POLITECNICO DI TORINO

Collegio di Ingegneria Chimica e dei Materiali

Corso di Laurea Magistrale
in Ingegneria Chimica e dei Processi Sostenibili

Tesi di Laurea Magistrale

Study of the Crystallization Behavior of Organic Molecules: A Combined Computational and Experimental Approach

Relatori

Prof.ssa Elena SIMONE
Prof. Antonio BUFFO
Prof. Nicodemo DI PASQUALE
Dott. Emmanuele PARISI

Candidata

Martina Gaia BAELI

Ottobre 2025

Italian Summary

Il presente studio si concentra sul comportamento morfologico e di crescita dello 9H-Xanthen-9-one in tre diversi solventi (acetone, acetonitrile e toluene), esaminato da una doppia prospettiva: sperimentale e computazionale. Lo xanthone è un composto organico aromatico eterociclico, facente parte della famiglia polifenolica degli xantoni, noto per le sue proprietà farmacologiche e per le sue potenziali applicazioni nei materiali funzionali.

L'influenza del solvente nelle cinetiche di crescita e dei cambiamenti che essi apportano alla morfologia del cristallo stesso è un aspetto di particolare interesse, soprattutto nei settori farmaceutico, chimico e ingegneristico, derivante dalla consapevolezza che diverse morfologie cristalline di una stessa molecola possono presentare proprietà differenti e, di conseguenza, adattarsi in modo più efficiente a specifiche applicazioni. Tali differenze influenzano aspetti critici come la cinetica del processo, l'efficienza delle operazioni di downstream e la resa finale. Nei processi di downstream ad esempio, alcune morfologie, come quella aghiforme, sono indesiderate, in quanto comportano problemi nella fase di filtrazione del prodotto. Al contrario, morfologie più compatte, come quella prismatica, possono risultare vantaggiose in applicazioni in cui si desidera aumentare la superficie di contatto del cristallo, ad esempio per migliorare la dissoluzione o l'interazione con altri componenti del sistema.

L'analisi computazionale e sperimentale sono state integrate al fine di verificare se fosse possibile ottenere informazioni sul comportamento morfologico dello xanthone basandosi esclusivamente su dati di simulazione molecolare e, di conseguenza, facilitarne il design del processo di cristallizzazione. In particolare, si è stimato il valore dell'energia interna del sistema cristallo–solvente all'interfaccia di tre diverse facce cristallografiche. L'energia interna rappresenta il primo termine dell'energia libera di Helmholtz; nello specifico, si è cercato di correlare tale termine a proprietà macroscopiche del sistema, quali la cinetica di crescita e il grado di anisotropia morfologica, al fine di valutare se l'impiego esclusivo delle simulazioni computazionali possa essere sufficiente a predire, almeno qualitativamente, tali caratteristiche, riducendo la necessità di test preliminari. Si sottolinea tuttavia che l'energia libera di Helmholtz comprende anche un contributo entropico, che in questo lavoro non è stato considerato nei calcoli. Tale omissione comporta una semplificazione

del modello, potenzialmente limitandone la capacità predittiva.

I solventi investigati sono stati acetone, acetonitrile e toluene. In tutti gli esperimenti, condotti in soluzione sovrasatura con rapporto $1.09\frac{C}{C_{eq}}$ tramite protocollo di seeding (massa dei seed: 5% in massa della concentrazione di soluto in equilibrio), è stato applicato un gradiente di raffreddamento costante di 0.5 $°C$/min. Per ciascun esperimento sono stati prelevati nove campioni (i primi otto ogni ora, l'ultimo a distanza di 14 ore per misurare le dimensioni delle principali direzioni di crescita e i valori di assorbanza raggiunti asintoticamente), su cui è stata condotta l'analisi di immagine tramite il software *ImageJ*, al fine di determinare le dimensioni caratteristiche dei cristalli (lunghezza e larghezza). I dati raccolti sono stati successivamente elaborati tramite analisi statistiche per ricavare i parametri caratteristici delle distribuzioni dimensionali dei cristalli nei diversi campioni. Indicatori come valore medio, deviazione standard e intervallo di confidenza hanno permesso di valutare la significatività statistica del campionamento effettuato, fornendo una base quantitativa per valutare quanto i risultati ottenuti siano rappresentativi della popolazione cristallina nel suo complesso: dal confronto tra gli intervalli di confidenza calcolati e la risoluzione del microscopio ottico emerge che la discrepanza è limitata e, pertanto, pur essendo possibile migliorare il campionamento, i risultati ottenuti risultano comunque adeguati a supportare le valutazioni qualitative condotte.

Nonostante i limiti quantitativi dell'analisi statistica, i dati ottenuti permettono comunque di trarre considerazioni di natura qualitativa, che risultano rilevanti ai fini dell'interpretazione del sistema. Gli esperimenti mostrano infatti che i differenti solventi influenzano in modo significativo la cinetica di crescita, portando alla formazione di morfologie finali dello xanthone sensibilmente diverse, sia in termini di dimensione generale dei cristalli, sia per quanto riguarda il rapporto lunghezza/larghezza. La valutazione di questo parametro consente di indagare il comportamento anisotropo del cristallo, ovvero quanto le sue proprietà (ad esempio la lunghezza) varino in funzione della direzione cristallina considerata. L'analisi ha evidenziato una maggiore anisotropia nel toluene, dove il rapporto lunghezza-/larghezza risulta più elevato, mentre negli altri due solventi, in particolare nell'acetonitrile, il cristallo presenta una crescita più uniforme tra le diverse direzioni.
In tutti e tre i casi, inoltre, è stato osservato il fenomeno di nucleazione secondaria, evidenziando la presenza di meccanismi complessi nel processo di crescita cristallina.

Per la parte computazionale, le simulazioni sono state effettuate con il software *LAMMPS*, inizialmente validando le box monocomponenti in ensemble NPzAT mediante il confronto con dati di letteratura di densità, funzione di distribuzione radiale (RDF) e diffusività dei tre solventi. Successivamente, è stato simulato il sistema completo (cristallo - solvente) in ensemble NVT, calcolando l'energia interna per ciascuna faccia ortogonale dello xanthone, (100), (010), (001) secondo gli indici cristallografici, nei diversi solventi. La differenza tra

l'energia interna della box unita e la somma delle energie interne delle due box separate ha permesso di stimare il contributo energetico all'interfaccia.

I risultati computazionali hanno mostrato che il toluene presenta il valore più basso di energia interna, seguito da acetone e infine acetonitrile. In particolare, le facce energeticamente favorite sono risultate essere la (001) per toluene e acetonitrile, e la (010) per acetone. Da tali valori ci si aspetterebbe che la crescita cristallina sia più rapida nel toluene e più lenta nell'acetonitrile, con crescita più rapida in direzione perpendicolare alle facce per cui rispettivamente si è calcolata un'energia minore.

L'analisi sperimentale ha effettivamente confermato che lo xanthone cresce più rapidamente e in modo più anisotropo in toluene, con cristalli fino a 13 volte più grandi rispetto a quelli sviluppati in acetonitrile; tuttavia, si è riscontrata una discrepanza tra le previsioni basate solo sull'energia interna e lo sviluppo morfologico reale delle singole facce: in alcuni casi, infatti, le facce che sperimentalmente crescono meno, in quanto facce più ampie e dunque presentanti maggior competività tra le interazioni soluto-solvente e soluto-soluto, risultano invece energeticamente favorite nei calcoli computazionali, condizione che suggerirebbe, al contrario, una crescita più rapida. Di conseguenza, sebbene l'energia interna risulti un buon indicatore per confrontare qualitativamente la tendenza alla crescita nei diversi solventi, essa non è sufficiente per predire con precisione la morfologia finale del cristallo o la crescita specifica delle singole facce.

Pertanto, l'analisi integrata tra dati sperimentali e computazionali ha fornito una comprensione più approfondita delle dinamiche di cristallizzazione dello xanthone; ma le discrepanze osservate evidenziano i limiti dell'approccio basato esclusivamente sull'energia interna, mentre la buona corrispondenza nelle tendenze generali di crescita convalida l'utilizzo della dinamica molecolare come strumento preliminare di screening.
Per ottenere predizioni morfologiche più accurate, in futuro sarà necessario includere anche il termine entropico dell'equazione di Helmholtz. Inoltre, il calcolo delle cinetiche di crescita specifiche per ciascun sistema e la ripetizione degli esperimenti permetterà una caratterizzazione più precisa dei meccanismi coinvolti e dei comportamenti cristallini.

Abstract

The present work investigates the morphological behavior of 9H-Xanthen-9-one, an aromatic heterocyclic compound from the polyphenolic xanthone family, known for its pharmaceutical properties and its potential applications in functional materials. In particular, the study examines how crystal morphology and growth kinetics depend on three different solvents (acetone, acetonitrile, and toluene). This aspect is of great relevance in the pharmaceutical, chemical, and engineering fields, as different crystal morphologies of the same compound can exhibit distinct properties and suitability for specific applications. For example, needle-like crystals are often undesirable in downstream processes due to filtration issues, whereas prism-like crystals may be advantageous when increased surface area is required for dissolution or interaction with other system components.
The analysis was conducted through a dual approach: experimental crystallization and molecular dynamics (MD) simulations. The aim is to assess whether the internal energy at the crystal–solvent interface alone, obtained computationally, could serve as a reliable predictor of experimental crystal growth trends and interfacial properties.

Experimentally, xanthone was crystallized in the three solvents using identical supersaturation and seeding conditions. Crystals were sampled at regular time intervals over a 24-hour period and analyzed via image processing to determine characteristic size distributions. To evaluate the statistical representativeness of each sample, the number of crystals analyzed was examined through methods such as the cumulative running average and moment analysis of the size distributions: by comparing the calculated confidence intervals with the resolution of the optical microscope, it emerges that the discrepancy is minor and thus, despite possible improvements in sampling, the results remain adequate to support the qualitative assessments performed in this work.
Experimental results showed a clear solvent-dependent trend: growth was fastest and most anisotropic in toluene, followed by acetone, with acetonitrile yielding the slowest and least anisotropic growth.

From a computational perspective, MD simulations were conducted using the LAMMPS package. The simulations focused on evaluating the interfacial internal energy, the first term of the Helmholtz free energy equation, for the three orthogonal crystal faces (100), (010), and (001). After validating the simulation setup via density, Radial Distribution

Functions, and diffusivity benchmarks obtained from literature, energy calculations were performed on both mono-component (crystal or solvent only) and crystal–solvent combined systems.
The simulation results echoed the experimental hierarchy: toluene yielded the lowest internal energy values, followed by acetone and then acetonitrile. However, discrepancies arose when comparing face-specific growth: in some cases, the faces that experimentally grew the least were those with the lowest internal energy, contrary to thermodynamic expectations.
This mismatch indicates that internal energy alone is insufficient to predict directional growth: although the computed energies can reliably forecast overall solvent effects on crystal growth kinetics, they do not capture anisotropic development.

In conclusion, the study demonstrates the viability of linking computational and experimental approaches to evaluate crystal morphology and growth kinetics, confirming the potential of molecular dynamics as a qualitative screening tool for the design of crystallization processes. Nonetheless, a more comprehensive energetic model including the entropic contribution is necessary for accurate morphological predictions. Future work should therefore focus on thermodynamic integration to quantify the entropic term and on crystal growth kinetics modeling to deepen the understanding of solvent-specific behavior and dependence parameters.

Ich glaube, heute Abend wird es regnen
So wie ich dich kenne passt dir das ganz gut

- *AnnenMayKantereit*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The influence of the solvent on crystal growth kinetics and the morphological changes it induces is a subject of particular interest, especially in the pharmaceutical, chemical, and engineering fields. This interest stems from the understanding that different crystal morphologies can exhibit distinct properties and, consequently, be better suited for specific applications.[1]
Such morphological differences affect critical aspects of the process, including growth kinetics, downstream efficiency, and overall yield. For example, in downstream processing, certain morphologies, such as needle-like forms, are generally undesirable, as they pose challenges during product filtration. Conversely, more compact or flat morphologies, such as prism-like crystals, can be advantageous in applications where an increased surface area is functionally beneficial, for instance to improve dissolution or interaction with other components in the system.
Understanding and controlling crystallization through solvent selection is therefore crucial to obtain the most suitable morphology for a specific application, especially when aiming for enhanced performance, bioavailability, or stability of the final product.

In parallel, a computational approach is employed to evaluate the facet-specific energetic contributions within the system. Molecular Dynamics simulations are conducted on both single-component systems (comprising only the solvent or the crystal) and combined systems (crystal–solvent) to estimate the primary term in the interfacial free energy equation.

The central hypothesis is that internal energy at the crystal–solvent interface can serve as a meaningful predictive descriptor for macroscopic properties such as growth kinetics and anisotropy.
An attempt was made to correlate this term with these properties in order to assess whether the exclusive use of computational simulations may be sufficient to qualitatively predict such characteristics, thereby reducing the need for preliminary experimental testing.

1.1 Thesis Objective

The present work aims to experimentally investigate the crystallization behavior of xanthone in different solvents. It has been observed that the crystal exhibits distinct morphologies depending on the crystallization medium. Therefore, this study focuses on analyzing the characteristic size, shape distribution and growth kinetics of xanthone crystals.

In parallel, a computational approach is employed to evaluate the different energy contributions in order to obtain useful information for solvent selection and process design. Molecular dynamics simulations are performed on both single-component systems (containing only the solvent or the crystal) and combined systems (solvent-crystal).

This thesis is thus structured around two complementary approaches: experimental and computational. This dual nature is consistently maintained throughout the chapters, each of which presents the topic from both perspectives. A comparative and integrative analysis is then provided in the results chapter, where the two methodologies are brought together to find connections and correlations.

The structure of the thesis is as follows:

- Chapter 2 – State of the Art: An overview of the literature on crystallization and molecular dynamics. The main parameters, techniques, and theoretical models relevant to this work are introduced, including the key equations used in the computational analysis.
- Chapter 3 – Materials and Methods: A detailed description of the experimental procedures and computational methods employed.
- Chapter 4 – Results and Analysis: Presentation and discussion of the experimental and computational results, first individually and then as a whole.
- Chapter 5 – Conclusions and Future Work: Summary of the findings and suggestions for future developments or improvements to the study.

Chapter 2

State of the Art

2.1 State of the Art in Crystallization

2.1.1 Crystallization Principles

Crystallization is a widely employed technique in purification and separation processes across a broad spectrum of materials, ranging from fine chemicals to pharmaceutical products. Over the past decades, scientific interest in crystallization has significantly increased, not only due to its inherent complexity, stemming from its multicomponent and multiphase nature, as well as the wide scale of the phenomena involved, but also in response to the expanding range of application fields, including protein crystallization and the development of nutraceutical compounds.

The process involves a phase transition in which a solute undergoes transformation from the liquid state to the solid crystalline phase, adopting a periodic and ordered structure with fixed intermolecular distances.

A solution is a homogeneous mix of two or more substances. In this context, it is considered as formed by the dissolution of a solid solute in a solvent, the latter acting as the bulk medium. The extent to which the solute dissolves is governed by solubility, which defines the equilibrium concentration of solute in the solvent as a function of temperature. For each solute–solvent pair, a specific solubility curve exists, typically exhibiting a positive temperature dependence: that is, solubility increases with temperature (Figure 2.1).

