev. Chem. Eng.*, 2018, **35**, 73–108.
16 S. M. Vegas Mendoza, E. Avella Moreno, C. A. Guerrero Fajardo and R. Fierro Medina, *Water*, 2019, **11**, 1452.
17 C. Negin, S. Ali and Q. Xie, *Petroleum*, 2017, **3**, 197–211.
18 M. S. Kamal, *J. Surfactants Deterg.*, 2016, **19**, 223–236.

19 J. J. Sheng, *J. Nat. Gas Sci. Eng.*, 2015, **22**, 252–259.
20 A. Y. Mehandzhiyski, E. Riccardi, T. S. van Erp, T. T. Trinh and B. A. Grimes, *J. Phys. Chem. B*, 2015, **119**, 10710–10719.
21 A. Y. Mehandzhiyski, E. Riccardi, T. S. van Erp, H. Koch, P.-O. Åstrand, T. T. Trinh and B. A. Grimes, *J. Phys. Chem. A*, 2015, **119**, 10195–10203.
22 O. Sundman, S. Simon, E. L. Nordgard and J. Sjöblom, *Energy Fuels*, 2010, **24**, 6054–6060.
23 E. Riccardi, K. Kovalchuk, A. Y. Mehandzhiyski and B. A. Grimes, *J. Dispersion Sci. Technol.*, 2014, **35**, 1018–1030.
24 K. Kovalchuk, E. Riccardi, A. Mehandzhiyski and B. Grimes, *Colloid J.*, 2014, **76**, 564–575.
25 K. Kovalchuk, E. Riccardi and B. A. Grimes, *Ind. Eng. Chem. Res.*, 2014, **53**, 11691–11703.
26 E. Riccardi and T. Tichelkamp, *Colloids Surf., A*, 2019, **573**, 246–254.
27 S. S. Jang and W. A. Goddard, *J. Phys. Chem. B*, 2006, **110**, 7992–8001.
28 L. Rekvig and D. Frenkel, *J. Chem. Phys.*, 2007, **127**, 134701.
29 L. Leal, *Phys. Fluids*, 2004, **16**, 1833–1851.
30 M. Dudek, D. Fernandes, E. H. Herø and G. Øye, *Colloids Surf., A*, 2020, **586**, 124265.
31 D. Langevin, *Curr. Opin. Colloid Interface Sci.*, 2019, **44**, 23–31.
32 R. Pal, *AIChE J.*, 1996, **42**, 3181–3190.
33 K. C. Neuman and A. Nagy, *Nat. Methods*, 2008, **5**, 491–505.
34 A. Ashkin, *Phys. Rev. Lett.*, 1970, **24**, 156.
35 A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm and S. Chu, *Opt. Lett.*, 1986, **11**, 288–290.
36 J. R. Moffitt, Y. R. Chemla, S. B. Smith and C. Bustamante, *Annu. Rev. Biochem.*, 2008, **77**, 205–228.
37 R. Huang, I. Chavez, K. M. Taute, B. Lukić, S. Jeney, M. G. Raizen and E.-L. Florin, *Nat. Phys.*, 2011, **7**, 576–580.
38 Y. Deng, J. Bechhoefer and N. R. Forde, *J. Opt. A: Pure Appl. Opt.*, 2007, **9**, S256.
39 P. Polimeno, A. Magazzu, M. A. Iati, F. Patti, R. Saija, C. D. E. Boschi, M. G. Donato, P. G. Gucciardi, P. H. Jones, G. Volpe, et al., *J. Quant. Spectrosc. Radiat. Transfer*, 2018, **218**, 131–150.
40 A. D. Ward, M. G. Berry, C. D. Mellor and C. D. Bain, *Chem. Commun.*, 2006, 4515–4517.
41 W.-A. C. Bauer, J. Kotar, P. Cicuta, R. T. Woodward, J. V. Weaver and W. T. Huck, *Soft Matter*, 2011, **7**, 4214–4220.
42 J. Nilsen-Nygård, M. Sletmoen and K. I. Draget, *RSC Adv.*, 2014, **4**, 52220–52229.
43 A. Chen, S.-W. Li, F.-N. Sang, H.-B. Zeng and J.-H. Xu, *J. Colloid Interface Sci.*, 2018, **532**, 128–135.
44 A. Chen, S.-W. Li, D. Jing and J.-H. Xu, *Chem. Eng. Sci.*, 2019, **193**, 276–281.
45 A. Chen, F. Wang, Y. Zhou and J.-h. Xu, *Langmuir*, 2020, **36**, 4664–4670.
46 M. R. Otazo, R. Ward, G. Gillies, R. S. Osborne, M. Golding and M. A. Williams, *Soft Matter*, 2019, **15**, 6383–6391.
47 T. Danner and H. Schubert, *Food Colloids: Fundamentals of Formulation*, ed. E. Dickinson and R. Miller, 2001, pp. 116–124.
48 H. Liu and Y. Zhang, *Phys. Fluids*, 2011, **23**, 082101.
49 J. E. Melzer and E. McLeod, *ACS Nano*, 2018, **12**, 2440–2447.
50 R. Beetstra, M. A. van der Hoef and J. Kuipers, *AIChE J.*, 2007, **53**, 489–501.
51 A. Buosciolo, G. Pesce and A. Sasso, *Opt. Commun.*, 2004, **230**, 357–368.
52 E. Evans, *Biophys. Chem.*, 1999, **82**, 83–97.
53 S. Orvalho, M. C. Ruzicka, G. Olivieri and A. Marzocchella, *Chem. Eng. Sci.*, 2015, **134**, 205–216.
54 J. Kamp and M. Kraume, *Chem. Eng. Sci.*, 2016, **156**, 162–177.
55 R. Kirkpatrick and M. Lockett, *Chem. Eng. Sci.*, 1974, **29**, 2363–2373.
56 S. C. Ozan and H. A. Jakobsen, *Int. J. Multiphase Flow*, 2019, **119**, 223–236.
57 F. Gebauer, J. Villwock, M. Kraume and H.-J. Bart, *Chem. Eng. Res. Des.*, 2016, **115**, 282–291.
58 C. Shi, L. Zhang, L. Xie, X. Lu, Q. Liu, J. He, C. A. Mantilla, F. G. Van Den Berg and H. Zeng, *Langmuir*, 2017, **33**, 1265–1274.
59 C. Tang, P. Zhang and C. K. Law, *Phys. Fluids*, 2012, **24**, 022101.
60 T. Frising, C. Noïk and C. Dalmazzone, *J. Dispersion Sci. Technol.*, 2006, **27**, 1035–1057.
61 T. Krebs, D. Ershov, C. Schroen and R. Boom, *Soft Matter*, 2013, **9**, 4026–4035.
62 A. Lervik, E. Riccardi and T. S. van Erp, *J. Comput. Chem.*, 2017, **38**, 2439–2451.
63 E. Riccardi, A. Lervik, S. Roet, O. Aarøen and T. S. van Erp, *J. Comput. Chem.*, 2020, **41**, 370–377.
64 O. Aarøen, H. Kiær and E. Riccardi, *J. Comput. Chem.*, 2020, 435–446.

Paper IV

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Thin film breakage in oil–in–water emulsions, a multidisciplinary study

Ola Aarøen,*[a] Enrico Riccardi,**[b] Titus S. van Erp[c] and Marit Sletmoen[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

Coalescence is a critical phenomenon in separation and transport processes. An improved understanding of coalescence can enhance current models predicting emulsion stability and separation. We here report a combined experimental and simulation study to investigate the thin film prior to its rupture in a coalescence event. Optical tweezers measured the influence of ions and of surface-active agents on coalescence time and the forces acting between colliding oil droplets. Molecular simulation described the composition and constituent distribution of the thin films in systems comparable with the ones investigated via optical tweezers. We identify a potential relationship between the disruption of the electrical double layer and the formation of nanocrystals with the thin film breakage times and depletion forces.