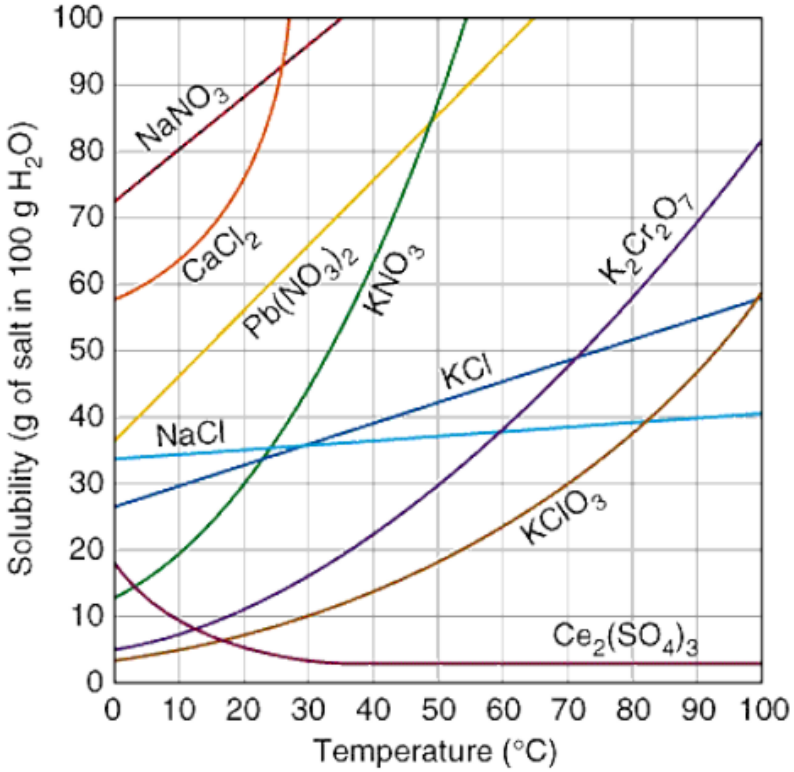


Figure 2.1: Example of solubility curves for different salts from [2]

By examining the solubility curve, it is possible to determine the state of a solution at any given temperature. If the concentration of solute lies below the curve, the system is classified as unsaturated; vice versa, if the concentration exceeds the equilibrium value, the solution is considered supersaturated, a condition under which crystallization may occur. A solution in which the solute concentration matches the solubility curve is deemed saturated. It is however possible for the solution to contain more solid than the equilibrium concentration and not to form solids or precipitates: this zone is known as metastable zone. Furthermore, the term induction time is commonly used to refer to the time required for the system to overcome the nucleation energy barrier. Figure 2.2 allows for a clearer picture of the conditions described.

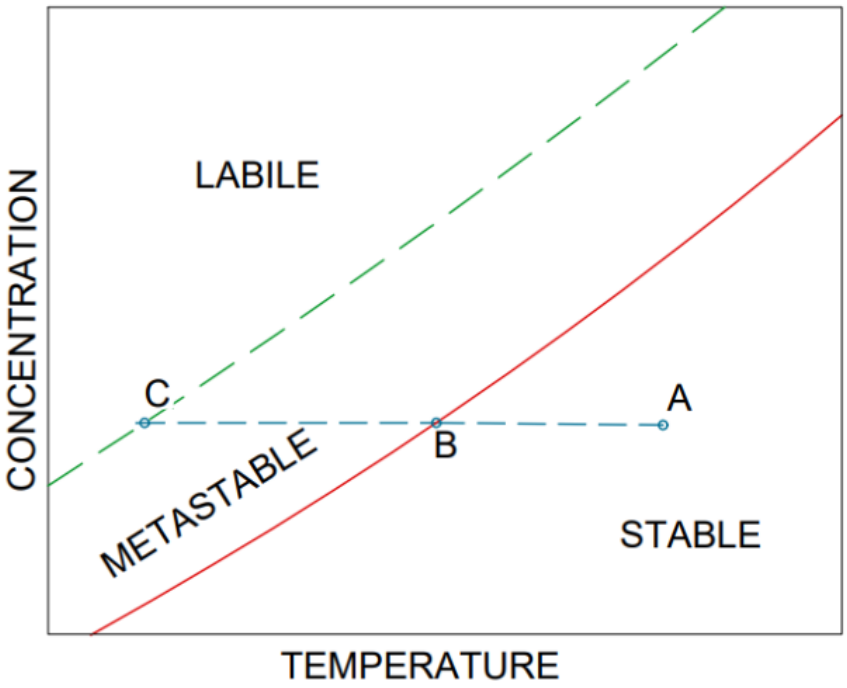


Figure 2.2: Depiction of solubility curve with stable, metastable and labile zone (from [3] and modified)

Starting from point A, in the stable undersaturated zone, a decrease in temperature allows to go past the solubility curve, reaching the metastable zone with a certain degree of supersaturation in point B: here spontaneous crystallization is improbable, but crystallization via other means is possible. By further decreasing the temperature, the over saturation limit is reached in point C, landing the solution in the labile zone, with consequent precipitation of part of the solute dissolved, in order to return to a state of equilibrium by decreasing the concentration of solid dissolved.
The degree of supersaturation can be described as difference in concentrations (Equation 2.1):

$$\Delta S = C - C^* \tag{2.1}$$

or ratio of concentrations (Equation 2.2):

$$\Delta S = \frac{C}{C^*} \tag{2.2}$$

or relative supersaturation (Equation 2.3):

$$\sigma = \frac{C - C^*}{C^*} = S - 1 \tag{2.3}$$

where C is the concentration of the given solution, and C* represents the concentration at equilibrium.
Supersaturation is the driving force for crystallization processes, as it is an essential

condition for all crystallization operations.

2.1.2 Crystallization Techniques

Numerous crystallization techniques have been developed throughout the decades, and they are distinguished based on the way supersaturation is created.[4]

- *Evaporative Crystallization*: the solvent is removed by evaporation from a solution, therefore obtaining supersaturation by diminishing the overall quantity of solvent present and consequently increasing the solute concentration. It is mainly used when the solubility does not show a strong dependence on temperature;
- *Cooling Crystallization*: supersaturation is reached by decreasing the temperature of the equilibrated solution, thus decreasing also the solubility (see Section 2.1.1). The equilibrium is then reestablished by growing crystals in the solution. It is a technique mostly used in the majority of crystallization processes, and it is particularly efficient if the solubility has a strong dependence on temperature;
- *Crystallization via antisolvent*: although the addition of an antisolvent, that is, a liquid not able to dissolve the solute, dilutes the mixture and reduces the concentration, it also reduces the overall solubility in the mixture, thus reaching supersaturation conditions and allowing crystals to form. In this instance a good mixing of the process is necessary to avoid large local supersaturation zones close to the inlet points of the antisolvent. However, mixing effects and spatially localized variations in system properties make the process control difficult;
- *Reactive Crystallization*: it occurs when two soluble materials are mixed and react, resulting in a sparingly soluble solute. If the concentrations are sufficiently high, and therefore also higher than the solubility, supersaturation is obtained. Same for the antisolvent method, this technique requires good mixing and homogeneity to avoid imbalances in the concentration of the solution.

2.1.3 Nucleation

Supersaturated conditions alone are not sufficient to allow the system to crystallize: it is necessary to overcome the nucleation energy barrier in order to have solid nuclei able to grow. This stochastic phenomenon of formation of crystalline matter in a supersaturated clear solution, whether spontaneous or not, is referred to as nucleation[3] [5]. Nucleation is distinguished in primary and secondary: in the first case, the crystals formed spontaneously in a solution that previously was homogeneous in the liquid phase; in the latter, other nuclei are formed near crystals already present in the supersaturated system. Furthermore, primary nucleation is defined as homogeneous when it happens spontaneously, and heterogeneous when it is induced by foreign particles, whether it being other crys-

tals of the same species (known as seeds), or impurities (dust in the equipment, external particles not belonging to the crystal, etc).

Homogeneous nucleation

Homogeneous nucleation is a process controlled only by the degree of supersaturation of the system, and it is appliable only in the case of certain absence of impurities or external bodies. It is therefore extremely difficult in an experimental environment to be sure that this phenomenon has occurred. Nonetheless, it is the basis for many nucleation theories, such as the classical nucleation theory (CNT), stemming from Gibbs (1948), Volmer (1939) and others. [3]
This theory visualizes the system as formed by a certain number of molecules that at one point will form clusters resulting mainly from a sequence of bimolecular additions, following the scheme:

$$\begin{aligned} A + A &\rightleftharpoons A_2 \\ A_2 + A &\rightleftharpoons A_3 \\ &\vdots \\ A_{n-1} + A &\rightleftharpoons A_n \end{aligned}$$

where A_n is the size of the critical cluster, leading to nucleation and subsequent crystal growth: if the nucleus achieves a certain critical size, it is considered stable and the addition reactions thermodynamically favored.
The nucleation rate, J, as in the number of nuclei formed per unit time per unit volume, is expressed in the form of the Arrhenius reaction velocity in Equation 2.4:

$$J = A \exp\left(-\frac{\Delta G}{kT}\right) \tag{2.4}$$

where A is the pre-exponential factor, k the Boltzmann constant and ΔG the Gibbs free energy therm.
In order to explain the thermodynamic behind this process, a further analysis of the Gibbs free energy is necessary. For simplicity, the nuclei are considered as a sphere of radius r. The overall Gibbs free energy between a solid particle in the solute and the solute itself is given by the sum of the surface excess free energy ΔG_s, which is positive (so not thermodynamically favored) and proportional to the radius of the nucleus following the relation described by equation 2.5, and the volume excess free energy ΔG_V, negative (thermodynamically favored) and also proportional to the radius following Equation 2.6.

$$\Delta G_s \propto r^2 \tag{2.5}$$

$$\Delta G_V \propto r^3 \tag{2.6}$$

As such, the overall term is illustrated in Equation 2.7.

$$\Delta G = \Delta G_s + \Delta G_V = 4\pi r^2 \gamma + \tfrac{4}{3}\pi r^3 \Delta G_v \tag{2.7}$$

where ΔG_v is the free energy change of the transformation per unit volume and γ is the interfacial free energy between the developing crystal surface and the supersaturated solution.

Since the two components of Gibbs free energy depend differently on r, the free energy function passes through a maximum, ΔG_{cr}, corresponding to the critical nucleus dimension $r*$, which represents the minimum size of a stable nucleus, as shown in Figure 2.3.

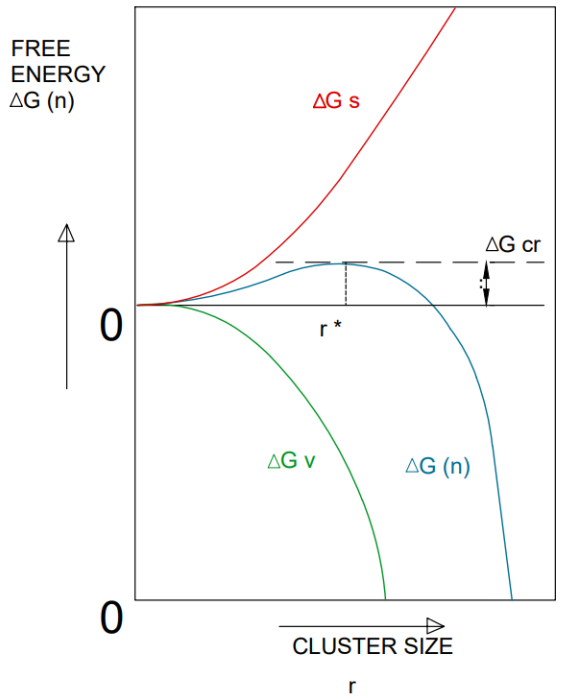


Figure 2.3: Gibbs free energy variation dependent on the size of the crystal nucleus r (from [4] and modified)

Through the analysis of Figure 2.3 it is possible now to explain the thermodynamic tendencies of crystal growth: for small particles, with $r < r*$, the positive free surface energy factor prevails, resulting in the process of molecule addition to the cluster not being thermodynamically favored, and thus particles smaller than $r*$ will dissolve. Vice versa, for $r > r*$ the free volume energy is the main contributor, increasing the free energy of the system and leading to cluster disaggregation.

The critical radius can be found through minimization of the Gibbs free energy function

(Equation 2.8), resulting in Equation 2.9.

$$\frac{d(\Delta G)}{dr} = 8\pi r^* \gamma + 4\pi r^{*2} \Delta G_v \tag{2.8}$$

$$r^* = -\frac{2\gamma}{\Delta G_v} \tag{2.9}$$

Considering the growth of the crystal cluster as governed by Gibbs-Thompson Equation (Equation 2.10) and substituting this term in Equation 2.4, the final equation for the nucleation rate is obtained (Equation 2.11), through which it is also possible to highlight the main variables that govern the phenomenon:

$$\Delta G_{cr} = \frac{16\pi\gamma^3\nu^2}{3k^3T^3(\ln S)^2} \tag{2.10}$$

$$J = A \exp\left[-\frac{16\pi\gamma^3\nu^2}{3k^3T^3(\ln S)^2}\right] \tag{2.11}$$

where T represents the temperature, S the degree of supersaturation, γ the interfacial free energy, ν is the molecular volume and k the Boltzmann constant.

It is also important to note that temperature has an effect on the critical Gibbs free energy, as shown in equation 2.10: lower temperatures lead in fact to lower values of critical energy, which results in a maximum point of the function shifted to the left and, therefore, to a lower value of critical radius, as depicted in Figure 2.4. This is also linked to the effect that temperature has on the degree of sovrasaturation: the higher the temperature, the higher the solubility of the crystals and consequently the easier it is to reach higher degrees of supersaturation, thus facilitating nucleation by reducing the energetic barrier and promoting faster crystal formation.

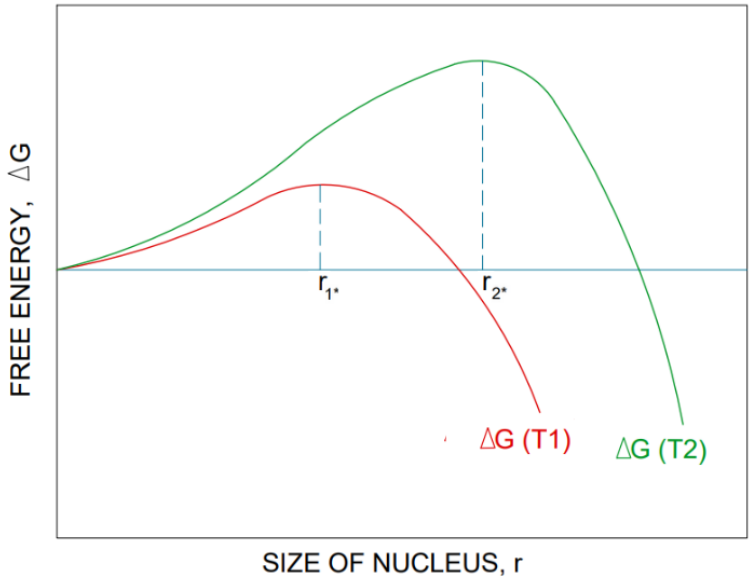


Figure 2.4: Effect of Temperature on the critical radius value (From [3] and modified)

Heterogeneous nucleation

Homogeneous nucleation is not a common event, most cased in which it is assumed that it has occurred are later deemed to be induced in some ways, whether by unknown seeding or presence of impurities in the system, since there is a greater chance of samples being contaminated with active heteronuclei. As those nuclei constitute a source of nucleation, the overall free energy change associated with heterogeneous nucleation, $\Delta G'$, is less than the one associated with homogeneous conditions.

In this case, equation 2.11 can be modified as shown in equation 2.12, which includes a factor $f(\theta)$ that keeps track of the variation in superficial tension:

$$J_{\text{het}} = A \exp\left[-\frac{16\pi\gamma^3\nu^2 f(\theta)}{3k^3T^3(\ln S)^2}\right] \tag{2.12}$$

During nucleation, both mechanisms happen at the same time, as such, the actual rate of nucleation is the sum of both homogeneous and heterogeneous nucleation rates (Equation 2.13).

$$J = J_{hom} + J_{het} \tag{2.13}$$

Secondary nucleation

The term secondary nucleation refers to the phenomenon in which nucleation occurs more readily when crystals of the solute are already present in the solution. There are many mechanisms thank to which this can happen, such as:

- Attrition: the crystals already in the solution collide with each other, the walls of

the reactor or the stirrer, resulting in crystals more strained and stressed and thus having an increased chemical potential;

- Initial breeding;
- Needle breeding;
- Fracture of the crystals;
- Fluid shear;

The size of the crystal seeds is influential in secondary nucleation: in fact, large seeds generate more secondary nuclei in agitated system than smaller seeds because of their greater contact probabilities and collision energy. As a result, increased supersaturation increases secondary nucleation and decreases the latent period.
Some other process parameters, like temperature, pressure, composition, stirring and cooling rate, can affect the probability of secondary nucleation.This phenomenon is usually considered undesirable when it occurs repeatedly within the same system, particularly in experimental studies, as it may result in a non-homogeneous crystal size distribution. However, nucleation induced by deliberate seeding is still classified as secondary nucleation and is generally preferred over primary nucleation, since the latter produces a large number of crystals with low growth rates, which is not desirable when the aim is to evaluate growth kinetics.

2.1.4 Crystal growth

After nucleation, the crystals in the solution grow by including surrounding solute molecules and increasing their dimension. When trying to describe the growth rate from a mathematical point of view, the structure in the crystal is approximated by individuating its characteristic dimension when compared to a known geometrical form, as shown in Figure 2.5. This is particularly true in the case of non-monodimensional crystals.

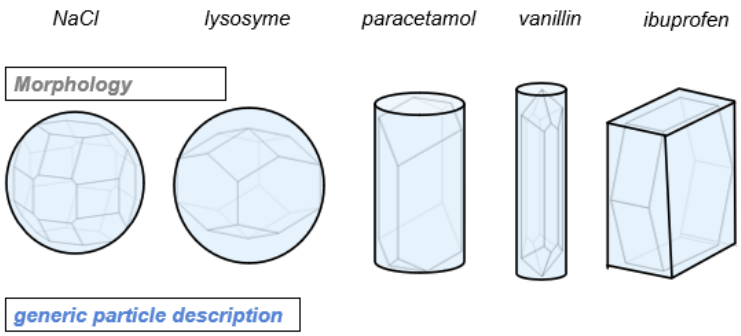


Figure 2.5: Crystal dimension approximation

Generally, then, the growth rate is described as the velocity in which the particle's char-

acteristical dimensions L_i increases as a function of time (equation 2.14).

$$G = \frac{dL_i}{dt} \tag{2.14}$$

The velocity of growth of a crystal face is measured by the outward rate of movement in direction perpendicular to that face. The many theories in which growth occurs and their implications are detailed below.

Surface energy theories [3]

The surface energy theories are based on the principle that the total free energy of a crystal in equilibrium with its surroundings at a constant temperature and pressure would be a minimum for a given volume (Equation 2.15). Therefore, when a crystal grows in a supersaturated solution, it should develop in a manner as to ensure that the whole crystal has a minimum total surface free energy: consequently, crystal faces should grow at rates proportional to their respective surface energies, and thus, the velocity of growth will vary from face to face.

$$\sum_{i=1}^{n} A_i \gamma_i = \min \tag{2.15}$$

where A_i represents the area of a specific face, and γ_i the surface energy per unit area of the ith surface.

Experimentally then, smaller faces grow faster in comparison to large faces, and so it is possible for a crystal not to maintain geometric similarity during its growth, as faster-growing faces gradually disappear, as pictured in Figure 2.6.

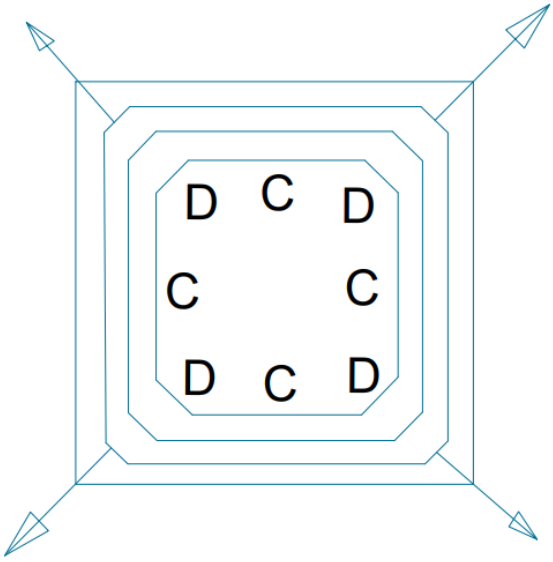


Figure 2.6: Velocities of crystal growth faces, here the D face of the crystal is shown to gradually disappear during its growth, from [3] and modified

Adsorption layer theories

The adsorption layer theories are based on the concept of existence of an adsorbed layer of solute atoms or molecules on a crystal face: when crystallizing units arrive at the crystal face, they are not immediately integrated, but instead lose one degree of freedom and are able to migrate over the crystal face via surface diffusion, where they will link into the bulk crystal in positions where the attractive forces are greatest.[3]

The presence of dislocations and imperfections on the bulk crystal surface can then promote attraction of solute molecules and face growth, solving the difficult thermodynamic problems of needing a two-dimensional nucleation for the adsorption growth to occur. This mechanism, in presence of a screw dislocation, is known as spiral growth mechanism: screw dislocations, as in "steps" in the crystal surface, provide energetically favorable access points for incorporating molecules. Since the step originates from a fixed position, the crystal particles that aggregate make such step spiral around the dislocation [6].

Diffusion-reaction theories

The diffusion-reaction theories define the deposition of solid on the face of a growing crystal as a two-step process:

- Diffusion process: the solute molecules are transported from the bulk of the fluid phase to the solid surface (Equation 2.16):

$$\frac{dm}{dt} = k_d \, A \, (c - c_i) \tag{2.16}$$

where k_d is a coefficient of mass transfer by diffusion and c_i the solute concentration in the solution at the interface;

- First-order reaction process: the solute molecules arrange themselves into the crystal (Equation 2.17)

$$\frac{dm}{dt} = k_r\, A\, (c_i - c^*) \tag{2.17}$$

in which k_r a rate constant for the surface reaction (integration) process, and $c*$ the concentration at equilibrium.

In both cases the driving force is a concentration difference, better illustrated in Figure 2.7.

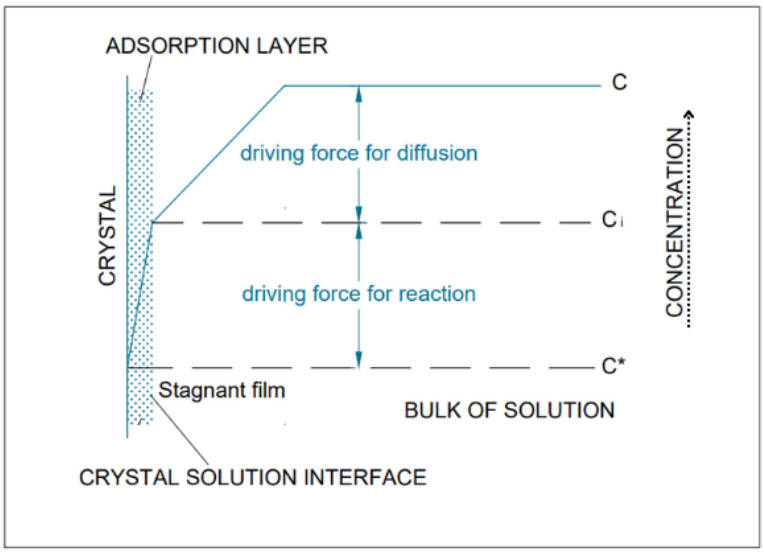


Figure 2.7: Concentration driving forces, from [3] and modified

A general equation for crystallization based on the overall driving force can be written as:

$$\frac{dm}{dt} = K_G\, A\, (c - c^*)^g \tag{2.18}$$

where K_G is the overall crystal growth coefficient, g is the order of the overall crystal growth process.

2.1.5 Surface Structure and Interfacial Energies in Crystal Growth

It exists an important connection between the structure of a crystal face and the modality of growth adopted. This property is characterized by the surface roughness or surface entropy factor, also known as the alpha factor (Equation 2.19).

$$\alpha = \frac{\xi \Delta H}{kT} \tag{2.19}$$

where ξ is an anisotropy factor related to the bonding energies in the crystal surface layers,ΔH is the enthalpy of fusion and k is the Boltzmann constant.
Low values of α are indicative of a rough crystal surface that will allow continuous growth to proceed, whereas higher values indicate smoother surfaces, which makes it more difficult for solute molecules to attach to the face of a growing crystal.
Interface structure changes can sometimes explain the changes undergone by the crystal in terms of solvent dependance: generally, the higher the solubility of the solute, the lower the α-value and consequently rougher the surface [7]. The surface roughness, as explained above (see Section 2.1.4), influences the velocity and probability of attachment of solute molecules on the surface of the pre-existing crystal, thus influencing the growth rate and impacting the final shape of specific faces.

Wullf construction

A way to predict the equilibrium shape of a crystal, hence its morphology during growth, is the Wulff construction [8] [9]. It is a graphic method that requires the surface energy to be known: said surface energies, which vary in order to minimize the free energy following equation 2.15, are selected in reference to the surface area, assumed as constant (equation 2.20). By individuating the values of surface energies, it is then possible to construct the shape of that face of the crystal at equilibrium, as depicted in Figure 2.8.

$$\gamma = \frac{E_{surf}}{A} \tag{2.20}$$

The surface free energy, howeve,r is complex to determine through conventional empirical methods, hence why in the lates decades molecular dynamic simulations and thermodynamic integrations have been employed (see Section 2.2).

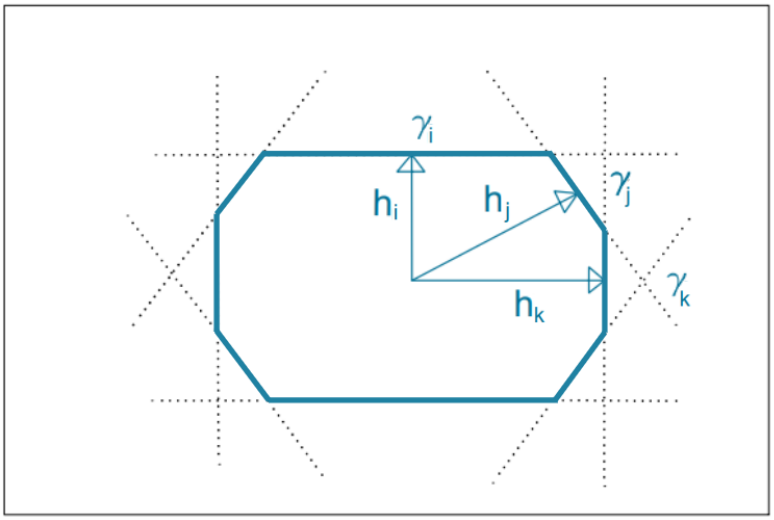


Figure 2.8: Example of Wullf construction, γ are the interfacial free energies and h the directions perpendicular to the accounted faces. The ticker, black outline shows which shape the face will undertake in equilibrium conditions, from [9] and modified

2.1.6 Growth rate measurement

To determine the growth rate of a specific crystal face, measurements made on population of crystals are often used to determine the overall mass transfer rate under controlled conditions and determine the bulk volume of the crystals.[10]
The general models that take in consideration the development of crystals in a solution, such as behaviors like birth, growth, agglomeration and death, are known as population balance equations (PBE), of which the general form is:

$$\frac{\partial n(L,t)}{\partial t} + G(L,t)\,\frac{\partial n(L,t)}{\partial L} = B(L,t) - D(L,t) \tag{2.21}$$

where the first term considers the evolution of the population density (also known as number density) of the crystals through time, taking into account the characteristic dimension of the crystals, L, and time t; G is the growth rate, allowing the second term to highlight how the crystals population varies in their size through time due to growth; $B(L,t)$ crystal birth (due to nucleation or breakage) and $D(L,t)$ the crystal death (due to agglomeration or dissolution). A simplified schema of the process is shown in Figure 2.9

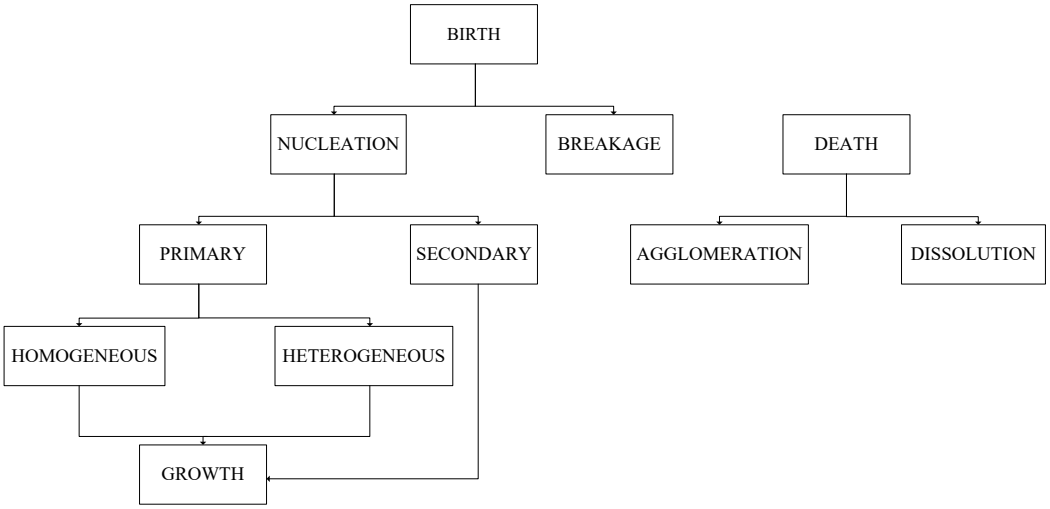


Figure 2.9: Parameters considered in PBE, from [10] and modified

Using simplifying assumptions, analytical solutions can be found for the population balance equation. However, the simplest and most efficient way to solve them is to implement the method of moments (MOM).

Method of Moments

The method of moments involves converting the PBE into moments equation of the number density and using those values to find global properties of the distribution. The i^{th} moment is defined in equation 2.22, and the general form of the PBE with the MOM application is given in equation 2.23.

$$\mu_i = \int_0^\infty f(L,t)\, L^i\, dL \tag{2.22}$$

$$\frac{dm_i(t)}{dt} = (0)^i B(t) + \int_0^\infty iL^{i-1}(L)n(L,t)\, dL + B_{i,a}(t) - D_{i,a}(t) + B_{i,b}(t) - D_{i,b}(t) \tag{2.23}$$

where m_i is the i^{th} moment ($\frac{m^i}{m^3}$), $B(t)$ is the birth rate due to nucleation, $B_{i,a}(t)$ is the birth rate due to agglomeration, $D_{i,a}(t)$ is the death rate due to agglomeration, and $B_{i,b}(t)$

and $D_{i,b}(t)$ are the birth and death rate due to breakage, respectively.

Solving the first four moments of the population balance equation provides a comprehensive description of the crystallizing system. For simplicity, assuming that only the growth rate G and the nucleation rate B are considered, the meaning of the first four moments is as follows, where the general multidimensional equation is shown in equation 2.26:

- Zeroth moment μ_0: Represents the total number of crystals per unit volume (equation 2.24);
- First moment μ_1: Corresponds to the total length of all crystals per unit volume (equation 2.25);
- Second moment μ_2: Provides information related to the total surface area of the crystals, assuming a simplified geometry (equation 2.25)
- Third moment μ_3: It is proportional to the total volume (or mass) of crystals present in the system (equation 2.25)

$$\dot{\mu}^{(0)} = B_0 \tag{2.24}$$

$$\dot{\mu}_{i_1 \ldots i_k} = -\sum_{r=1}^{k} g_{i_r}\, \mu_{i_1 \ldots i_{r-1}\, i_{r+1} \ldots i_k} \;+\; B_0 \prod_{r=1}^{k} \ell_{0,i_r} \tag{2.25}$$

where g represents the crystal decay term, quantifying the strength of the dependence of the $k-th$ moment μ_k on the $(k-1)-th$ moment μ_{k-1}. In the mono dimensional case, g is a scalar; in multidimensional formulations it can instead take the form of a vector ($\mathbf{g} \in R^d$) or, more generally, a matrix ($A \in R^{d \times d}$).

Each moment integrates the particle size distribution across different size classes, allowing for a simplified but meaningful description of the crystallization process. These values are fundamental for process control and scale-up, as they relate directly to measurable physical properties.

The method of moment can also be extended into the Quadrature Method of Moments (QMOM) by approximating the integral with a quadrature, as shown in equation 2.26. When discrete crystal size distributions are available, the QMOM provides clear advantages over the classical method of moments. QMOM reconstructs the particle size distribution from a small set of moments, ensuring realizability and enabling more accurate tracking of nucleation, growth, and aggregation dynamics. In contrast, standard MOM relies on moment closures and may produce non-physical results, especially when higher-order dynamics are relevant.

$$m_i = \int_0^{\infty} f(L)\, L^i\, dL \;\approx\; \sum_{j=1}^{n} w_j\, L_j^i \tag{2.26}$$

where L_j^idenotes the representative crystal size of the $j-th$ class and w_j the associated weighting factor, corresponding to its relative frequency in the distribution.