Introduction

Coalescence is controlled by the drainage and eventual rupture of the thin liquid film that develops between approached droplets[1]. Thin film drainage times and film rupture mechanism are strongly dependent on the forces between the molecules present at the interfaces. surface-active molecules can, therefore, significantly influence the processes. When amphiphilic molecules are present, they tend to form layers that cover the surface of the droplets [1–3]. Moreover, the solvent molecules interacting with the surfactants determine the behavior of the thin film formed between the two inter-phase regions. Solvent rearrangement and fluctuations in local surfactant density are effects that have been long neglected in the current emulsion models [4] with consequential limitation in the predictability of emulsion behaviors. In particular, emulsions comprised of crude oil and brine (petroemulsions) are significantly affected by the dynamics of coalescence between droplets, regarding their performance of transport and separation processes [1–3].

During the past decade, extensive experimental studies have been undertaken to characterize the rheological, thermodynamic, and structural properties of interfacial layers composed of resin and asphaltene surfactants. The focus has mainly been posed on model oils employing model resin and asphaltene surfactants to facilitate well-controlled studies [5] as well as on real crude oils containing a myriad of indigenous resin and asphaltene surfactants in practical studies [2]. Molecular Dynamics simulations (MD) have provided further insight into the system. Riccardi and co-workers [5; 6] recently found that the interface region behaves as a third phase in between the two adjacent bulk phases, and its properties are different from what could be predicted based on the knowledge of its composition and a simple linear component averaging. Dedicated investigations are therefore required to determine its behavior. As an example, it has been confirmed both theoretically and experimentally [7] that emulsions containing asphaltene type surfactants are effectively stabilized against coalescence when the interfacial concentration of the asphaltene approaches and exceeds 80% of its maximum attainable interfacial concentration. Other studies have reported that the inter-phase region in between two droplets formed by asphaltene and resin type surfactants give rise to a positive disjoining pressure [1; 7–9].

Despite the experimental evidence [2] described above, current theoretical models for film drainage utilize relatively simple descriptions of surfactant behavior: London dispersion forces in oil continuous emulsions and through Derjaguin, Landau, Verwey, and Overbeek (DLVO) theory for water continuous emulsions [1; 8]. In addition to the electrostatic and van der Waals forces in DLVO theory, the force of depletion arising from osmotic pressure due to concentration difference between bulk and thin film [10] will also affect the droplet collision behavior. Furthermore, limitations on available experimental techniques to either control or measure the distance separating the interphase regions of two approaching droplets, as well as the forces acting between the droplets, have hampered a reliable validation of film drainage

[a]Department of Biotechnology and Food Science, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway
[b]Department of Informatics, UiO, Gaustadalléen 23B, 0373 Oslo, Norway
[c]Department of Chemistry, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway
* ola.aaroen@ntnu.no
** enricori@ifi.uio.no

models for coalescence [1; 8]. Due to the complexity of the systems, it has not been possible to determine the relation(s) between the macroscopic behavior and the molecular mechanisms. Recent investigations have shown at a molecular level the importance of describing the structural and conformational rearrangements in the thin film generated between two droplets [9]. The interphase region is very sensitive to system modification in each neighboring phase and the induced structural modification must be described and accounted for. The description of the structural rearrangements occurring in the inter-phase region as a response to changes in various parameters (surfactant concentration and type, ionic strength, and type) requires molecular resolution [11]. The current continuous resolution models often fail to predict overall system behavior because they adopt a too superficial description of the disjoining pressure and film rupture mechanics. For petroemulsions in particular, the considered state-of-the-art models for film drainage during coalescence [1; 8], are not in accordance with the stability mechanisms suggested by experimental studies [2].

Optical tweezers applied to emulsions have the potential to significantly reduce the previously mentioned experimental limitations. Ashkin was the first to demonstrate optical manipulation of micron-sized objects in laboratory experiments [12]. These famous experiments revealed that the radiation pressure of a beam of light could impart momentum on particles and push them along the path of light propagation. Later he showed how a small particle could be trapped in a small spacial volume by introducing a counteracting gradient force to the radiation pressure [13]. The gradient force is due to the Lorenz force and draws the particle toward the area of highest light intensity, i.e. toward the beam axis in case of a Gaussian beam profile, and toward the focus of the laser if the beam is focused. These discoveries lead to the "single-beam gradient force optical trap" [13], also referred to as an optical trap, or optical tweezers (OT), because of its ability to pick up and hold microscale particles. An optical trap is created by directing a laser beam with a Gaussian intensity profile through a lens of high NA, giving a tightly focused beam. A dielectric particle near the focus will experience a force due to the transfer of momentum from the scattering of the high number of incident photons. The resulting optical force has traditionally been decomposed into two components: a scattering force acting in the direction of light propagation and a gradient force acting in the direction of the higher light intensities along the intensity gradient. The gradient force is proportional to both the polarizability of the dielectric particle and the intensity gradient across the laser beam. For stable trapping in all three dimensions, the gradient force must be sufficient to balance the scattering force, and this is achieved in situations of a steep intensity gradient across the cross-section of the laser beam, produced by sharply focusing the trapping laser beam using the high numerical aperture objective lens.

The forces acting on a trapped particle can be described as a harmonic oscillator [14], where the particle is relatively firmly held, with the freedom to diffuse by Brownian motion [15]. The spring constant of the trap is obtained from a study of the restrictions induced by the optical trap on the Brownian motion [16]. By observing the outgoing beam profile using a QPD (quadrant photodiode) [17], any changes in the position of the trapped particle compared to the center of the optical trap, can be multiplied with the known spring constant of the trap. This (combined with accurate software-controlled optics) gives the OT set-up the ability to position particles with high precision while simultaneously measuring any 3D displacements of said particle with sub-micrometer and sub-millisecond accuracy and resolution, and also exert and quantify forces up to and beyond 10^{-10} N on particles ranging in size from nanometers to micrometer.

Since its invention, OT has been used in a wide range of research fields – from physics [18; 19], to biology and medicine [20–22]. Several studies have also included a proof-of-concept for the applicability of OT in the field of emulsion research. These have focused on, among other things, emulsion droplets with ultralow interfacial tension, droplet deformation [23; 24] and the formation of nanofluidic networks interconnecting droplets [25], pH dependant interactions between individual emulsion droplets [26], as well as the forces acting on droplets of micro and macromolecular stabilization [27]. Chen and coworkers have contributed with several advancements, including the quantification of interaction forces as functions of varying NaCl and sodium dodecyl sulfate (SDS) concentrations [28], later with non-ionic surfactant coating [29], and a switchable surface-active colloids system [30]. Otazo and coworkers have used OT for the study of coalescence. More precisely, they performed a study of partially crystalline emulsion droplets, with the aim to determine their propensity to aggregate or coalesce after being brought into contact [31]. Mitsunobo and coworkers studied temperature-induced coalescence on trapped emulsion droplets. [32] The coalescence of aerosol particles has been explored thoroughly in a single-beam optical trap [33], and using a holographic optical tweezer set-up [34; 35]. Several previous papers focus on the use of optical tweezers to study depletion forces between colloidal particles [27–29; 36]. Initial interactions between approaching interfaces will influence the probability of a coalescence event occurring, as they can increase the contact time between emulsion droplets in bulk. Previous work has also shown the effects of approach velocity on depletion forces and coalescence times in pristine oil-in-water emulsions [37]. In this work, we used OT to quantify the effect of ionic strength and surfactant density on the depletion force as well as the droplet coalescence times.

To interpret our experimental results, we complemented our study with atomistic computer simulations. The coalescence between two droplets requires the thinning and breakage of the thin film formed in between. While an actual coalescence event (the breaking of the thin film) is relatively quick, the steps leading up to the establishment of the thin film from initial contact and up to film breakage may take several seconds, even for unstabilized droplets [37]. Simulating coalescence at an atomistic scale would require unreachable amounts of computational resources. Dedicated strategies, therefore, need to be implemented. We adopted a step-wise approach to mimic the thin film thinning and breakage events. The approach provides also a necessary set-up to perform simulations with a more elaborated but demanding

simulation set-up as rare events [38–41]. These methods permit the investigation of events that are extremely unlikely but require knowledge of the transition under investigation and its potential energy landscape [42–44].

In the present work, we examine the thin film breakage in petroemulsions by the use of both optical tweezers based determination of depletion forces and coalescence times as well as molecular dynamics simulations aiming at determining the properties and thickness of the thin film prior to rupture.

Materials and methods

Preparation of emulsions and liquid cells

Emulsions were made by mixing dodecane oil, $C_{12}H_{26}$, (Sigma-Aldrich, St. Louis, USA) and de-ionized water (Milli-Q, >18 MΩ cm^{-1}, Millipore). Additionally, NaCl (Sigma-Aldrich, St. Louis, USA) and the surfactant SDS (Sigma-Aldrich, St. Louis, USA) were added to the emulsions prior to mixing, bringing the final concentrations of 5, 20, and 100 mM NaCl, and 0.3, 3, 30 and 300 μM SDS. O/W emulsions of various stability were made by adding all components of the sample to an Eppendorf tube and performing a simple homogenization by vigorously shaking the sample tube on a vortex mixer (Fisher Scientific, Pittsburg, US) at 2000 rpm for at least 3 minutes. O/W emulsions were prepared in specific oil–to–water ratios in order to obtain suitable droplet density for trapping in the sample chamber. The droplet density in the sample cell must be sufficiently low to reduce the risk of trapping more than one droplet to an acceptable level, and sufficiently high to provide a sufficient density of droplets to trap. Based on an evaluation of the current method of emulsification and the desired size range of the oil droplets, a 1:10 ratio of oil to total sample volume was chosen. Immediately post homogenization, small volumes of emulsion were transferred to pre-prepared OT fluid cells and mounted in the microscope set up of the optical tweezers. The fluid cells used for the OT experiments were made by attaching a circular Borax glass (35mm, thickness no. 1, VWR, Pennsylvania, US) to a rectangular cover glass (22x50mm, thickness no. 1, VWR, Pennsylvania, US) with two thin layers of double-sided tape, spaced roughly 0.5 – 1.0 cm apart. The height of the cell, therefore, equals the height of the two layers of tape and was in the range 100–200 μm, resulting in an internal volume of 20–30 μL.

Determination of depletion force and coalescence times using OT

The NanoTracker™ 2 optical tweezer instrument (JPK Instruments, Berlin, Germany) was used for the optical manipulation of the oil droplets. The instrument is mounted on a Zeiss Axio Observer Inverted optical microscope. It is equipped with a TEM00 laser, with a maximum power of 3W. Each laser beam has a Gaussian profile, and they can be controlled independently of each other. A quadrant photodiode detector (QPD) located at the back-focal plane of the condenser objective detects any 3D displacement of the trapped objects relative to the laser focal point. Prior to measurements, two droplets that were observed to have a size within the chosen size range were trapped, one in each laser trap. The trap stiffness was determined for each trap based on the power spectrum, which was obtained by tracking the random 3D Brownian motion of the trapped droplets.

Droplets of a size that fell within the pre-specified size range were trapped and moved away from the glass surface to avoid sticking to the surface during and prior to the experiments. As the OT instrument allows for precise position control of the optical traps, the experiments aiming at determining depletion forces were performed by moving one of the traps towards the other until the droplets came in contact. Droplet point of contact was observable as an increase in tweezers force output, in the range of pico-Newtons. Although the force applied from the tweezers now press the droplet interfaces towards each other, it is not sufficient for any droplet deformation, which would require an ultra-low interfacial tension of the droplets [24]. Droplets were maintained in contact for 0.5 seconds before moving the traps back to their initial positions, where they were kept for 0.5 seconds to remove drift due to drag force. An illustration of trapped droplets approaching (retracting), and extending can be seen in Fig. 1 (*top left* and *top right*, respectively). Fig. 1 *Bottom left* shows the recorded force output of both segments, and *Bottom right* the difference, the depletion curve. These steps were repeated until 20 force-displacement curves were recorded. Recorded force-displacement curves were plotted and processed using a Savitzky-Golay filter (bins:200, order:3), and the smoothed retract curve was subtracted from the extend curve to obtain the difference in forces acting on droplets during extend segments as compared to retract segments. This procedure is expected to result in flat curves containing a positive peak due to the attractive force acting between the oil droplets when separating them after first having brought them in contact. The height of this peak reflects the depletion force necessary to separate droplets.

For the determination of coalescence times, the optical traps were subjected to a different movement scheme. Instead of stopping the approach at droplet contact, traps were approached to the full overlap of focal points, ensuring full contact of droplets and emulating collisions occurring during droplet sedimentation. Droplets were left in contact and the time duration from reaching the contact point until coalescence occurred, was determined. The droplets were left in contact for up to a maximum of 6 minutes, a limitation of the system resources of the computer controlling the instrument. If no coalescence was observed during this time the experiment was considered a failure. For each sample composition, the coalescence time was determined for three independent droplet pairs and the results were averaged.

During both schemes, the approach velocity was fixed at 10 μm/s, as it is an important factor influencing both the depletion force and the coalescence time. The choice of concentration range for both NaCl and SDS was based on our previous observations on coalescence in optical traps [37]. The concentration range chosen covered the range within which coalescence is expected to occur. It also extended into the range where coalescence is not expected, due to too high concentrations of surfactant as well as an insufficient ionic strength of the bulk compared to the expected requirements for thin film formation and subsequent breakage.

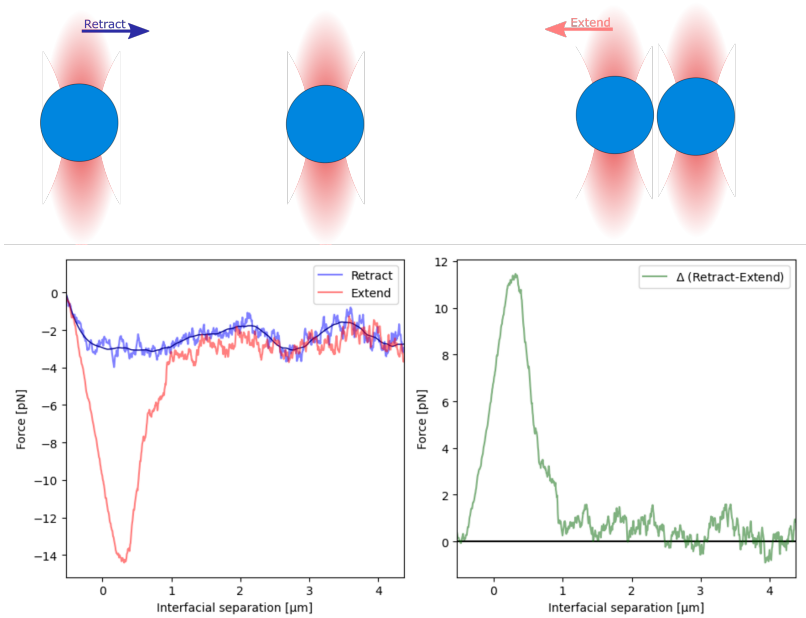


Fig. 1 *Top:* Illustration of trapped droplets retracted and extended and *Bottom left:* force versus distance plots for the two movements, colored accordingly, and *Bottom right:* the difference in force between retract and extend movements, giving the depletion force (the maximum point of the curve).

MD simulation of the rupture process of thin films

The interphase layer between two approaching dodecane droplets was studied via full atom molecular dynamics simulations. This layer is composed of water, ions, and surfactants. The ions considered in the current study are Na^+ and Cl^-, while the surfactant is SDS. The ions and the dodecane molecules were simulated according to the OPLS-AA force field [45], water molecules with the TIP4p/2005 model [46], while SDS was simulated with an improved version of OPLS-AA for the surfactant [47]. A set of 36 simulations at different SDS and ion concentrations were performed. For a water/oil surface area of 16x16 nm^2, 8, 32, 72, 128, 288, and 512 SDS molecules were included. Each system contained 8, 32, 72, 128, 288, and 512 ion pairs (Na^+ and Cl^-) in a water phase composed of 24000 water molecules. The non-polar phase was composed of 2400 dodecane molecules. The equilibrated slab of water and ion molecules was first prepared. The desired number of surfactant molecules was added both on the top and bottom of the slab, in a 2D ordered lattice. A slab of pre-equilibrated dodecane molecules was then placed on top of the 2D slab with sufficient distance to avoid overlaps between molecules. Initial simulations at 100 bar of pressure were performed to connect the various phases. The system equilibration procedure was completed by relaxing the overall system at normal conditions for 50 ns.