2.1.7 Techniques for crystal characterization and analysis

Polarized Light Microscopy

Polarized Light Microscopy (PLM) is an optical technique used to enhance image contrast when analyzing samples containing birefringent materials. This method is particularly effective for the qualitative and quantitative examination of anisotropic substances, such as crystalline compounds, due to its high sensitivity to optical anisotropy.[11]
The technique relies on the interaction between polarized light and birefringent structures within the specimen. A typical polarized light microscope is equipped with two key optical elements: the polarizer, positioned below the sample, and the analyzer, placed above the objective in the optical path. When plane-polarized light passes through a birefringent material, it splits into two orthogonal wave components that travel at different speeds. These components are later recombined by the analyzer, resulting in constructive and destructive interference patterns.
This interference enhances contrast and reveals detailed structural features that are often invisible under standard brightfield illumination. For this reason, PLM is especially useful for observing crystals, as it helps clearly show their shape, orientation, and how they interact with light. The technique makes it easier to see small differences in the structure that would be hard to detect with regular light microscopy.

Turbidity analysis

During a crystallization process, turbidity analysis is important both to evaluate the beginning of the process and, indirectly, the occurrence of other phenomena, i.e. secondary nucleation. This method is based on the principle of measuring the transmission of light though a turbid solution, in order to account for variations in terms of transmittance and absorbance. The two parameters are linked to one another following equation 2.27, and have opposite trends: an increased value of transmittance leads to a decrease in the absorbance, and vice versa.

$$A = -log_{10}T \tag{2.27}$$

Transmittance (T) is defined as the fraction of incident light that passes through a sample:

$$T = \frac{I}{I_0} \tag{2.28}$$

where I_0 is the intensity of the incident light and I is the intensity of the transmitted light. It is commonly expressed as a percentage ($\%T$), where higher values indicate clearer solu-

tions (i.e., less absorbance or scattering). On the other hand, absorbance is a logarithmic measure of the amount of light absorbed by the sample, and it is directly proportional to the solute concentration, given constant path length and molar absorptivity. As crystallization proceeds and solute leaves the solution to form crystals, transmittance increases and absorbance decreases, indicating a lower concentration of dissolved solute. [12]
For this type of analysis immersion probes are often used, in which a constant LED ray goes through the process fluid: the ray is intercepted by a sensor that measures how much the intensity of the transmitted light has been toned down due to turbidity. An example of such instrumentation is depicted in Figure 2.10.

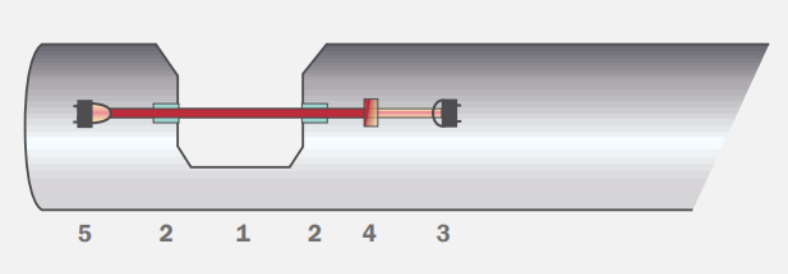


Figure 2.10: Insertion sensor, where 1) Optic path, 2) Optics in sapphire, 3) Measuring sensor, 4) Ambient light Filter, 5) LED source

Statistical Reliability of Crystal Size Distribution Measurements

To ensure the reliability of any analysis conducted on crystal populations, it is essential to verify that the investigated sample is statistically significant; namely, that the number of particles analyzed is representative of the entire population batch or, in certain cases, that the statistical estimators can properly represent the sample analyzed.[13]
To this end, two key statistical estimators are employed:

- $\bar{x}_n$, cumulative running average (Equation 2.29): This parameter reflects the cumulative mean of a property across the first n analyzed particles. Ideally, the running average should converge as n increases. Such convergence indicates that the analyzed sample is large enough to statistically represent the population of the sample at the given time:

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2.29}$$

- s, cumulative running standard deviation (Equation 2.30): This value measures the dispersion of the observed property across the first n particles. Similar to the average, convergence of the standard deviation towards a stable value suggests that the variability within the sample is well captured and additional sampling is unlikely

to alter the population statistics significantly:

$$s = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{2.30}$$

To further validate the reliability of the calculations, a comparison between the values obtained through running statistics and those derived from evaluation of the moments of the sampled distribution was performed. In particular, the mean and variance calculated via the moments should match the cumulative estimators if the calculations are correct.

As such, the mean and variance values may be written as a function of the sample raw moments, following equation 2.31 and 2.32:

$$\bar{L} = \frac{M_1}{M_0} \tag{2.31}$$

$$S^2 = \frac{M_0 M_2 - M_1^2}{M_0(M_0 - 1)} \tag{2.32}$$

where M_0 is equal to the total number of sampled particles N, and M_i represents the ith moment for the given system of crystals.

From Equation 2.32, the common equation for standard deviation can be written as:

$$S = \sqrt{S^2} \tag{2.33}$$

In cases where more than one dimension is analyzed (e.g., both length and width of the crystals), the same statistical criteria apply. However, the most descriptive property may vary: for instance, in a 2D context, the average projected area might be the most relevant parameter, as it captures both dimensions' effects simultaneously. As such, equations 2.31 and 2.32 can be generalized for n-dimensions as:

$$\bar{X} = \frac{M_n}{M_0} \tag{2.34}$$

$$S^2 = \frac{M_0 M_{2n} - M_n^2}{M_0(M_0 - 1)} \tag{2.35}$$

The moments are primarily employed to estimate statistical descriptors of a population. In the context of this work, it is particularly useful for validating the representativeness of the samples by estimating the confidence interval (CI) associated with the measured property.

The confidence interval defines the range within which the true value of a population parameter is expected to lie, at a given confidence level α. It is delimited by an upper

limit (U_{limit}) and a lower limit (L_{limit}), as shown in Equation 2.36, while the interval width (CI_W) is given by their difference. The value of α indicates the proportion of measurements expected to fall within the interval (typically between 80% and 99%). The corresponding t-value, $t_{\infty,\alpha/2}$, is derived from the two-tailed Student's t-distribution with infinite degrees of freedom (a valid approximation when N is large).

$$CI_W = U_{limit} - L_{limit}, \quad where: \quad \begin{cases} U_{limit} \\ L_{limit} \end{cases} = \begin{cases} \bar{X} + 2t_{\infty,\alpha/2} \frac{S}{\sqrt{N}} \\ \bar{X} - 2t_{\infty,\alpha/2} \frac{S}{\sqrt{N}} \end{cases} \tag{2.36}$$

where S is the sample standard deviation, $\bar{X}$ is the sample mean of the measured property, and N is the number of sampled particles.
The evaluation of the confidence interval can be performed using the Central Limit Theorem (CLT), which states that, within a population of collected samples N, the sampling distribution of the mean $\bar{x}$ is increasingly concentrated around the mean of the population μ and converges to a standard Gaussian distribution as N increases, as better shown in Figure 2.11[14].

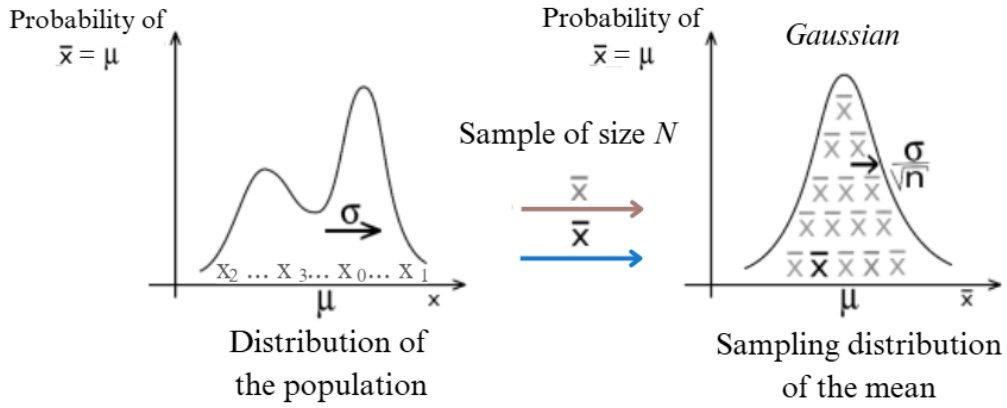


Figure 2.11: Rapresentation of the Central Limit Theorem, from [14] and modified

From Equation 2.36, it is also possible to invert the formulation in order to estimate the minimum number of particles (N) required to achieve a target level of precision, expressed as a desired confidence interval:

$$N = (\frac{2t_{\infty,\alpha/2}}{benchmark})^2 \tag{2.37}$$

This formulation is essential when planning experiments, as it provides a statistical basis for determining whether the number of samples is sufficient to make reliable conclusions.

In addition to the confidence interval, the calculation of the Relative Confidence Interval

(RCI) provides a dimensionless measure of statistical uncertainty by normalizing the interval width with respect to the mean. Unlike the absolute confidence interval, the RCI enables direct comparison between samples with different average crystal sizes, allowing the reliability of the measurements to be assessed on a relative basis. Its equation is:

$$RCI = \frac{U_{limit} - L_{limit}}{\bar{X}} \tag{2.38}$$

where $\bar{X}$ is the average value, and U and L are the upper and lower limit of the confidence interval, respectively.
In order to enable a qualitative assessment of sample representativity, an arbitrary threshold value for the relative confidence interval (RCI) was defined. A cut-off of 15% was adopted: intervals with RCI $< 15\%$ were classified as narrow, whereas those with higher values were considered wide. The threshold of 15% was selected based on the common criterion in scientific practice whereby a relative error below 10% is generally regarded as acceptable according to the literature. By adopting a slightly higher limit, the definition of narrow confidence intervals (RCI $< 15\%$) ensures consistency with established standards while accounting for the experimental constraints of the present study.

2.2 State of the Art in Molecular Modeling Applied to Crystallization

2.2.1 Molecular Modeling and Molecular Mechanics

Molecular modeling enables the investigation of molecular structures across a wide range of systems by combining theoretical concepts with computational simulations. A rigorous description of atoms requires analysis at the quantum level; however, such an approach becomes impractical for systems involving thousands of atoms over very short time scales (on the order of tens of nanoseconds). This limitation persists even when employing approximations such as the Born–Oppenheimer approximation, which assumes that electrons, due to their much smaller mass, adjust instantaneously to changes in the positions of the nuclei, which constitute the heavier part of the atom.
Therefore, the solution is to resort to classical approximations, representing molecular systems through simplified models that omit certain complex features, such as explicit electronic structure, while retaining the essential physics of the problem. This approach allows the prediction not only of macroscopically measurable properties, but also of more subtle quantities, such as the potential energy and its individual contributions.

Molecular modeling is applied across numerous STEM disciplines, including computational biology, computational chemistry, spectroscopy, and materials science. Within this

framework, Molecular Mechanics (MM) represents molecular systems by modeling atoms as semi-rigid spheres of radius r, centered at the positions of the nuclei. These spheres interact with one another through forces derived from classical potential functions that approximate quantum-mechanical interactions. Molecular Mechanics simulations therefore involve evaluating the interactions among these spheres across different configurations; namely, by sampling multiple relative positions. It is important to note that, for practical purposes, molecular mechanics models must be parametrized: choosing appropriate parameters is a critical step to accurately simulate the molecular environment and obtain reliable results.

In this context, tracking the positions of molecules is inherently connected to monitoring their changes over time, particularly when studying conformational rearrangements or dynamic interactions. To investigate such complex phenomena, molecular dynamics provides an essential computational framework.

2.2.2 Molecular Dynamics

Molecular Dynamics (MD) is a computational technique used to simulate the time evolution of a system of interacting particles (typically atoms or molecules), based on the laws of classical mechanics. It allows the analysis of the time-dependent behavior of molecular systems, by updating the positions and velocities of particles over time by numerically solving Newton's equations of motion, while the forces acting on the particles are determined by empirical potentials (force fields), which depend on the geometric configuration of the system.[15]

Newton's equations of motion are applied to a system of N particles. In particular, the definition of the force F_i acting on a particle is expressed through Newton's second law, which states that it is proportional to its mass and acceleration (expressed as the second derivative of the particle's position r_i):

$$F_i = m_i \frac{\partial^2 r_i}{\partial t^2} \tag{2.39}$$

and also as the gradient of the system's potential energy $E(R)$:

$$F_i = -\nabla_{r_i} E(R) \tag{2.40}$$

giving as a result a set of differential equations that must be solved for all particles in the system, in order to determine the evolution through time of their velocities and positions (Equation 2.41).

$$m_i \frac{\partial^2 r_i}{\partial t^2} = -\nabla_{r_i} E(R) \tag{2.41}$$

Since equation 2.41 is usually not analytically solvable, due to the complexity in estimating the different parameters, numerical integration methods are employed. Common schemes include Euler, Verlet, and Velocity-Verlet, which integrate the equations over a chosen timestep (Δt).
The update of particle positions and velocities uses either the current and next timestep (t and $t+\Delta t$, in the Euler scheme) or the previous, current, and next timestep ($t-\Delta t$, t, and $t+\Delta t$ in the Verlet scheme). As such, the choice of the timestep is a crucial point for the simulation of the system, and mainly depends on the oscillation frequency of the studied phenomena: it must be sufficiently small to accurately describe the system's dynamics, but too little of a timestep would result in computationally inefficient simulations. In practice, it is on the order of femtoseconds (10^{-15} seconds), corresponding to the oscillation period of the H atom bonded in a molecule.

This approach allows for the investigation of a wide range of physical and chemical phenomena at the atomic and mesoscopic scale, such as phase transitions, diffusion processes, thermomechanical behavior, and structural properties. The versatility of the method makes it applicable across various scientific fields, from crystalline materials to polymers and complex biological systems.

The parameters and functions that must be defined to ensure a realistic and reliable simulation are discussed in the following subsections.

Force Fields

As previously stated, a force field is a set of mathematical functions and associated parameters used to describe the potential energy function (PEF) of a molecular system as a function of the positions of its atoms[16]. For a system composed of N atoms, each represented by a position vector r_i, the overall potential energy, also known as potential energy surface (PES), is expressed as:

$$V(r_N) = V_{\text{bond}}(r_N) + V_{\text{non-bond}}(r_N) \tag{2.42}$$

where $V_{\text{bond}}(r_N)$ describes forces with atoms connected through chemical bonds, including aspects as bond stretching, angle bending and torsional rotations; and $V_{\text{non-bond}}(r_N)$ represents forces between atoms not directly bonded, such as Van der Waals attractions and electrostatic interactions. A summary of the different interactions is illustrated in Figure 2.12.

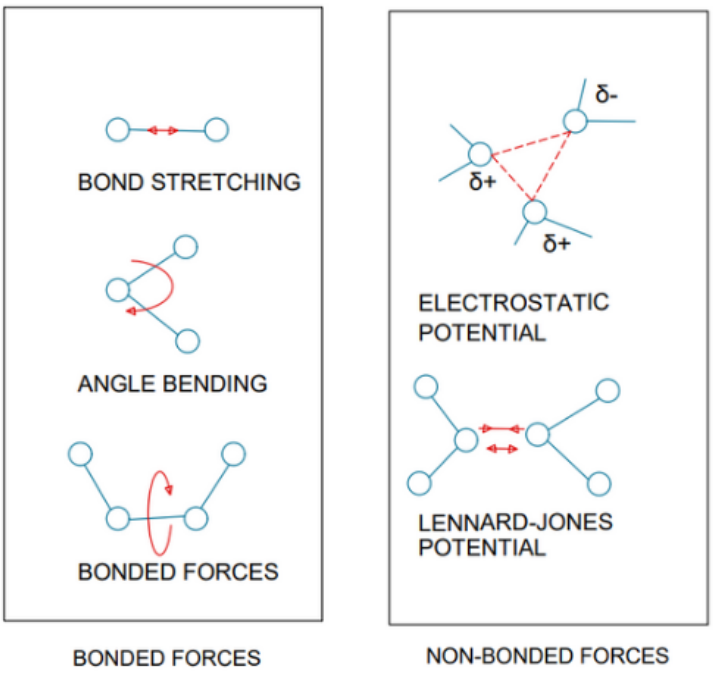


Figure 2.12: Bonded (left) and non-bonded (right) interactions considered in MD simulations, from [17] and modified

In MD simulations, examples of commonly used force fields include TraPPE[18], OPLS[19], AMBER[20], and CHARMM[21]. Each of these has been parametrized for specific classes of systems and applications: for instance, TraPPE[18] (Transferable Potentials for Phase Equilibria) is optimized for accurate prediction of thermophysical properties in phase equilibrium studies, while OPLS[19] (Optimized Potentials for Liquid Simulations) focuses on reproducing liquid-state properties. AMBER[20] and CHARMM[21], on the other hand, are widely used in biomolecular simulations, offering extensive parameter sets for proteins, nucleic acids, and other biological macromolecules. The choice of force field is therefore dictated by the chemical nature of the system under study and the target properties to be predicted.