The MD simulations of the overall system were performed with GROMACS 5.0.4 simulation package in a $NA_{xy}T$ ensemble [48; 49], where A_{xy} is the area of the simulation box parallel of the water/dodecane interface. The temperature was set to 300 K and was controlled with a velocity rescaling method [50] employing a coupling time of 0.1 ps. The pressure, P, was controlled by the Berendsen barostat and its normal component was maintained constant at 1 bar, with a time constant of 1.0 ps and compressibility coefficient of $4.5 \cdot 10^{-5}$ bar^{-1}. Electrostatic interactions were considered via Particle-Mesh Ewald summation, cut-off distances for nonbonded interactions were taken to be equal to 1.2 nm and the leapfrog algorithm solved the Newton equation of motion with 1 fs time steps interval. Periodic boundary conditions were applied along all spatial dimensions. The x and y directions of the simulation box for all the systems have a length of 16 nm. The height (z direction) depends upon the stage of the simulations.

For each system, the thin film was progressively reduced in a step-wise manner reflecting an artificial drainage procedure. At each step, a constant number of water molecules (400) were removed and the system re-equilibrated until the breakage of the thin film occurred. Each step followed a relaxation time of 10 ns and induced a thin film reduction of circa 0.4 Å. The procedure was iterated until the thin film was ruptured, and oil phases came into contact. The thin film thickness at the step preceding the film breakage event was used to estimate the critical thickness. When using this procedure the number of ions in the system remains constant, and the removal of water molecules, leading to reduced thin film thickness, will therefore lead to an increased local ion concentration. This approach is a judicious choice considering simulations limitations. In an experimental set-up, the ion concentration in the thin film is not known. It depends upon coalescence progression and on system properties beyond the reach of the molecular dynamics simulations (e.g. ion concentration in the bulk, total droplet surface area, surface surfactant concentration). Our simulation design aimed to provide a minimal disturbance to the system that undergoes a stochastic thin film breakage. The simulation captures the effect of the ion concentration on the thin film breakage by running multiple simulations at different ion concentrations.

Results and Discussion

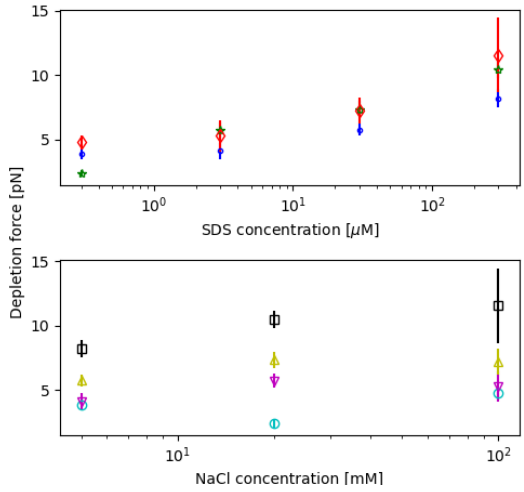


Fig. 2 Depletion force as a function of SDS concentration and NaCl concentration. *Top:* Depletion force as a function of SDS concentration. The NaCl concentration was kept constant at 5 mM (blue datapoints), 20 mM (red datapoints) and 100 mM (green datapoints). *Bottom:* Depletion force presented as a function of NaCl concentration. The SDS concentration was kept constant at 0.3μM (cyan), 3.0μM (purple), 30.0μM (yellow) or 300.0μM (black). Each data point represent the average of 20 measurements. The error bars in the plots represent the standard deviation for each data point, and all droplets investigated had a diameter in the range 8–9 μm.

When using optical tweezers to study emulsions containing oil droplets of a wide size range, care must be taken when choosing droplet pairs for trapping so that their sizes do not change significantly between samples. The emulsion preparation procedure chosen has previously been shown to provide droplet populations with a suitable size range, and is a low-cost, time-efficient alternative to other techniques which did not prove to provide more narrow size-distribution [37]. While the methods used gave a wide spread of sizes in the droplet population, the droplet pairs used in the experiments were carefully selected based on microscopic imaging, ensuring a narrow spread of diameters for the droplets included in the OT-based investigations of coalescence.

Effect of ionic strength and surfactant density on depletion force

Immediately after emulsion homogenization, the liquid cell used in the OT experiments was filled with freshly prepared emulsion. An oil droplet was trapped in each of the two optical traps and the traps were calibrated using the procedures explained in previous sections. The two oil droplets were brought in contact by moving one optical trap until contact was obtained. This point was determined based on the obtained force-displacement curves, by inspecting each curve for signs of a collision force on the extension of traps. As shown in prior experiments [37], this point of contact might be challenging to identify, and overestimation of the displacement distance required might cause droplets to coalesce. However, since the droplets used in the current study are stabilized by surfactants, the optimal contact points were, compared to our previous study [37], easier to achieve. The visual determination of size entails an inaccuracy due to the diffraction-limited resolution of light microscopy. The optical limit entails a 250 nm over or underestimation of the droplet diameter. However, these errors are not expected to significantly influence the calibration result for droplet pairs with diameters of $\sim$9 μm.

The depletion force between droplet pairs was recorded as explained in the experimental section. Samples containing varying concentrations of ions and surfactant and droplet pairs in the size range 8–9 μm were inspected. The results revealed that the depletion force increases with increasing surfactant concentration (Fig. 2, top figure). This behavior was observed for all concentrations investigated.

The experimental series obtained by keeping the SDS concentrations constant and increasing the ionic strength confirmed this behavior in the sense that the depletion force increases with increasing SDS concentration for all ionic strengths (Fig. 2(*Bottom*). The sample containing 0.3μM SDS shows an unexpected low depletion force for the intermediate NaCl concentration of 20 mM. This is likely due to an underestimation of the depletion force due to insufficient contact of the droplets pairs during OT experiments. Furthermore, it is interesting to notice that the samples containing $\leq 30\mu$M SDS reach a maximum depletion force at the intermediate NaCl concentration investigated (20 mM). Whether this behavior is also followed by the 300 μM SDS series can not be determined based on the experimental data, due to the large error margin in the last data point in this experimental series. There may be several explanations for this effect. An increased density of the ionic surfactant on the interface might lead to a reduced charge difference between the interface and the bulk and thus reduced depletion force. Alternatively, the observations may be due to the increased pliability of the interface and thus the thin film, induced by the increasing surfactant density. This might lead to increased mobility of ions into the thin film, and thus effectively reduce the experienced depletion force. A third hypothesis to explain the findings involves the effect of the ionic surfactants on the mobility of ions when approaching the droplet interfaces to form a thin film. The presence of the surfactants reduces the mobility of the ions and therefore counteracts the establishment of a difference in concentration between bulk and thin film, and thus the osmotic pressure and depletion force. An elucidation of to what extent these explanations are valid, either alone or in combination, requires an improved understanding of the underlying mechanics on an atomistic scale, which we will return to in later sections of this paper.

Effect of ionic strength and surfactant density on droplet coalescence times

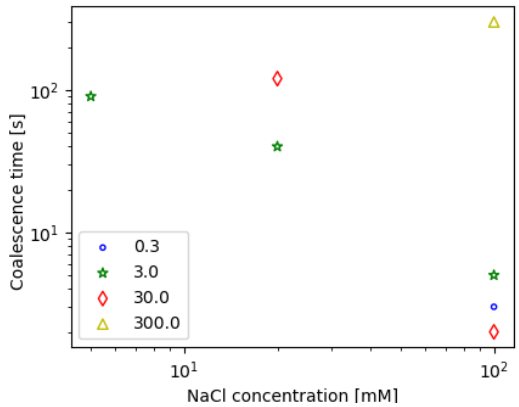


Fig. 3 Experimentally observed coalescence times under conditions of fixed SDS concentrations (colored series: blue=0.3μM, green=3.0μM, red=30.0μM, yellow=300.0μM) as a function of increasing concentration of NaCl. The data points shown in the figure are the average of 2–3 coalescence time measurements, and all oil droplets investigated had a diameter in the range 8–9 μm. A missing data point indicates that no coalescence was observed within the time period investigated (6 minutes).

For the experimental conditions (SDS concentration and ionic strength) under which coalescence was observed, a second series of samples were prepared and investigated in order to determine the coalescence times. The coalescence times presented and discussed in the following represent average values calculated based on these observations.

In Fig. 3, the average coalescence time is reported as a function of increasing NaCl concentration for different SDS concentrations. At low salt and low surfactant concentration, the coalescence time exceeds the experimental limit of 6 minutes. This agrees with previous observations [37], where we noted that a certain bulk ionic concentration was necessary in order to observe coalescence within the time window investigated on OT. At the highest surfactant concentration investigated (300 mM), the coalescence time also exceeds the experimental limit, for most of the sample series. This illustrates that with the increasing density of surfactants a threshold will be reached where the interface is efficiently stabilized by the surfactants. However, the stability of the thin film appears compromised by a high concentration of salt at an intermediate concentration of surfactant (NaCl concentration in the range between 5 mM and 20.0 mM and SDS concentration in the range between 3.0 μM and 30.0 μM). We infer that this instability might be caused by a local perturbation

driven by the surfactant molecules at the interface when present in sufficient density. The charged head-groups might promote a local over-concentration of ions that can facilitate the interaction between the two interfaces delimiting the thin film, promoting its breakage.

Properties and composition of the thin film prior to rupture

Thirty-six systems were studied to investigate the dependence of ion and surfactant concentration on thin film breakage. The thickness of the aqueous thin film was for each system gradually reduced from an initial thickness that was sufficient to secure that the two oil-water interfaces did not interact with each other. The reduction in thickness was obtained by removing a fixed number of water molecules at each step. More precisely, 400 water molecules, resulting in a reduction in the film thickness equal to 0.4Å, were removed every 10 ns. If the water thin film were showing a uniform 2D density for the 10 ns step, a new step of water removal was performed. It should be stressed that the criterion for the stability assessment in simulations based on the 10 ns interval is considerably smaller than the one used in experiments based on the 6 minutes interval. Furthermore, simulations were aimed to estimate the critical thin film thickness, while optical tweezers measure the time elapsed from droplets contact to their coalescence. Thin film breakage is too immediate to be recorded separately (sub-microsecond scale).

Table 1 reports the composition and thickness of the thin film between the two oil droplets at the last simulation step prior to breakage and it is reported according to two different definitions: (i) *Top* thin film composed of water only, (ii) *Bottom* thin film composed by water, ions and surfactants (here defined as *critical thin film thickness*). To compute the thickness of the thin film reported in Table 1, the dimensions of the simulation box composed by different combinations of the system constituents has been computed, at the same conditions of pressure, temperature, and lateral box dimensions (16x16 nm) as in the main simulation set up. In the table, the overall simulation box thickness is obtained by accounting only for the contribution of the relevant constituents.

The surface surfactant densities were calculated by dividing the total number of molecules by the total interface area, $2\times$[16 nm$\times$16 nm], and the reported ion concentrations were calculated prior to the first steps of the shrinking procedure, i.e. a 8.99x8.99x8.99 nm^3 water box.

From the Table 1, it can be asserted that the thin film rupture has a significant stochastic component, as noticeable deviations from linear trends have been reported. At the net of the statistical fluctuation, water presence in the thin film prior to breakage seems to not significantly be dependent upon surfactants nor ion concentration. The amount of water required to stabilize the thin film slightly increases for an increasing amount of surfactants.

For the critical thin film thickness, reported in Table 1(Bottom), more significant trends appear. The critical thin film thickness increases with increasing concentrations of surfactants and, to a lesser extent, of NaCl. To facilitate their visualization, the data reported in Table 1(*Bottom*) is also presented as a 3D surface (Fig. 4).

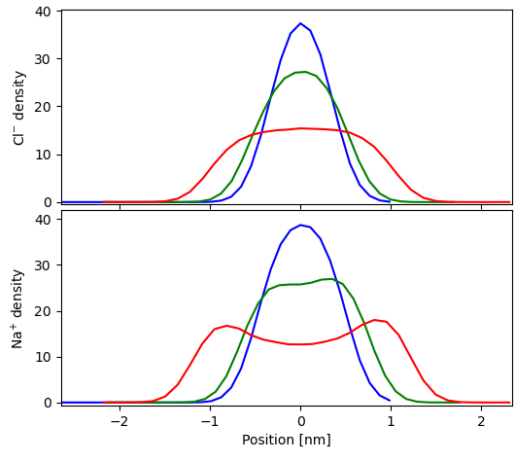


Fig. 5 Local concentration distribution of ions along the direction normal to the interface at different thin film thickness. *Top:* Cl^- ions and *Bottom:* Na^+ ions, ionic concentration 0.293 mol/L, SDS coverage 14.976 mol/mm^2. *Red line:* thin film thickness 3,939 nm, *Green line:*, thin film thickness 3,159 nm, *Blue line:* critical thickness of water film, 1.999 nm.

When removing water molecules from the thin film, the local ion concentration increases. The ion mobility is reduced in the thin film due to the presence of charged surfactants and of the finite space. It should be reminded here that the local ion concentration depends also on bulk properties, but our relatively simplified simulation could not include these effects. Yet, it is worth reporting that the formation of small clusters of ions at the interface was observed for all systems with ionic concentrations larger or equal to 0.293 mol/L. This appears to be a heterogeneous nucleation promoted by the high local ion concentration promoted by the surfactants and by the finite space. In Fig. 5, the ion concentration profiles along the normal direction to the interface are plotted for different thin film thickness, for a simulation system of 0.293 mol/L NaCl and 14.976 mol/mm^2 SDS. It can be observed how the ions tend to accumulate in the water region, with an increasing concentration. In particular, Na^+ ions reach, from a bi-modal distribution to a Gaussian distribution for small thin films. Around the thin film critical value, the concentration of Na^+ and Ca^+ ions become nearly identical, neutralizing the electrical double layer. This effect could, in essence, contribute to lower the thin film stability and, hereby, promote coalescence. Furthermore, the increased ion concentration provides the necessary conditions to promote crystal seed formation. Ordered nano-cluster of ions were observed in different simulations, at sufficiently high local salt concentrations. These clusters can locally disrupt the interface properties. Surfactant molecules cannot freely rearrange, and the interfaces surrounding the thin film are, in effect, in contact via the salt grain, possibly inducing alterations in the overall thin film thickness. These observations

Table 1 Thin film thickness prior to breakage calculated for different NaCl concentration and surfactant surface concentration. *Top:* thickness of water layer only, and *Bottom:* thickness of thin film composed by water, ions, and surfactant (here defined as critical thin film thickness).

SDS mol/mm^2 \ NaCl mol/L	0.018	0.073	0.165	0.293	0.658	1.170
0.936	1.620	1.578	1.677	1.674	1.762	1.655
3.744	1.648	1.639	1.599	1.717	1.692	1.695
8.424	1.658	1.767	1.668	1.687	1.697	1.796
14.976	1.603	1.705	1.654	1.808	1.890	1.681
33.696	1.740	1.615	1.715	1.756	1.699	1.891
59.984	1.756	1.812	1.818	1.786	1.828	1.950

SDS mol/mm^2 \ NaCl mol/L	0.018	0.073	0.165	0.293	0.658	1.170
0.936	1.622	1.579	1.687	1.690	1.807	1.748
3.744	1.692	1.684	1.643	1.777	1.781	1.833
8.424	1.751	1.861	1.770	1.796	1.835	1.982
14.976	1.777	1.880	1.837	1.999	2.109	1.948
33.696	2.142	2.019	2.127	2.175	2.147	2.387
59.984	2.501	2.557	2.572	2.547	2.617	2.787

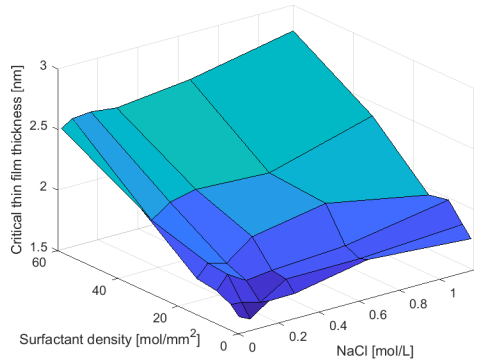


Fig. 4 Surface plot of data shown in **Table 1**(Bottom). The calculated critical thin film thickness is presented as a function of the concentration of NaCl in the solution as well as the interface density of surfactant on x- and y-axes respectively.

are consistent with the reduction in the depletion force observed in Fig. 2 for the higher NaCl concentrations. Furthermore, the restricted mobility of the ions supports the hypothesis proposed above, stating that the presence of the surfactants reduces the mobility of the ions and therefore counteracts the establishment of a charge difference between bulk and thin film, explaining the observed reduced depletion force.