Analyzing each term of equation 2.42 singularly, the bonded interactions in MD refer mainly to atoms connected by covalent bonds, and are divided in three categories:

- Bond stretching, involving two atoms, it is expressed through Hooke's law:

$$V_{\text{bonds}}(d) = \frac{k_i}{2}\,(d_i - d_{i,0})^2 \tag{2.43}$$

  where k is the stretching constant of the bond, and d_0 is the reference bond length between the two particles considered;
- Angle bending, involving three atoms, it is often represented using a harmonic potential, and it expresses the energy linked to deviations from the ideal valence

angle defined within three atoms:

$$V_{\text{angles}}(\theta) = \frac{h_i}{2}\left(\theta_i - \theta_{i,0}\right)^2 \tag{2.44}$$

in which h is a force constant and θ_0 is the reference value for the valence angle;

- Torsional or dihedral interactions, involving four atoms, it explicates the energy changes occurring during the rotation around a bond involving four atoms, and it is modeled using a sinusoidal function:

$$V_{\text{dihedrals}}(\phi) = \frac{V_n}{2}\left(1 + \cos(n\phi - \varphi)\right) \tag{2.45}$$

the parameters of this equation are the barrier term V_n, expressing the energy required for rotation, the multiplicity n, indicating the number of minimum potential points encountered in a full rotation of the molecule, and the phase factor ϕ that defined the position of the different minima.

As such, the general equation for bonded interaction is:

$$V_{\text{bond}}(r_N) = \sum_{\text{bonds}} V_{\text{bonds}}(r_N) + \sum_{\text{angles}} V_{\text{angles}}(r_N) + \sum_{\text{dihedrals}} V_{\text{dihedrals}}(r_N) \tag{2.46}$$

Turning back to the original equation for the overall potential energy surface (Equation 2.42), the non-bond interaction refers to atoms not connected by covalent bonds, and in MD are modeled using functions that depend on the distance between the particles (r). It consists mainly of two terms, resulting in Equation 2.47:

$$V_{\text{non-bond}}(r_N) = \sum_{i=1}^{N}\sum_{j=i+1}^{N} V_{\text{elect}}(r_{ij}) + \sum_{i=1}^{N}\sum_{j=i+1}^{N} V_{\text{VdW}}(r_{ij}) \tag{2.47}$$

The individual terms are as follows:

- Electrostatic forces: long-range interactions resulting from differences in electronegativity among atoms, leading to unbalanced distribution of charges within the molecule itself, they are mainly expressed using Coulomb's law:

$$V_{\text{elect}}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0\, r_{ij}} \tag{2.48}$$

where q_i and q_j are the net charges of the two atoms analyzed, r_{ij} is the distance between them and ϵ_0 is the dielectric constant;

- Van der Waals forces: they describes short-range interactions and its trend is com-

monly described using the Lennard-Jones potential:

$$V_{\mathrm{VdW}}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{2.49}$$

The Lennard-Jones potential is characterized by two terms, better represented in Figure 2.13: σ, which represents the distance at which the potential energy is zero, also known as collision diameter, and ϵ, that indicates the minimum of the potential.

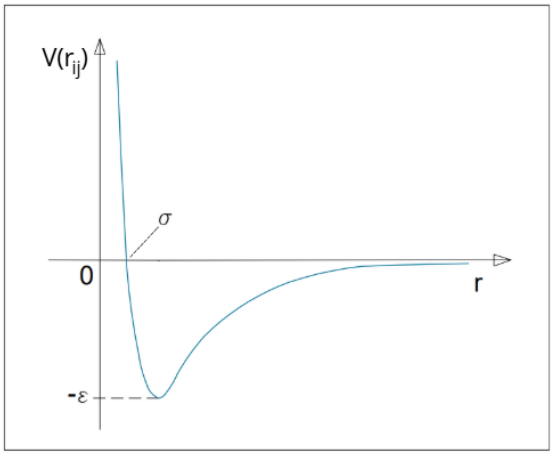


Figure 2.13: Trend of Lennard-Jones potential, from [22] and modified

Through Figure 2.13 it is possible to analyze the characteristic behavior of the Lennard–Jones potential, and consequently of Van der Waals interactions. At short intermolecular distances the potential rises steeply, reflecting strong repulsive forces, whereas at longer distances it asymptotically approaches zero. The Lennard-Jones potential has a minimum for a specific value of r, r_0: for $r_{ij} < r_0$,

$$\frac{dV(r_{ij})}{dr} < 0, \tag{2.50}$$

and the corresponding force

$$F(r_{ij}) = -\frac{dV(r_{ij})}{dr}, \tag{2.51}$$

is positive, indicating repulsion. Viceversa, for $r_{ij} > r_0$ the derivative of the potential is positive, so the force is negative, corresponding to attractive interactions. In other words, the ϵ parameter represents the imaginary line of distinction between attractive forces (for $r_{ij} > \epsilon$ and repulsive forces (for $r_{ij} > \epsilon$).

Realistically, even if moderately, each atom interacts with all the other atoms present in the system, leading to a simulation that would require the calculations of all atom-pair combinations: however, simulating and calculating the atomic interactions in this way

would not only lead to a too high computational cost, but also would not result in models potentially appliable elsewhere than the single simulation itself. As such, there are several techniques available to reduce the computational cost and to accelerate non-bonded force computations, the most used of which is the introduction of a cut-off radius.

Cut-off radius

In an MD simulation, the system is typically represented as a simulation box containing the atoms of interest, which can be further subdivided into smaller cells to reduce computational complexity. To evaluate non-bonded interactions between atom pairs, a cut-off distance is defined (Figure 2.14): this spherical truncation allows only the interactions between particles with a distance smaller than the cut-off radius (R_c) to be considered and atoms lying beyond this radius to be excluded from the computation, as their contribution is assumed to be negligible. As such, the potential energy considered follows Equation 2.51.

$$V_{\text{trunc}}(r) = \begin{cases} V_{\text{LJ}}(r), & r \leq R_c \\ 0, & r > R_c \end{cases} \tag{2.52}$$

Because atomic positions change over the course of the simulation, the list of atom pairs within the cut-off distance is updated at each computational step to ensure that only the relevant interactions are calculated. This approach maintains computational efficiency while preserving the physical accuracy of the simulation, also due to the choice of a spherical geometry of the cut-off region, which ensures that interactions are considered isotropically in all directions, avoiding directional bias that would arise from non-spherical shapes (e.g., cubic boundaries), due to the fact that the distance between the atom-pair considered would vary.

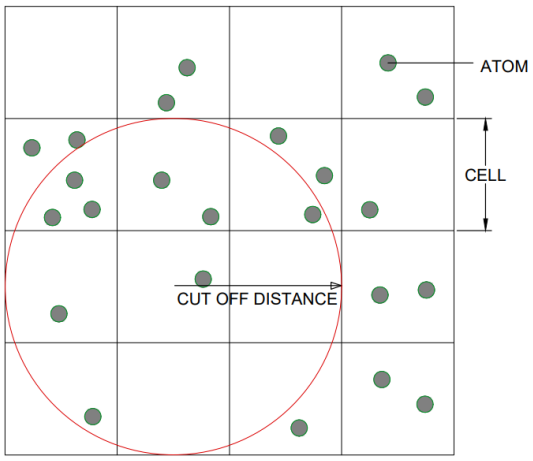


Figure 2.14: Division of the simulation box in cells and representation of the cut-off radius, from [17] and modified

Boundary Conditions

The interest of a MD simulation is to obtain the bulk properties of a liquid or solid system, as such, despite having the possibility to simulate the system of N particles in isolation surrounded by a vacuum, it is better to impose boundary conditions. The use of rigid walls is allowed but generally discouraged, as the surface effects would impact on the estimation of the real physics of the bulk system: this occurs because the fraction of atoms near the walls is proportional to $N^{-\frac{1}{3}}$. While this fraction is negligible for large systems, it becomes particularly relevant the smaller the investigated box is: for example, a box with $N = 1000$ particles would have 45% of its elements close to the walls, while in a bigger sample, with $N = 10^6$ particles, only 15% of the particles would lay in the walls region. Since not always the investigation of a bigger system is wished or necessary, it is better to use periodic boundary conditions (PBC) in any case [16].

The base principle of PBC is to surround the simulation box with an infinite number of replicas, as shown in Figure 2.15. In this case, only the atoms inside the main cell are considered, but an image particle enters from the opposite side of the box as soon as one of the atoms leaves the cell boundaries.

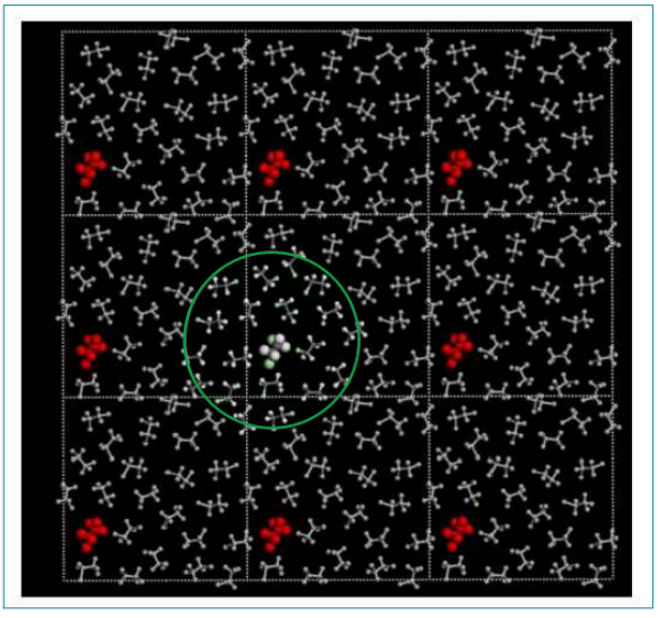


Figure 2.15: Representation of PBC (in 2D) in a MD simulation using a cubic box, from [16] and modified

PBC are used together with the minimum image convention (MIC), meaning that only the interactions with the nearest atom or image of it are considered within the boundaries of the chosen cut-off radius R_c. The cut-off radius cannot be larger than $\frac{L}{2}$ (where L is

the side of the cell) because otherwise interactions between a particle and more than one images would occur.

Ensembles

Macroscopic properties of interest can be obtained from MD simulations using statistical mechanics, and in particular the concept of ensembles, i.e., the theoretical representation of the physical system that best represents the system's conditions. Observable properties are obtained as time-averaged values over the course of the simulation, in contrast to other approaches, such as the Monte Carlo method, where properties are estimated via statistical sampling of randomly generated, time-independent configurations.
Given that macroscopic systems at equilibrium can be represented by ensembles, the value of a macroscopic thermodynamic state variable (A) is expressed as the mean value over the infinite phase space points of the ensemble:

$$\overline{A} = \langle A \rangle_{\text{ensemble}} \tag{2.53}$$

As exploring "infinite" microstates of the system is impractical, equation 2.53 can be modified by applying the ergodic hypothesis. This hypothesis states that, given a sufficiently long period of evolution, the system will pass through all of its accessible states. Therefore, the ensemble average of a state variable corresponds to its time-averaged value:

$$\overline{A} = \langle A \rangle_{\text{ensemble}} \tag{2.54}$$

In the context of Molecular Dynamics, this equivalence is crucial: it allows the calculation of macroscopic thermodynamic properties from a single, sufficiently long simulation trajectory, rather than having to explicitly sample an ensemble of systems. By averaging over time along the trajectory, properties such as temperature, pressure, or potential energy can be estimated with the same statistical meaning as ensemble averages.

The commonly used ensembles include:

- NVE (microcanonical): keeps constant the number of particles (N), the system volume (V), and the total energy (E);
- NVT (canonical): keeps N, V, and temperature (T) constant, often by introducing a thermostat;
- NPT (isothermal–isobaric): keeps N, T, and pressure (P) constant, allowing the system volume to fluctuate;
- μVT (grand canonical): keeps the chemical potential (μ), volume, and temperature constant, typically used to simulate open systems in contact with a thermal reservoir.

2.2.3 Output of interest in Molecular Dynamics Simulations

Molecular dynamics simulations can produce a broad spectrum of output data, encompassing structural, energetic, and dynamic properties of the system under investigation. In the context of the present work, attention will be devoted exclusively to those quantities that are of direct relevance to the objectives of the study. For each selected output, its computational significance, underlying physical meaning, and contribution to the overall interpretation of the results will be presented in the subsequent sections.

Density

In MD, as with most engineering fields, density ρ is derived directly from the total mass of the system, expressed in g and the simulation volume (in $\mathring{A}^3$):

$$\rho = \frac{m_{\text{tot}}}{v_{\text{box}}} \tag{2.55}$$

Evaluating the density is a fundamental first step to validate both the realism of the constructed system and the adequacy of the chosen force field. Significant deviations from experimental reference values may indicate an inappropriate force field or an incorrect initial box volume.
In addition, monitoring the density profile over time provides valuable insight into the system's equilibration. In an NPT ensemble, stabilization of the density around a constant value generally indicates that equilibrium has been reached. Conversely, a lack of stabilization may suggest that the NPT parameters impose volume changes too abruptly, therefore requiring adjustment.

Radial Distribution Function

The Radial Distribution Function (RDF) is a mathematical function used to describe the spatial organization of atoms within a system by evaluating the probability of finding a particle at a given distance from another. It analyzes the relative atomic density as a function of radial distance (cutoff radius) from a reference atom and enables identification of local order and interactions between the atomic species present.

This property is particularly useful for assessing the accuracy of the force field used in the simulation, as it allows comparisons between the atomic arrangement predicted by the model and the expected configuration for the simulated physical state.

RDF graphs (Equation 2.56) exhibit a characteristic pattern, with peaks corresponding to the average interatomic distances in the various interaction "shells".

$$g(r) = \frac{dN}{dV\,\rho} = \frac{\rho_L(r)}{\rho} \tag{2.56}$$

where dN is the number of particles within an infinitesimal volume dV, $\rho_L(r)$ density at a radial distance r and ρ is the macroscopic density ($\frac{N}{V}$).

Typical RDF plots show a sequence of peaks corresponding to the average interatomic distances within successive coordination shells. In ordered systems such as crystals, these peaks are sharp and extend over long distances; in liquids, the ordering decays beyond the first shell, and $g(r)$ tends towards 1 at large r. For a better understanding of the function trend, an example of a typical RDF function for water is shown in Figure 2.16.

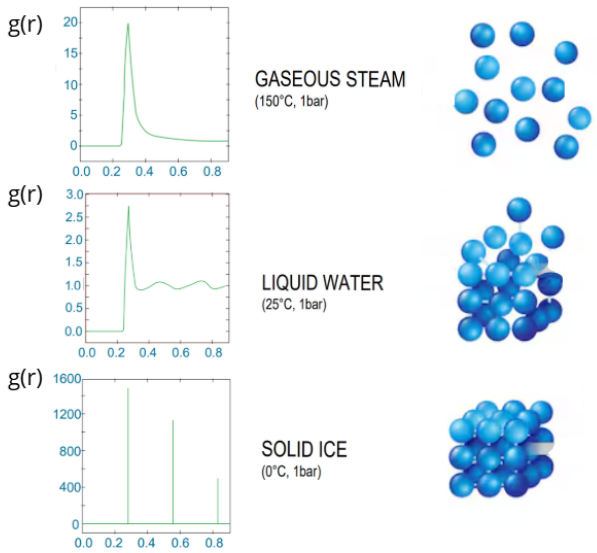


Figure 2.16: Example of RDF for water in different phases: gaseous water (top), liquid water (middle), solid ice (bottom), from [23] and [24] and modified

Diffusivity

Diffusivity describes the rate at which molecules move within a system and provides critical information about liquid mobility, Brownian motion, and molecular relaxation times.

In molecular dynamics, this property is typically determined using two main approaches:

- Velocity Autocorrelation Function Method (VACF), a method based on the time integration of the velocity autocorrelation function:

$$D = \frac{1}{3} \int_0^\infty \langle v(t)v(0) \rangle \, dt \tag{2.57}$$

  where $v(t)$ is the velocity vector of a particle at time t, and $\langle v(t)v(0) \rangle$ denotes the ensemble and time average. The factor $\frac{1}{3}$ accounts for averaging over the three spatial degrees of freedom. In practice, the integration is carried out numerically up to the point where the correlation decays to zero, excluding statistical noise;

- Mean Square Displacement Method (MSD), based on the mean square displacement of the particles:

$$D = \frac{\text{MSD}}{6} = \lim_{t\to\infty} \frac{1}{6t} \left\langle \, |r(t) - r(0)|^2 \, \right\rangle \tag{2.58}$$

where $r(t)$ denotes the particle's position at time t. Diffusivity is obtained from the slope of the MSD curve in its linear regime, typically evaluated via linear regression over long-time data.