To further elucidate the coalescence process, 2D density maps of water, surfactants, and ions have been generated for three different stages of step-wise water reduction and reported in Fig. 6 for the following case: density of SDS equal to 14.976 mol/mm^2, and concentration of NaCl equal to 0.658 mol/L. The plots were generated by averaging the *xy* position of the center of mass of each molecule/atom obtained from the trajectory produced in a shrinking step. The first row in Fig. 6 reports the 2D-densities at the critical thin film thickness, the second row reports the system conditions when the water layer is not any longer continuous (we consider the thin film to have been already ruptured, as the stability of this phase is extremely low) and the third row when the thin film is further compromised. The 2D density of water and surfactants is similar for all the studied cases in the system prior to the point where the critical thin film thickness is reached. The evolution of the thin film breakage can be, instead, different for different simulations of the same system. Depending on the initial phase of the breakage, the size and the shape of the water vacancy in the 2D density plot might vary. The stability of the second stage is also rather low, as most simulations evolve rather immediately to the third stage (third row) reported in Fig. 6. The water phase minimizes its surface area, generating a cylinder (for the periodic boundary condition assumption) or a sphere. For the case SDS=14.976 mol/mm^2 and NaCl=0.658 mol/L reported, ion nuclei can be observed already at the critical thin film stage. The nuclei seem to correlate with a region of water over-concentration in the second row of the 2D plot, probably contributing to the thin film breakage. In the third row, the ion nuclei are fully immersed in the water drop region.

It is worth reminding here that our simulation strategy is aimed to capture a structural description of the coalescence process. For instance, for simulation times longer than 10 ns, or different simulation setups, slightly different critical thin film thickness would result. A more accurate prediction of the film thickness prior to rupture could be obtained via rare event simulations where the probability of thin film breakage (and its rate), can be estimated for each thin film thickness. These approaches require a more involving set-up but they could directly account for the stochastic contribution in the process. The feature is also essential when studying the formation of nanocrystals that have been observed in the present simulations. While this phenomenon is of interest, the determination of the limiting conditions for its occurrence as well as the nucleation rate requires a dedicated investigation.

Conclusions

The effects of ionic strength and surfactant coverage on the depletion force acting between approached droplets were examined with optical tweezers and MD simulations. The stabilized emulsions investigated in the current study showed a similar response to depletion force by increasing ion concentrations as pristine emulsions [37]. This indicates that the presence of charges at the droplet interface did not impede the formation of a charge difference between thin film and bulk. The depletion force increased with increasing SDS concentration for all NaCl concentrations in the interval 5 to 100 mM. The samples containing $\leq 30\mu$M SDS

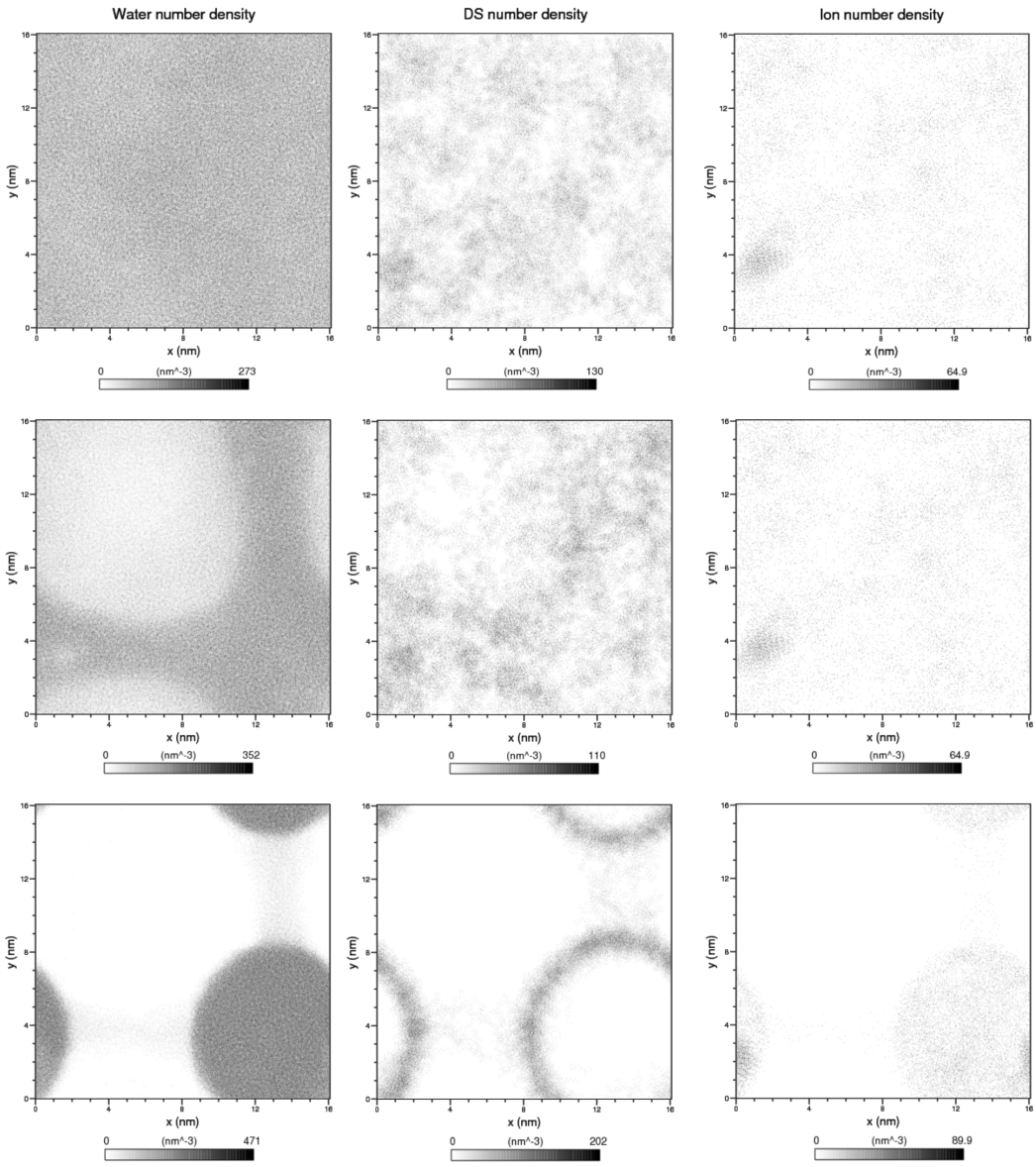


Fig. 6 Snapshots of MD simulation (Concentrations: SDS=14.976 mol/mm^2, NaCl=0.658 mol/L), showing the 2D density of ions (*left column*), water (*middle column*), and SDS (*right column*), for simulation trajectories during step-wise water molecule removal. *Top row:* # Water molecules (current film thickness) = 3200, oil phases are still separated. *Middle row:* # Water = 3100, system is beginning to show dimpling. *Bottom row:* # Water = 3000, the thin film is broken and the phases merge.

reach a maximum in depletion force at the intermediate NaCl concentration investigated (20 mM), a behavior potentially also followed by the sample containing a higher concentration of SDS. This indicates that in the absence of a sufficient ionic strength required for the reduction of the double layer extent, the ionic surfactants provided a similar function to the liquid thin film. Furthermore, this effect appeared in our coalescence experiments, where samples of comparatively low NaCl concentration to previous work [37], still coalesced. Even with the higher interfacial coverage of stabilizing surfactants, the thin film was established and broke, for droplets approached to close proximity. The work here has thus improved upon the relatively brief exploration on approaching emulsion droplets with OT [27–30], and also with aid from theoretical simulations to explain the observed behavior in the liquid thin film.

The simulation data presented in the current paper indicate that the presence of the ionic surfactants reduces the mobility of the ions and therefore counteracts the establishment of a charge difference between bulk and thin film. This may be one factor explaining the observed decrease in the depletion force, although this would merit further investigation.

Both the experimental data and the simulations reveal that the coalescence process is strongly influenced by a stochastic behavior, leading to an inherent broad distribution of several variables, including the coalescence time. Still, the data reveal a reduced coalescence time with increased ionic strength of the surrounding solution, as expected based on previously published studies. The stabilizing effect of SDS on the droplet interface was sufficient to hinder coalescence when SDS was present at a concentration equal to or exceeding 300 μM, in combination with low ionic strength $\leq$ 20 mM.

Coalescence was not observed for the sample series containing the lowest SDS concentration investigated, but it was observed for SDS concentrations equal to 3 and 30 μM. A possible explanation was provided by the MD simulations. We estimated the thin film thickness prior to breakage (here defined as the critical film thickness), for a wide range of ions and surfactant concentrations. A dedicated approach to reproduce the mechanism of the coalescence process has been presented and adopted. The results presented reveal that the critical thin film thickness increases with increasing concentrations of surfactants and, to a lesser extent, NaCl. It appears that in the presence of surfactant, the water present in the thin film to stabilize the inter-phase slightly increases for increasing concentration of ions and surfactants. Furthermore, ions appear to increase the stability of the thin film. The electrical double layer is neutralized prior to the thin film breakage.

The increased local ion concentration and their reduced mobility in the thin film promoted the formation of small clusters of ions. The agglomerate presented, in some cases, an ordered structure, representing nanocrystals or crystal seeds. Their presence appears to influence thin film stability, promoting coalescence in samples of intermediate SDS concentrations (3 and 30 μM). While the formation of the ions agglomerate/nanocrystal is of certain interest, its simulation would require dedicated set-up and methodologies able to consider and quantify the stochastic contribution of the process.

Note that since the thin film rupture is also a stochastic process, different speeds of the artificial water removals will result in somewhat different estimates of the critical thin film thickness. More advanced simulation techniques [40; 41; 51–53], that we plan to report on in future publications, would allow the study of the thin breakage process without inducing any artificial bias, and analysis of the reactive paths [54–56] could be used to quantify whether the nanocrystal formation is in fact a trigger of the film breakage. This might provide a new mechanistic insight in the film breakage process that could possible be exploited to design operational techniques to reduce or accelerate coalescence.

While the stability of emulsions depends on a large set of factors, the thin film breakage is one of the most significant bottlenecks in a coalescence process. An enhanced control and understanding of the mechanism can lead to better theoretical models and to the selection of more effective working conditions.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Research Council of Norway through their funding of the PETROMAKS 2 project number 267669.

References

1 J. C. Slattery, L. Sagis, and E.-S. Oh, *Interfacial transport phenomena*. Springer Science & Business Media, 2007.

2 I. Kralova, J. Sjöblom, G. Øye, S. Simon, B. A. Grimes, and K. Paso, "Heavy crude oils/particle stabilized emulsions," *Advances in colloid and interface science*, vol. 169, no. 2, pp. 106–127, 2011.

3 J. G. Speight, *The chemistry and technology of petroleum*. CRC press, 2014.

4 K. Kovalchuk, E. Riccardi, and B. A. Grimes, "Multiscale modeling of mass transfer and adsorption in liquid–liquid dispersions. 1. molecular dynamics simulations and interfacial tension prediction for a mixed monolayer of mono-and tetracarboxylic acids," *Industrial & Engineering Chemistry Research*, vol. 53, no. 29, pp. 11691–11703, 2014.

5 E. Riccardi, K. Kovalchuk, A. Y. Mehandzhiyski, and B. A. Grimes, "Structure and orientation of tetracarboxylic acids at oil–water interfaces," *Journal of dispersion science and technology*, vol. 35, no. 7, pp. 1018–1030, 2014.

6 E. Riccardi and T. Tichelkamp, "Calcium ion effects on the water/oil interface in the presence of anionic surfactants," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 573, pp. 246–254, 2019.

7 B. Grimes, C. Dorao, S. Simon, E. Nordgård, and J. Sjöblom, "Analysis of dynamic surfactant mass transfer and its relation-

ship to the transient stabilization of coalescing liquid–liquid dispersions," *Journal of colloid and interface science*, vol. 348, no. 2, pp. 479–490, 2010.

8 L. Leal, "Flow induced coalescence of drops in a viscous fluid," *Physics of fluids*, vol. 16, no. 6, pp. 1833–1851, 2004.

9 S. S. Jang and W. A. Goddard, "Structures and properties of newton black films characterized using molecular dynamics simulations," *The Journal of Physical Chemistry B*, vol. 110, no. 15, pp. 7992–8001, 2006.

10 Y. Mao, M. Cates, and H. Lekkerkerker, "Depletion force in colloidal systems," *Physica A: Statistical Mechanics and its Applications*, vol. 222, no. 1-4, pp. 10–24, 1995.

11 A. I. Liapis, E. Riccardi, and J.-C. Wang, "Effects on the dynamic utilization of the adsorptive capacity of chromatographic columns induced by non-uniform ligand density distributions," *Journal of separation science*, vol. 33, no. 17-18, pp. 2749–2756, 2010.

12 A. Ashkin, "Acceleration and trapping of particles by radiation pressure," *Physical review letters*, vol. 24, no. 4, p. 156, 1970.

13 A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, "Observation of a single-beam gradient force optical trap for dielectric particles," *Optics letters*, vol. 11, no. 5, pp. 288–290, 1986.

14 J. R. Moffitt, Y. R. Chemla, S. B. Smith, and C. Bustamante, "Recent advances in optical tweezers," *Annual review of biochemistry*, vol. 77, pp. 205–228, 2008.

15 R. Huang, I. Chavez, K. M. Taute, B. Lukić, S. Jeney, M. G. Raizen, and E.-L. Florin, "Direct observation of the full transition from ballistic to diffusive brownian motion in a liquid," *Nature Physics*, vol. 7, no. 7, pp. 576–580, 2011.

16 Y. Deng, J. Bechhoefer, and N. R. Forde, "Brownian motion in a modulated optical trap," *Journal of Optics A: Pure and Applied Optics*, vol. 9, no. 8, p. S256, 2007.

17 A. Pralle, M. Prummer, E.-L. Florin, E. Stelzer, and J. Hörber, "Three-dimensional high-resolution particle tracking for optical tweezers by forward scattered light," *Microscopy research and technique*, vol. 44, no. 5, pp. 378–386, 1999.

18 L. Anderegg, L. W. Cheuk, Y. Bao, S. Burchesky, W. Ketterle, K.-K. Ni, and J. M. Doyle, "An optical tweezer array of ultracold molecules," *Science*, vol. 365, no. 6458, pp. 1156–1158, 2019.

19 R. M. Pettit, W. Ge, P. Kumar, D. R. Luntz-Martin, J. T. Schultz, L. P. Neukirch, M. Bhattacharya, and A. N. Vamivakas, "An optical tweezer phonon laser," *Nature Photonics*, vol. 13, no. 6, pp. 402–405, 2019.

20 X. Wang, S. Chen, M. Kong, Z. Wang, K. D. Costa, R. A. Li, and D. Sun, "Enhanced cell sorting and manipulation with combined optical tweezer and microfluidic chip technologies," *Lab on a Chip*, vol. 11, no. 21, pp. 3656–3662, 2011.

21 N. Thammawongsa, F. D. Zainol, S. Mitatha, J. Ali, and P. P. Yupapin, "Nanorobot controlled by optical tweezer spin for microsurgical use," *IEEE transactions on nanotechnology*, vol. 12, no. 1, pp. 29–34, 2012.