Helmholtz free energy

Building upon the concept of interfacial free energy γ (Section 2.1.5), it is possible to also define the variation of Helmholtz free energy (ΔF) between to states of the system: with and without interface. As such, Helmholtz free energy variation can be written as:

$$\Delta F = A\Delta\gamma \tag{2.59}$$

Alternatively, incorporating the internal energy U, the configurational entropy at constant volume S and the temperature T:

$$A\gamma = \Delta F = \Delta U - T\Delta S \tag{2.60}$$

With the goal of isolating the interfacial free energy term, as it is difficult to evaluate, the formula is weighed on the surface area:

$$\gamma = \frac{\Delta F}{A} = \frac{\Delta U}{A} - T\frac{\Delta S}{A} \tag{2.61}$$

Resulting in:

$$\Delta\gamma = \Delta f = \Delta u - T\Delta s \tag{2.62}$$

Where the internal energy term is expressed as:

$$\Delta U = U_{\text{interface}} - U_{\text{bulk}} \tag{2.63}$$

In which $U_{\text{interface}}$ is the internal energy at the solvent-solute interface in the combined system, and U_{bulk} is the internal energy of the solvent and solute in isolated systems.

This formulation allows for the estimation of the interfacial free energy γ based solely on computable parameters, such as internal energy and entropy, thus providing a valuable thermodynamic descriptor for the characterization of the solute-solvent interface.
In the current work, only the first term of the calculation (Equation 2.63) was estimated, with the goal of correlating the experimental results with computational simulations of interfacial energy. In particular, the hypothesis states that, for the observed growth to be

consistent with the simulations, the calculated interfacial energy must be lower than the sum of the surface energies of the crystal and the solvent considered individually.

Chapter 3

Materials and Methods

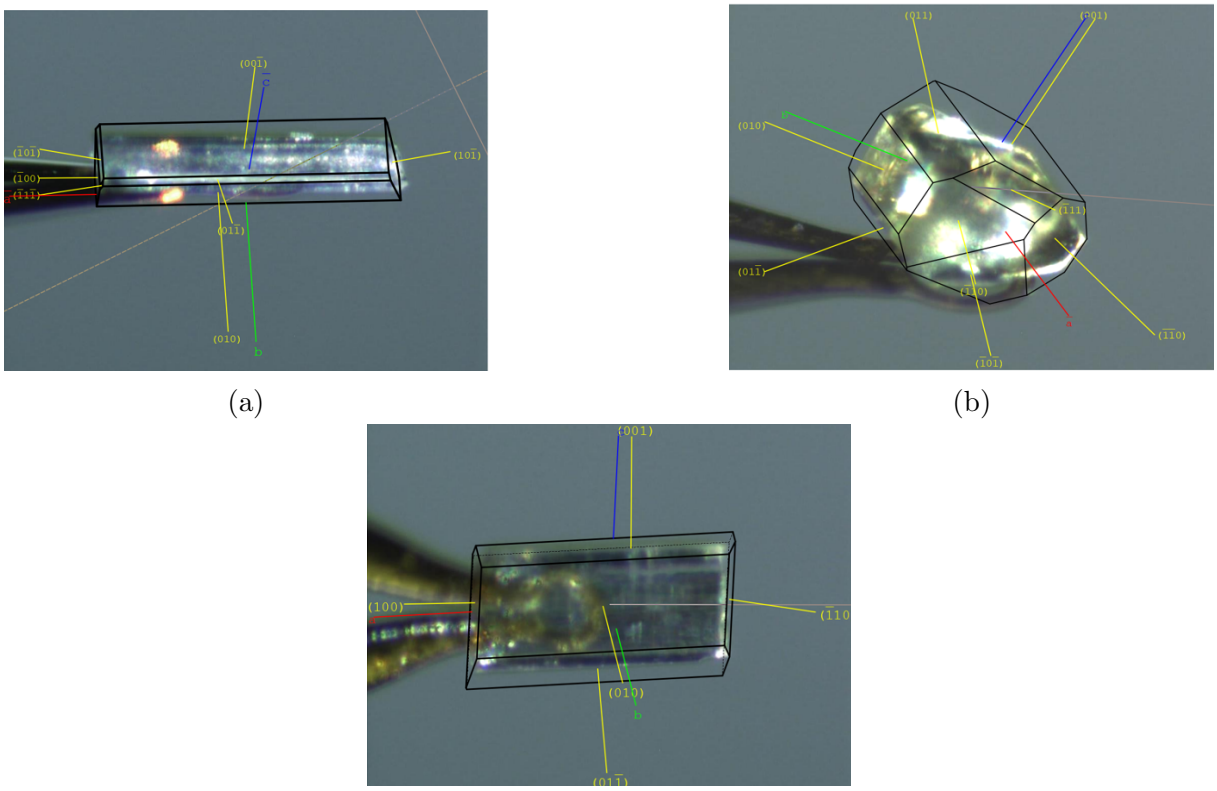
Before presenting the material and methods for each part of the work, it is important to clarify which xanthone crystal faces were considered in both the experimental and computational analyses, as these constitute the core of the interaction study between the crystal and the solvents.

Table 3.1 lists the analyzed faces: experimental nomenclature is reported using Miller indices, while the corresponding computational equivalents are shown alongside. Face indexing of xanthone crystals were obtained by collecting the diffracted intensities on single-crystals with a CMOS Photon II 2D detector on a Bruker D8 Venture diffractometer equipped with a kappa goniometer and an Oxford Cryostream [1].
To better appreciate the morphology of the different faces of xanthone in the three solvents, Figure 3.1 is provided.

Table 3.1: Nomenclature of xanthone faces used in experimental and computational analysis

Crystalline Face	**Computational Name**
(100)	BC
(010)	AC
(001)	AB

The selected faces correspond to the three orthogonal crystallographic directions, providing a representative basis for characterizing crystal–solvent interactions. This choice ensures coverage of the main growth orientations and offers a suitable starting point for the type of analysis performed.



(a) (b)

(c)

Figure 3.1: Morphology characterization of xanthone Crystals in acetone (a), acetonitrile (b), and toluene (c), from [1]

3.1 Materials and Methods for Experimental Analysis

Multiple experiments were carried out aimed at collecting images and data related to the crystal growth of xanthone in three different solvents; to ensure comparability among the results obtained across the three systems, a standardized experimental protocol was adopted, using identical equipment for data and image acquisition.

3.1.1 Materials

Xanthone (IUPAC name: 9H-xanthen-9-one) appears as a light brown powder and was supplied by Alfa Aesar(Ward Hill, Massachusetts, USA) with a reported purity of 98%. Acetone, acetonitrile, and toluene (Riedel-de-Haën, Honeywell Specialty Chemicals, Seelze, Germany) were employed as solvents for crystallization, with a reported purity of 99.9%.

3.1.2 Instrumentation

The materials used both during the experiments and during measurements were as follows:

- Jacketed Reactor: Reactor-Ready Radleys, 500 mL capacity;
- Overhead Stirrer: Velp Scientifica DHL;
- Thermocouple: Radleys Pt 100 PTFE 8 mm with Lemo connector;
- Spectrophotometer: Optek 4000;
- Temperature Controller: Huber Pilot One;
- Polarized Light Optical Microscope: Zeiss Axiolab 5 (Germany);
- Analytical Balance: Radwag (accuracy ± 0.1 mg);
- Filtration Assembly: KNF vacuum pump SC 920 F (Germany), 100 mL Büchner flask, filter, and Cordenos Perfecte 2 filter paper;
- Beaker: 1 L capacity;
- Hot Plate: IKA C-mag Hotplate;
- Mortar and Pestle;

Images of crystals were captured using a mobile device and subsequently analyzed with *ImageJ* software.The imaging was done with a Xiaomi Redmi Note 11 5G rear camera: 108 MP wide-angle lens (f/1.9) and 8 MP ultra-wide lens (f/2.2).

3.1.3 Crystallization Experiments

Seeds Production

The seeds were produced via anti-solvent technique from a saturated solution at 24°C in acetone, following the protocol below:

1. Preparation of a saturated xanthone-acetone solution in a 1 L beaker
2. Heating the solution on a hot plate to 40°C with magnetic stirring (agitation: 240 rpm)
3. Temperature maintenance at 40°C for 15 minutes
4. Removal of the beaker from the hot plate
5. Anti-solvent (distilled water) addition in a 1:1 ratio
6. Sample filtering

Steps 5 and 6 were repeated multiple times until the permeate appeared sufficiently clear and the retentate presented a minimal mass of crystals.

The resulting seeds were then ground with a mortar and pestle to reduce their size, thereby enabling, in solvents that allow it, growth along the three main crystallographic directions. In fact, excessively large needle-like crystals, such as those typically obtained through this method, tend to grow preferentially along one face, limiting the ability to observe crystal development in both directions.
The seeds thus obtained (Figure 3.2) were photographed and analyzed under optical

microscopy to allow characterization and identification of a dimensional baseline, that is, the initial average length and width, to effectively evaluate the growth of xanthone in the various solvents.

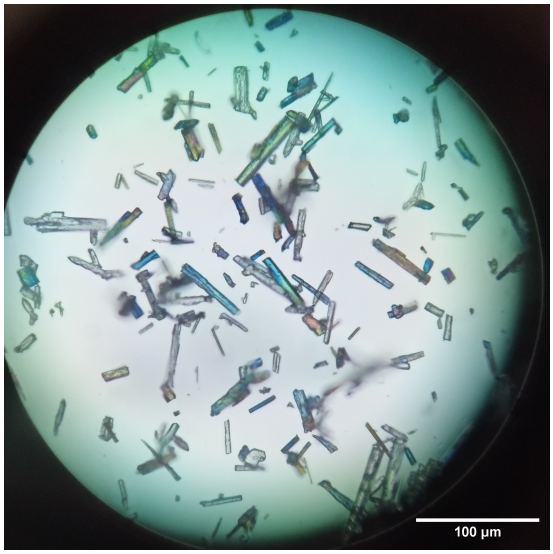


Figure 3.2: Seeds produced with anti-solvent method (objective 40x)

The mass quantities of the reagents are therefore summarized in Table 3.2.

Table 3.2: Reagent quantities used in seeding

Compound	**Quantity (g)**
Xanthone	5.2
Acetone	300
Distilled Water	300

Solubility Curves

The solubility curves of xanthone in the different solvents, necessary to calculate the equilibrium concentration and the corresponding excess solute mass required to obtain a supersaturated solution, were obtained using Crystal 16, an automated parallel crystallizer capable of handling up to 16 different crystallization processes. It is primarily used for solubility screening, crystallization studies, and small-scale process optimization.
The device is equipped with a turbidity sensor that measures the transparency of the solution, from which the software identifies dissolution points and extrapolates solubility curves as a function of temperature and concentration. These solubility curves provided by Crystal 16 constitute the source of solubility data employed in the experiments.

The solubility curves were thus obtained by simultaneously monitoring 10 stirred reactors (stirring rate: 780 rpm) with increasing concentrations, each subjected to a linear temperature ramp from 10 to 30°C at a rate of 0.3°C/min.

The various solutions were placed in $1mL$ vials equipped with small stir bars at the bottom, which served both to maintain agitation and solution homogeneity, and to prevent crystal residues from adhering to the vial walls.

For the purposes of this analysis, the solubility curve of xanthone in acetone was newly determined using the methodology described above. The resulting graph (Figure 3.3) and a summary table of the dissolution concentrations and temperatures for the different vials (Table 3.3) are reported below.

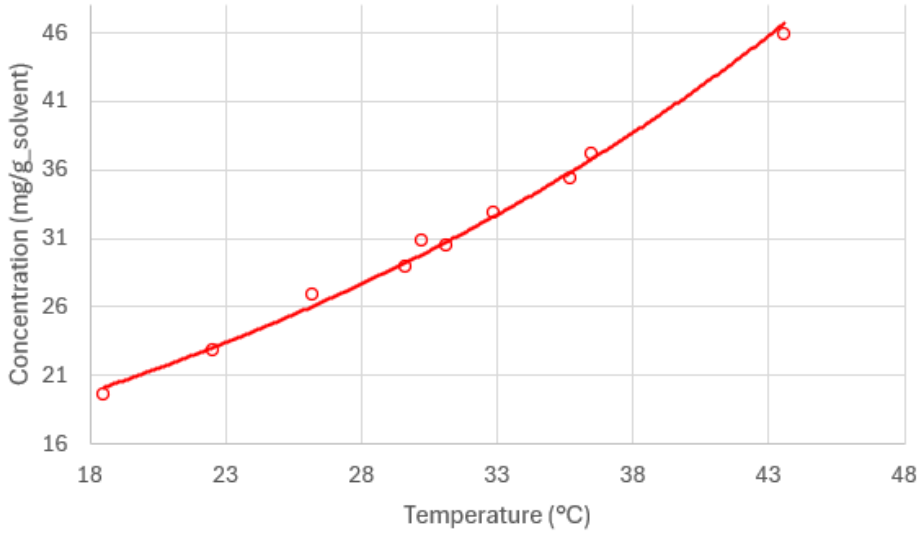


Figure 3.3: Solubility curve of xanthone in acetone

Table 3.3: Values of concentration of xanthone in acetone as a function of Temperature in solubility analysis

Vial	**Concentration (mg/g)**	**Temperature (°C)**
1	19.62	18.5
2	22.78	22.5
3	26.90	26.2
4	28.94	29.6
5	30.51	31.1
6	30.86	30.2
7	32.89	32.9
8	35.35	35.7
9	37.18	36.5
10	45.84	43.6

The curve related to acetonitrile (Figure 3.4) was instead expanded and refined based

on pre-existing data, in order to improve its resolution within the temperature range of interest (Table 3.4).

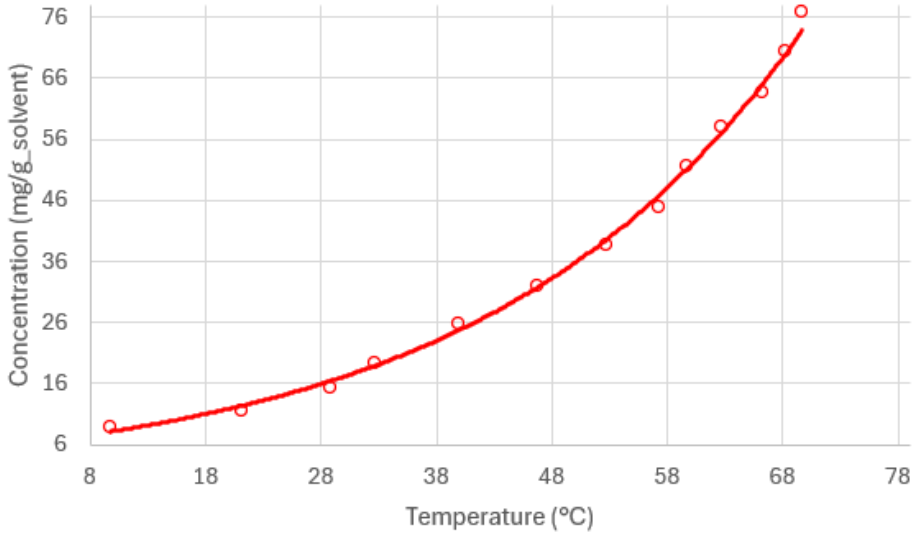


Figure 3.4: Solubility curve of xanthone in acetonitrile

Table 3.4: Values of concentration of xanthone in acetonitrile as a function of Temperature in solubility analysis

Vial	**Concentration (mg/g)**	**Temperature (°C)**
1	8.80	9.8
2	11.60	21.2
3	15.40	28.8
4	19.30	32.7
5	25.78	39.9
6	31.94	46.7
7	38.57	52.7
8	44.84	57.3
9	51.51	59.7
10	58.17	62.7
11	63.80	66.2
12	70.30	68.2
13	76.76	69.7

For toluene, previously obtained data were used, collected through an experimental protocol identical to that adopted in the present study. These data cover a broader temperature range and exhibit a bigger number of samples (23 in total) compared to those acquired for acetone and acetonitrile (10 and 13 samples, respectively).