22 S. Hadjialirezaei, G. Picco, R. Beatson, J. Burchell, B. T. Stokke, and M. Sletmoen, "Interactions between the breast cancer-associated muc1 mucins and c-type lectin characterized by optical tweezers," *PLoS One*, vol. 12, no. 4, p. e0175323, 2017.

23 A. D. Ward, M. G. Berry, C. D. Mellor, and C. D. Bain, "Optical sculpture: controlled deformation of emulsion droplets with ultralow interfacial tensions using optical tweezers," *Chemical communications*, no. 43, pp. 4515–4517, 2006.

24 A. L. Hargreaves, F. Gregson, A. K. Kirby, S. Engelskirchen, and C. D. Bain, "Microemulsion droplets in optical traps," *Journal of Molecular Liquids*, vol. 210, pp. 9–19, 2015.

25 D. A. Woods, C. D. Mellor, J. M. Taylor, C. D. Bain, and A. D. Ward, "Nanofluidic networks created and controlled by light," *Soft Matter*, vol. 7, no. 6, pp. 2517–2520, 2011.

26 W.-A. C. Bauer, J. Kotar, P. Cicuta, R. T. Woodward, J. V. Weaver, and W. T. Huck, "Microfluidic production of monodisperse functional o/w droplets and study of their reversible ph dependent aggregation behavior," *Soft Matter*, vol. 7, no. 9, pp. 4214–4220, 2011.

27 J. Nilsen-Nygaard, M. Sletmoen, and K. I. Draget, "Stability and interaction forces of oil-in-water emulsions as observed by optical tweezers–a proof-of-concept study," *RSC Advances*, vol. 4, no. 94, pp. 52220–52229, 2014.

28 A. Chen, S.-W. Li, F.-N. Sang, H.-B. Zeng, and J.-H. Xu, "Interactions between micro-scale oil droplets in aqueous surfactant solution determined using optical tweezers," *Journal of colloid and interface science*, vol. 532, pp. 128–135, 2018.

29 A. Chen, S.-W. Li, D. Jing, and J.-H. Xu, "Interactions between colliding oil drops coated with non-ionic surfactant determined using optical tweezers," *Chemical Engineering Science*, vol. 193, pp. 276–281, 2019.

30 A. Chen, F. Wang, Y. Zhou, and J.-h. Xu, "In situ measurements of interactions between switchable surface-active colloid particles using optical tweezers," *Langmuir*, vol. 36, no. 17, pp. 4664–4670, 2020.

31 M. R. Otazo, R. Ward, G. Gillies, R. S. Osborne, M. Golding, and M. A. Williams, "Aggregation and coalescence of partially crystalline emulsion drops investigated using optical tweezers," *Soft matter*, vol. 15, no. 31, pp. 6383–6391, 2019.

32 M. Mitsunobu, S. Kobayashi, N. Takeyasu, and T. Kaneta, "Temperature-induced coalescence of droplets manipulated by optical trapping in an oil-in-water emulsion," *Analytical Sciences*, vol. 33, no. 6, pp. 709–713, 2017.

33 R. Power, J. Reid, S. Anand, D. McGloin, A. Almohamedi, N. Mistry, and A. Hudson, "Observation of the binary coalescence and equilibration of micrometer-sized droplets of aqueous aerosol in a single-beam gradient-force optical trap," *The Journal of Physical Chemistry A*, vol. 116, no. 35, pp. 8873–8884, 2012.

34 B. R. Bzdek, L. Collard, J. E. Sprittles, A. J. Hudson, and J. P. Reid, "Dynamic measurements and simulations of airborne picolitre-droplet coalescence in holographic optical tweezers," *The Journal of chemical physics*, vol. 145, no. 5, p. 054502, 2016.

35 A. E. Haddrell, R. E. Miles, B. R. Bzdek, J. P. Reid, R. J. Hop-

kins, and J. S. Walker, "Coalescence sampling and analysis of aerosols using aerosol optical tweezers," *Analytical chemistry*, vol. 89, no. 4, pp. 2345–2352, 2017.

36 S. Liu, Y. Hu, J. Xia, S. Fang, and M. Duan, "In situ measurement of depletion caused by sdbs micelles on the surface of silica particles using optical tweezers," *Langmuir*, vol. 35, no. 42, pp. 13536–13542, 2019.

37 O. Aarøen, E. Riccardi, and M. Sletmoen, "Exploring the effects of approach velocity on depletion force and coalescence in oil-in-water emulsions," *RSC Advances*, vol. 11, no. 15, pp. 8730–8740, 2021.

38 T. S. Van Erp and P. G. Bolhuis, "Elaborating transition interface sampling methods," *Journal of computational Physics*, vol. 205, no. 1, pp. 157–181, 2005.

39 T. S. Van Erp, "Dynamical rare event simulation techniques for equilibrium and nonequilibrium systems," *Advances in Chemical Physics*, vol. 151, p. 27, 2012.

40 A. Lervik, E. Riccardi, and T. S. van Erp, "Pyretis: A well-done, medium-sized python library for rare events," *Journal of Computational Chemistry*, vol. 38, no. 28, pp. 2439–2451, 2017.

41 E. Riccardi, A. Lervik, S. Roet, O. Aarøen, and T. S. van Erp, "Pyretis 2: An improbability drive for rare events," *Journal of Computational Chemistry*, vol. 41, no. 4, pp. 370–377, 2020.

42 M. Moqadam, A. Lervik, E. Riccardi, V. Venkatraman, B. K. Alsberg, and T. S. van Erp, "Local initiation conditions for water autoionization," *Proceedings of the National Academy of Sciences*, vol. 115, no. 20, pp. E4569–E4576, 2018.

43 E. Riccardi, E. C. Van Mastbergen, W. W. Navarre, and J. Vreede, "Predicting the mechanism and rate of h-ns binding to at-rich dna," *PLoS computational biology*, vol. 15, no. 3, p. e1006845, 2019.

44 E. Riccardi, A. Krämer, T. S. van Erp, and A. Ghysels, "Permeation rates of oxygen through a lipid bilayer using replica exchange transition interface sampling," *The Journal of Physical Chemistry B*, 2020.

45 W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids," *Journal of the American Chemical Society*, vol. 118, no. 45, pp. 11225–11236, 1996.

46 J. L. Abascal and C. Vega, "A general purpose model for the condensed phases of water: Tip4p/2005," *The Journal of chemical physics*, vol. 123, no. 23, p. 234505, 2005.

47 S. Abdel-Azeim, "Revisiting opls-aa force field for the simulation of anionic surfactants in concentrated electrolyte solutions," *Journal of Chemical Theory and Computation*, vol. 16, no. 2, pp. 1136–1145, 2020.

48 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, "Gromacs: fast, flexible, and free," *Journal of computational chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005.

49 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 1, pp. 19–25, 2015.

50 G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of chemical physics*, vol. 126, no. 1, p. 014101, 2007.

51 T. van Erp, D. Moroni, and P. Bolhuis, "A novel path sampling method for the calculation of rate constants," *J. Chem. Phys.*, vol. 118, p. 7762, 2003.

52 T. van Erp, "Reaction rate calculation by tarallel path swapping," *Phys. Rev. Lett.*, vol. 98, p. 268301, 2007.

53 E. Riccardi, O. Dahlen, and T. S. van Erp, "Fast decorrelating monte carlo moves for efficient path sampling," *The journal of physical chemistry letters*, vol. 8, no. 18, pp. 4456–4460, 2017.

54 T. S. van Erp, M. Moqadam, E. Riccardi, and A. Lervik, "Analyzing complex reaction mechanisms using path sampling," *Journal of chemical theory and computation*, vol. 12, no. 11, pp. 5398–5410, 2016.

55 O. Aarøen, H. Kiær, and E. Riccardi, "Pyvisa: Visualization and analysis of path sampling trajectories," *Journal of Computational Chemistry*, vol. 42, no. 6, pp. Pages 435–446, 2020.

56 S. Roet, C. D. Daub, and E. Riccardi, "Chemistrees: data driven identification of reaction pathways via machine learning," *arXiv preprint arXiv:2104.11728*, 2021.

ISBN 978-82-326-5715-5 (printed ver.)
ISBN 978-82-326-6444-3 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU

Norwegian University of
Science and Technology