Despite this difference in the number of measurements, the comparative analysis remained valid, as the underlying methodological consistency was preserved and the resulting information proved compatible with the objectives of the study.

The solubility curve (Figure 3.5), accompanied by a table reporting the dissolution concentrations and temperatures (Table 3.5), was therefore used without further modification, ensuring methodological consistency with the other solvents analyzed.

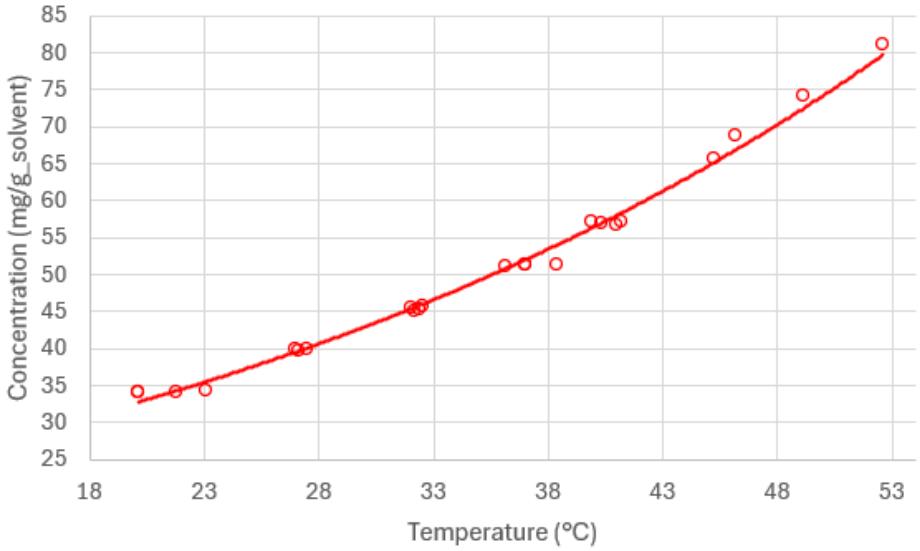


Figure 3.5: Solubility curve of xanthone in toluene

Table 3.5: Values of concentration of xanthone in toluene as a function of Temperature in solubility analysis

Vial	**Concentration (mg/g)**	**Temperature (°C)**
1	34.0	20.1
2	34.0	21.8
3	34.2	20.1
4	34.3	23.1
5	39.7	27.1
6	39.9	27.0
7	40.0	27.5
8	45.1	32.2
9	45.4	32.4
10	45.5	32.0
11	45.8	32.5
12	51.2	36.1
13	51.3	37.0

Vial	Concentration (mg/g)	Temperature (°C)
14	51.3	38.4
15	51.4	37.0
16	56.8	41.0
17	57.0	40.3
18	57.1	41.2
19	57.2	39.9
20	65.8	45.2
21	68.8	46.2
22	74.2	49.1
23	81.1	52.6

Experimental Protocol

The crystallization technique employed relies on secondary nucleation induced by seeding in a supersaturated solution of xanthone and solvent at $24°C$.
Following the solubility analysis and the determination of the equilibrium concentration, an initial supersaturation level of $1.09\frac{C}{C_{eq}}$ was arbitrarily selected, in order to standardize the initial conditions across all experiments and allow for a consistent and reliable comparison of the resulting growth kinetics. The seeding quantity was set at 5% of the xanthone mass corresponding to the equilibrium concentration in the solvent at $24°C$. The mass values of crystal and solvent used in the experiments are summarized in Table 3.6. After the introduction of the seed crystals, the system was subjected to controlled cooling with a thermal gradient of -0.05$°C$/min, from $24°C$ down to $10°C$. The final temperature of $10°C$ was then maintained constant for the next 14 hours.

The procedure adopted was as follows:

1. Introduction of the xanthone–solvent solution into the stirred reactor (stirring: 220 rpm)
2. Heating of the solution to $35°C$ to allow complete dissolution of the crystals
3. Maintaining the system at $35°C$ for 30 minutes
4. Cooling to $24°C$
5. Seeding
6. Sampling: 2 mL every hour (8 samples total), with an additional sample after 14 hours to assess the plateau in terms of growth and absorbance
7. Sample filtration (pressure: 400 mbar)
8. Weighing of the permeate before solvent evaporation

9. Weighing of the permeate after solvent evaporation
10. Morphological analysis of the crystals in the retentate using optical microscopy

Steps 8 and 9 are fundamental for determining the equilibrium concentrations of the solution at various temperatures—an essential parameter for the subsequent analysis of growth kinetics.

The forementioned value was obtained through Equation 3.1.

$$C = \frac{m_x}{m_s} \tag{3.1}$$

where m_x is the mass of xanthone in mg, calculated using Equation 3.2, and m_s is the mass of solvent in grams, obtained using Equation 3.3.

$$m_x = m_V - m_{V_{pe}} \tag{3.2}$$

$$m_s = m_P - m_V - m_x \tag{3.3}$$

where m_V and $m_{V_{pe}}$ represent the mass of the empty vessel and the mass of the vessel containing the sample after solvent evaporation, respectively, while m_P corresponds to the mass of the vessel containing the sample immediately after filtration.

3.1.4 Crystal Characterization and Image Analysis

Turbidity Probe

The use of a turbidity sensor enables the measurement of two key properties for assessing crystal growth trends: absorbance and transmittance (see Section 2.1.7).

Analyzing the evolution of these functions over time was essential to draw meaningful conclusions and formulate plausible hypotheses on the system's behavior. Monitoring these variations therefore provided an indirect yet effective way to detect secondary nucleation events and other dynamic phenomena during crystallization.

Optical Microscope

For the execution of step 10 of the experimental protocol (described in Section 3.1.3), a portion of the crystals retained on the filter paper was deposited onto a microscope slide for analysis. Observations were carried out using an optical microscope (see Section 3.1.2), with 40x, 10x, and 5x objectives selected based on the size of the crystals observed. Images (approximately 60 per sample) were captured using the camera of a Xiaomi Redmi

Note 11 5G mobile device (Xiaomi Inc., Beijing, China).

ImageJ

The images obtained were used to measure crystal size and shape distributions using the Fiji software (*ImageJ* 1.54p, NIH, Bethesda, USA)[25]. The analysis was performed with the ROI Analysis tool, which allows for manual tracing of visible crystals in the images (Figure 3.6) and subsequent measurement of their width and length (Figure 3.7). These dimensions were selected as xanthone crystals predominantly exhibit a needle-like morphology, which can be reasonably approximated by rectangular geometry.

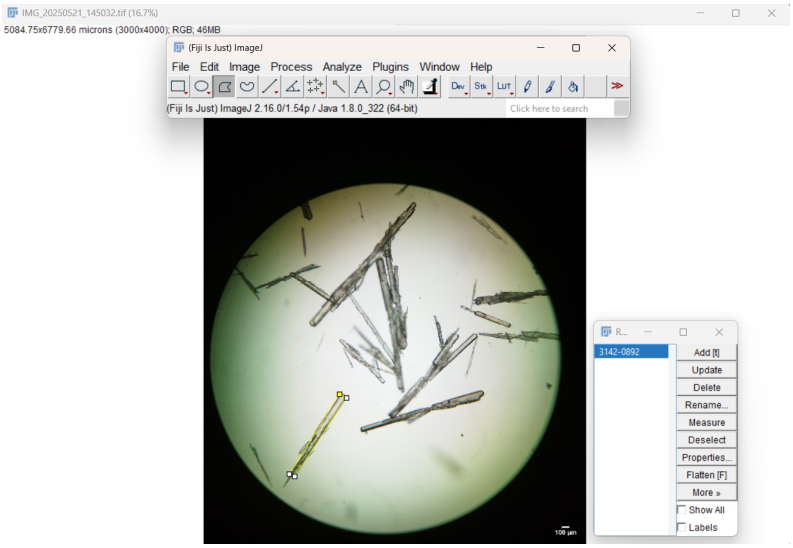


Figure 3.6: Crystal individuation through ROI Set Analysis Tool in *ImageJ*

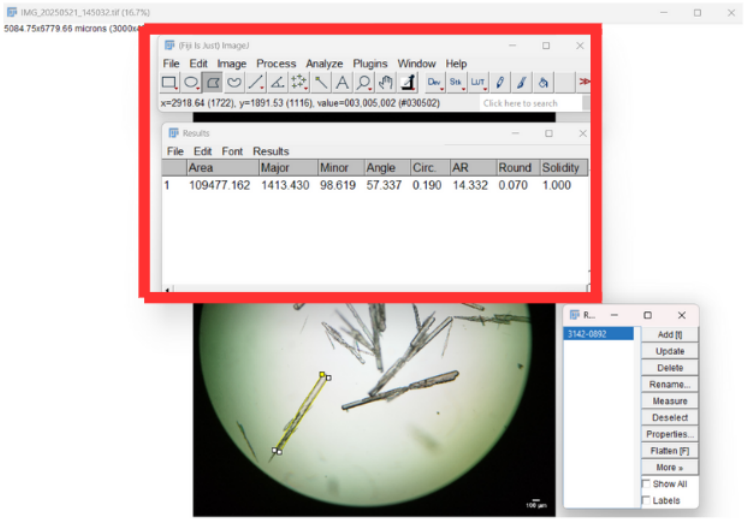


Figure 3.7: Crystal measurements through ROI Set Analysis Tool in *ImageJ*, the red square shows the measures obtained for the selected crystal

The data thus obtained were subsequently transferred to Excel (Microsoft Corp., Redmond, USA) for processing. Specifically, for each sample, the average length, width, and aspect ratio were calculated, as well as the variations in the distribution of these values over time. To monitor the morphological evolution of xanthone crystals during the experiment, the Aspect Ratio (AR) was calculated for each sample using Equation 3.4. This dimensionless parameter provides insight into the overall shape of the particles by quantifying the ratio between their length (L) and width (W):

$$AR = \frac{L}{W} \tag{3.4}$$

Values of AR close to 1 typically indicate equiaxed or near-spherical particles, whereas higher values are associated with elongated or needle-like morphologies. In the context of this study, the AR served as a useful indicator of anisotropic growth, allowing to evaluate whether the crystals developed more uniformly along both axes or displayed a preferential elongation along a specific direction.

3.1.5 Statistical Analysis

As previously discussed (see Section 2.1.7), cross-validation between the mean and standard deviation values obtained via the moments evaluation and those calculated through the cumulative running average and running standard deviation is essential; firstly, to verify the correctness of the estimations derived from the latter method, secondly, to assess whether the number of particles analyzed is statistically representative of the overall population.

Moreover, since experimental observations of the manually selected crystals indicated that xanthone predominantly grows along two crystallographic directions, the statistical analysis was performed using both a one-dimensional (1D) and a two-dimensional (2D) approach: the 1D method provided a first set of indicative results, while the 2D analysis offered a more comprehensive assessment of the population's geometrical features.

1D - Analysis

Following the equations detailed in Section 2.1.6, the evaluation of the moments in one dimension for a predominantly bi-dimensional crystal required the estimation of an equivalent linear dimension.

To this end, the crystal, whose actual three-dimensional geometry resembles a prism, was approximated as a sphere. This allowed for the calculation of an equivalent spherical diameter, which was then used as the characteristic size (L_1) for the one-dimensional formulation. The procedure adopted is summarized as follows:

- Individuation of the two principal growth directions, L_1 and L_2, through image analysis (see Section 3.1.4). The number of crystals analyzed per experimental sample is reported in Table 3.6.

Table 3.6: Number of crystals analyzed for each solvent and sample

Solvent	**Sample**	**Number of crystals analyzed (N)**
Acetone	1	732
	2	984
	3	1014
	4	1113
	5	541
	6	531
	7	812
	8	803
	9	459
Acetonitrile	1	414
	2	624
	3	331
	4	697
	5	360
	6	404
	7	381
	8	265
	9	232
Toluene	1	843
	2	902
	3	741
	4	416
	5	641
	6	513
	7	415
	8	360
	9	538

- Estimation of the actual crystal volume assuming a prism with a square base (L_1) and height L_2, using Equation 3.5:

$$V_{prism} = L_1^2 L_2 \tag{3.5}$$

- Calculation of the equivalent diameter D_e by equating the prism volume to that of a sphere and inverting the standard formula (Equation 3.6):

$$D_e = (\frac{6V_{prism}}{\pi})^{\frac{1}{3}} \tag{3.6}$$

The resulting values were used to compute the zero-through-sixth moments of the population. Specifically, moments 0–3 were used to derive the total number of crystals, total length, surface area, and volume (see Section 2.1.6). Moments 4-6 were also evaluated as parameter employed in the statistical formulations (see Equation 2.34-2.35 in Section 2.1.7) for the estimation of each property's standard deviation and confidence interval. Furthermore, Equations 2.31, 2.32, and 2.33 were used to estimate the average length and standard deviation of the entire crystal population. For the parameter N, representing the total number of sampled particles, the values reported in Table 3.6 were used as reference.

Cumulative Running Average and Cumulative Running Standard Deviation

In line with the 1D moments evaluation protocol, the equivalent diameter D_e was employed as the representative linear dimension for the statistical calculations. This enabled the estimation of both the cumulative (running) average length of the crystal population and the corresponding running standard deviation over time. To streamline the process and reduce computational effort, the calculations were automated via a custom Python script (see Appendix A). The code processed data directly imported from Excel worksheets, where the individual D_e values for each crystal in each sample had been previously recorded.

Confidence Interval and Minimum Number of Crystals

Using the equations described in Section 2.1.7, confidence intervals were calculated for the three main investigated properties (average length, average projected area, and average projected volume) through evaluation of the moments of the sampled distribution. The values used for the $t_{\infty,\alpha/2}$ parameter in Equation 2.36 are as follows[14]:

- for $\alpha = 90\%$: 1.645;
- for $\alpha = 99\%$: 2.576.

From the computed confidence intervals (CI), it was then possible to estimate the minimum number of crystals that would be required to reduce the confidence interval.
This approach serves a dual purpose: on one hand, it validates whether the number of particles analyzed in the samples is sufficient to statistically represent the population with a given confidence level; on the other, it ensures that the measurement uncertainty does

not fall below the instrument's physical resolution limit, since higher statistical precision than what the instrument can detect would be both misleading and meaningless.

To this end, the optical resolution of the three objectives (40x, 10x, and 5x) of the polarized light microscope used for image acquisition was calculated using Abbe's diffraction limit (Equation 3.7):

$$d = \frac{0.61\lambda}{NA} \tag{3.7}$$

where d is the resolution limit, λ is the wavelength of visible light (assumed to be 500 nm), and NA is the numerical aperture of the objective, as indicated on the objective itself. For the microscope used in this study, the NA values were 0.65, 0.25, and 0.15 for the 40x, 10x, and 5x objectives, respectively.

The calculated resolution values for each objective are reported in Table 3.7.

Table 3.7: Optical resolution values for different objectives employed in the crystals' image analysis

Objective	**Optical Resolution** (μm)
40x	0.52
10x	1.34
5x	2.24

2D - Analysis

Due to the complexity involved in modeling the correlation between the two main growth dimensions, the 2D statistical analysis was carried out by treating each dimension independently. In other words, rather than considering the crystal area as the product of the two correlated lengths (e.g., $L_1 \times L_2$), the estimators were calculated using simplified forms such as L_1^2 or L_2^2, depending on the dominant dimension being analyzed. This allowed for a more manageable evaluation of the crystal size distribution while still providing insight into the geometrical characteristics of the population.

To this end, similarly to the 1D analysis, firstly the zero-through-sixth moments were estimated independently for each of the two principal growth dimensions. Secondly, these parameter where then employed to assess the confidence interval, as described in the previous paragraph.

3.2 Material and Methods for Computational Analysis

To enable a comprehensive analysis of the crystal growth system, it was necessary to evaluate the internal energy of both the xanthone and the solvent boxes.
This was achieved through a series of molecular dynamics simulations, carried out in two stages: first, on separate systems, each containing only one component, and subsequently on combined systems, consisting of both the crystal and the solvent under investigation.

3.2.1 Softwares

LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator; Sandia National Laboratories, New Mexico, USA) is an open-source software for molecular dynamics (MD) simulations.
In this work, LAMMPS (ver. 2 August 2023) was used to perform classical atomistic simulations aimed at extracting properties useful for evaluating the compatibility between the computational model and experimental reality.
The simulations were carried out on a molecular system (hereafter referred to as box), in which the molecules are confined within a three-dimensional periodic geometry reproducing the non-orthogonal triclinic structure characteristic of xanthone. The objective is to obtain the internal energy of the system, then used as a parameter for analyzing interfacial energy.

To perform a simulation in LAMMPS, as previously described (see Section 2.2.2), it is necessary to define a force field: in the present case, the TraPPE model (Transferable Potentials for Phase Equilibria)[18] was employed, a force field parameterized to accurately reproduce the thermodynamic properties of organic fluids, particularly in the context of phase equilibrium. This model was chosen in order to ensure an optimal balance between accuracy and computational cost.

Another key aspect in configuring molecular dynamics simulations concerns boundary conditions. In all systems analyzed, Periodic Boundary Conditions (PBCs) were applied, whereby the simulated box is treated as part of an infinite system, periodically replicated along all three spatial directions. This setup solves the issue of edge effects, which would otherwise introduce significant artifacts in the calculations, especially in liquid systems. PBCs allow particles to exit one side of the box and re-enter from the opposite side, thus simulating a homogeneous and continuous environment more representative of bulk conditions.

Each simulation requires a preliminary energy minimization step to obtain a stable initial

configuration corresponding to a local minimum of the system's potential energy. Although LAMMPS includes a built-in minimization function (*minimize*), this command was deemed adequate only in specific cases (see Section 3.2.2). Various tests revealed that, for the systems considered in this study (see Section 3.2.2), the integrated approach was overly approximate and failed to ensure the necessary structural stability.

In such cases, an alternative strategy was adopted: a selective minimization performed via a custom script, in which the equations of motion were applied exclusively to the fluid portion of the system (i.e., the solvent). This approach allowed for optimizing the solvent configuration without altering the crystal structure. The resulting configuration was then used as the starting point for simulating the full system, including the dynamic evolution of the crystalline component.

LAMMPS simulations are based on an iterative evaluation of system parameters at regular and predefined time intervals, known in computational terms as timesteps. The timestep value was selected according to the specific objective of each simulation:

- For simulations aimed at minimizing the system's energy (Section 3.2.2), a very small timestep of 10^{-8} fs (with 1 fs = 10^{-15} s) was used in order to ensure numerical stability and prevent physically unrealistic behaviors such as energy divergence or structural collapse of the system;
- For simulations intended to evaluate the global physical properties of the system (Section 3.2.2), a timestep of 0.5 fs was applied, which is more appropriate for data collection purposes.

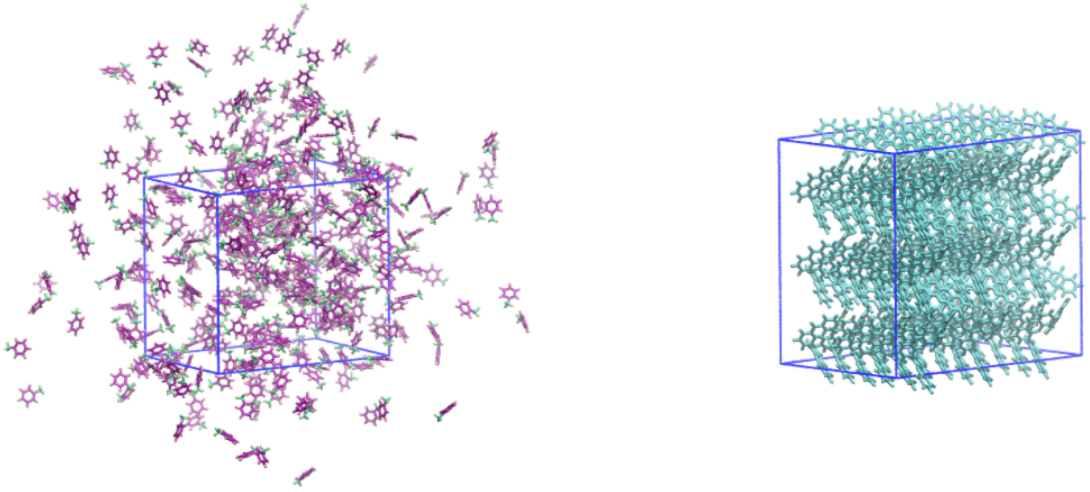
The total simulation time (running time) was also adapted to the type of analysis being conducted. Total durations ranged from 20,000 steps (for energy minimization simulations) to 500,000 steps (for simulations focused on thermodynamic or structural property analysis).

The values of the physical properties were obtained as averages of the quantities calculated at each step, since in the simulated context, LAMMPS does not provide a continuous dynamic evolution of the molecular structure, but rather computes the system's state at each discrete time point. This approach allows simulations to be treated as a series of repeated measurements on the same system, offering a statistical assessment of the target properties analogous to conducting multiple experimental trials on the same sample.

VMD (Visual Molecular Dynamics, Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign, USA) is an open-source software used for the visualization and analysis of molecular structures and trajectories generated from molecular dynamics simulations.
The program allows graphical representation of molecular systems using various styles

(wireframe, licorice, CPK, etc.) and integrates tools for identifying structural issues that may compromise the reliability or stability of simulations, such as molecular overlaps, interpenetration of distinct species, or unstable configurations.

In this work, VMD was primarily used for visualizing the simulation boxes and verifying the correct spatial configuration of crystal–solvent systems, with particular attention given to simulations involving selective minimization.
The boxes used had a non-orthogonal triclinic geometry, which made it less intuitive to interpret compared to a standard rectangular box. For this reason, during the initial stages of computational setup, graphical visualization was essential for identifying issues such as molecular escape or improper containment within the box; conditions that, in some cases, led to simulation errors or run failures.
An example of the boxes is given in Figure 3.8.



(a) (b)

Figure 3.8: Visualization example in VMD of toluene- (a) and xanthone- (b) only boxes

Python (Python Software Foundation, USA) is an open-source, high-level programming language widely used in scientific computing, data analysis, and automation of simulation workflows. Thanks to its extensive ecosystem of libraries, such as NumPy, pandas, Matplotlib, and MDAnalysis, Python enables efficient handling of large datasets, statistical evaluation, and graphical representation of results obtained from molecular dynamics simulations.

In the present work, Python was employed primarily for post-processing simulation outputs from LAMMPS, including the extraction of relevant physical properties (density, radial distribution functions, diffusivity) and the generation of plots for result visualization. Custom scripts were developed to automate repetitive analysis tasks, ensuring

reproducibility and minimizing human error in data handling. Particular attention was devoted to the Radial Distribution Function (RDF) and Mean Square Displacement (MSD) analyses, where Python scripts were used to compute ensemble-averaged values, estimate error bars, and compare trends with literature data.

Python played a key role also in the post-processing of experimental data, particularly for optimization routines and parameter estimation within the framework of the calculations of moments (see Section 2.1.6), as well as for the calculation of cumulative running averages and cumulative running standard deviations, and for the comparison of results across different datasets (see Section 2.1.7).

The specific scripts developed for each stage of the analysis are provided in Appendix A, to ensure reproducibility and to offer a transparent overview of the implemented procedures. In most cases, input and output data were managed through Excel worksheets, facilitating data handling and visualization.

3.2.2 Computational Protocol and simulation assessment

Xanthone and solvent boxes validation trough physical properties assessment

Before proceeding with the evaluation of the internal energy of the systems, it was necessary to verify the adequacy and consistency of the provided molecular structures (boxes), both with respect to other models available in the literature and to the physical constraints of the system.
To this end, independent simulations were performed on both crystalline and liquid boxes, with the aim of separately assessing the structural and thermodynamic properties of each phase and comparing them with known experimental and computational data from the literature.
The properties analyzed, selected for their physical relevance, are, in order: density, radial distribution function (RDF), and diffusivity.

Solvents' density was the first physical property to be validated, as it constitutes a fundamental macroscopic characteristic of the system. Its evaluation provides an initial check on the consistency between the simulated configurations and physical reality, serving as a preliminary validation of the chosen force field and simulation setup.

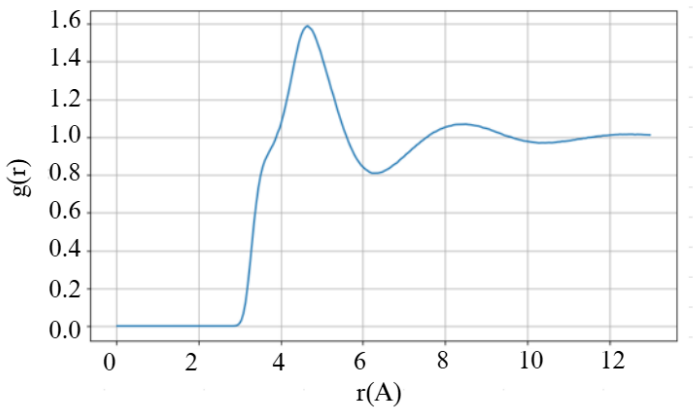
Simulations were conducted for 500,000 iterations (equivalent to approximately 2.5×10^{-10} s), using an NPzAT ensemble, a variant of the classic NPT ensemble, at 293.15 K and 1 atm. In LAMMPS, the standard NPT ensemble adjusts the dimensions of the simulation box along all three directions (x, y, and z) to maintain constant pressure. However, in this study, such behavior was deemed undesirable: evaluating the interfacial

energy between the solvent and the crystal requires the box to retain constant dimensions along the x and y axes, which correspond to the directions in direct contact with the crystal surface.

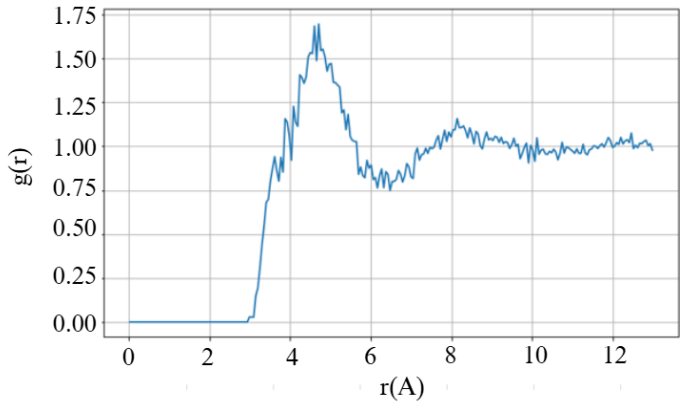
The adoption of the NPzAT ensemble, which allows dimensional variation only along the z-axis (perpendicular to the interface), enabled the pressure effect to be isolated without altering the contact area with the crystal surface. This approach provided a more realistic and reliable assessment of the simulated systems' density, which was obtained through the evaluation of the mean value throughout the simulation running time.

The analysis of Radial Distribution Function (RDF) profiles was carried out to verify the correctness of the interatomic distances in the simulation boxes. Agreement between the simulated RDF trends and reference data from the literature was taken as evidence that the constructed boxes reproduced the real molecular structure with high accuracy.

The simulations were performed using the NVT ensemble, with a total runtime of 400,000 iterations (equivalent to 200 ps, with a timestep of 0.5 fs). The use of this ensemble was made possible only after preliminary validation of the systems' density. The RDF output produced by LAMMPS was computed over windows of 250 iterations, generating a total of 1,600 RDF profiles for each analyzed atom–atom pair. To reduce noise and obtain representative curves, the average of the generated RDF profiles was calculated. The difference between the raw and averaged graphs is illustrated in Figure 3.9.



(a)



(b)

Figure 3.9: Averaged (a) and not-averaged (b) RDF for acetonitrile, atom-pair: C1-C1

Diffusivity was the final property evaluated to further assess the reliability of the simulation boxes. Of the two methods described in Section 2.2.3, the Velocity Autocorrelation Function (VACF) approach was deemed unsuitable for quantitative analysis in this work, mainly due to sample asymmetry and the intrinsic noise of the simulated data. Instead, the Mean Square Displacement (MSD) method, shown to be more robust for the present liquid systems, is adopted.

Although LAMMPS automatically provides an estimated value of D, for improved reliability the raw data were exported, plotted, and the slope manually calculated using spreadsheet tools.

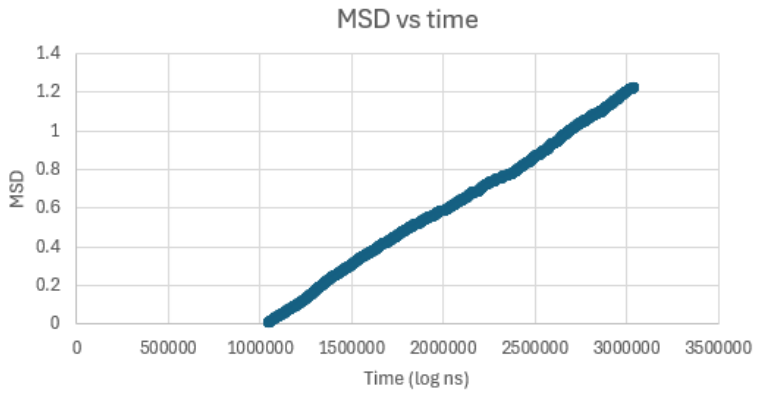


Figure 3.10: MSD vs. time plot for acetone (ac face) – Example of data interpolation

As with the RDF analysis, the simulation was carried out in the NVT ensemble at room temperature and atmospheric pressure, with a total runtime of 400,000 iterations.

Internal Energy calculation

Following the validation of the physical consistency of the simulated systems, simulations were carried out to calculate the internal energy of the individual components; namely, simulation boxes containing only crystalline xanthone and those containing the respective pure solvents (acetone, acetonitrile, toluene).
These simulations represent the first step required to determine the U factor, a key parameter in the calculation of the interfacial energy between crystal and solvent.

Simulations were performed in the NVT ensemble at a constant temperature of 293.15 K and a pressure of 1 atm, with a total runtime of 100,000 iterations (equivalent to 50 ps), for a total of 10 simulations:

- 3 simulations for each pure solvent (one for each crystal face analyzed);
- 1 simulation for the crystalline box.

The obtained internal energy was averaged over a time interval, excluding initial transients, to ensure the reliability of the value.

The next step was to construct the combined xanthone–solvent boxes, required to evaluate the second term in the overall U calculation (see Section 2.2.3). To this end, two main steps were necessary:

- The physical assembly of the simulation box ;
- Energy minimization simulations to identify the most stable configuration.

The initial assembly was carried out using a Bash script, which merged the crystal and solvent boxes by keeping the xanthone structure fixed and translating the coordinates of

the solvent box to position it adjacent to the selected crystal face.
The script allows the user to adjust two key parameters: the distance between the two systems (defined in the script as *extrashift*) and the lateral void space within the combined simulation box (in the script as *slsshift*).
These values had to be carefully tuned: if set too high, the components would fail to interact meaningfully, preventing LAMMPS from computing a reliable interaction energy. Moreover, excessive box dimensions would yield internal energy values that are not comparable with those obtained for the separate xanthone and solvent boxes, due to the mismatch in system volume. Conversely, values that were too small led to overlapping between the crystal and the solvent, resulting in abnormally high Van der Waals and Lennard-Jones interaction energies. This not only caused the simulation to fail but also prevented the generation of physically meaningful or reliable results for parameter estimation.

Several simulations were performed to explore different parameter combinations. Ultimately, a configuration using a value of 1.5 for both extrashift and slsshift was found to be the optimal compromise, ensuring both simulation stability and physical consistency of the calculated energies. An example of the final merged box is shown in Figure 3.11.

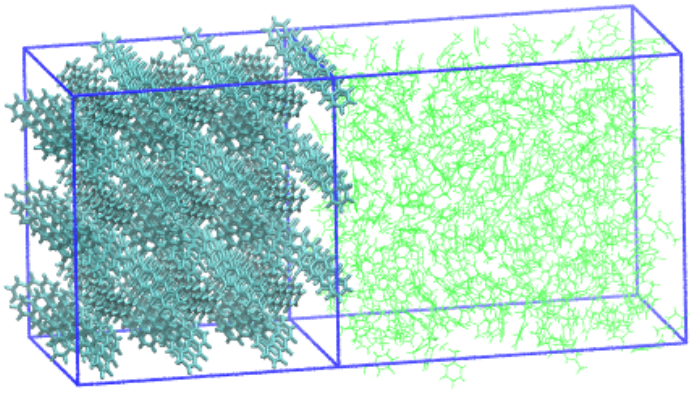


Figure 3.11: Visualization example in VMD of xanthone (left) - toluene (right) box, face ac

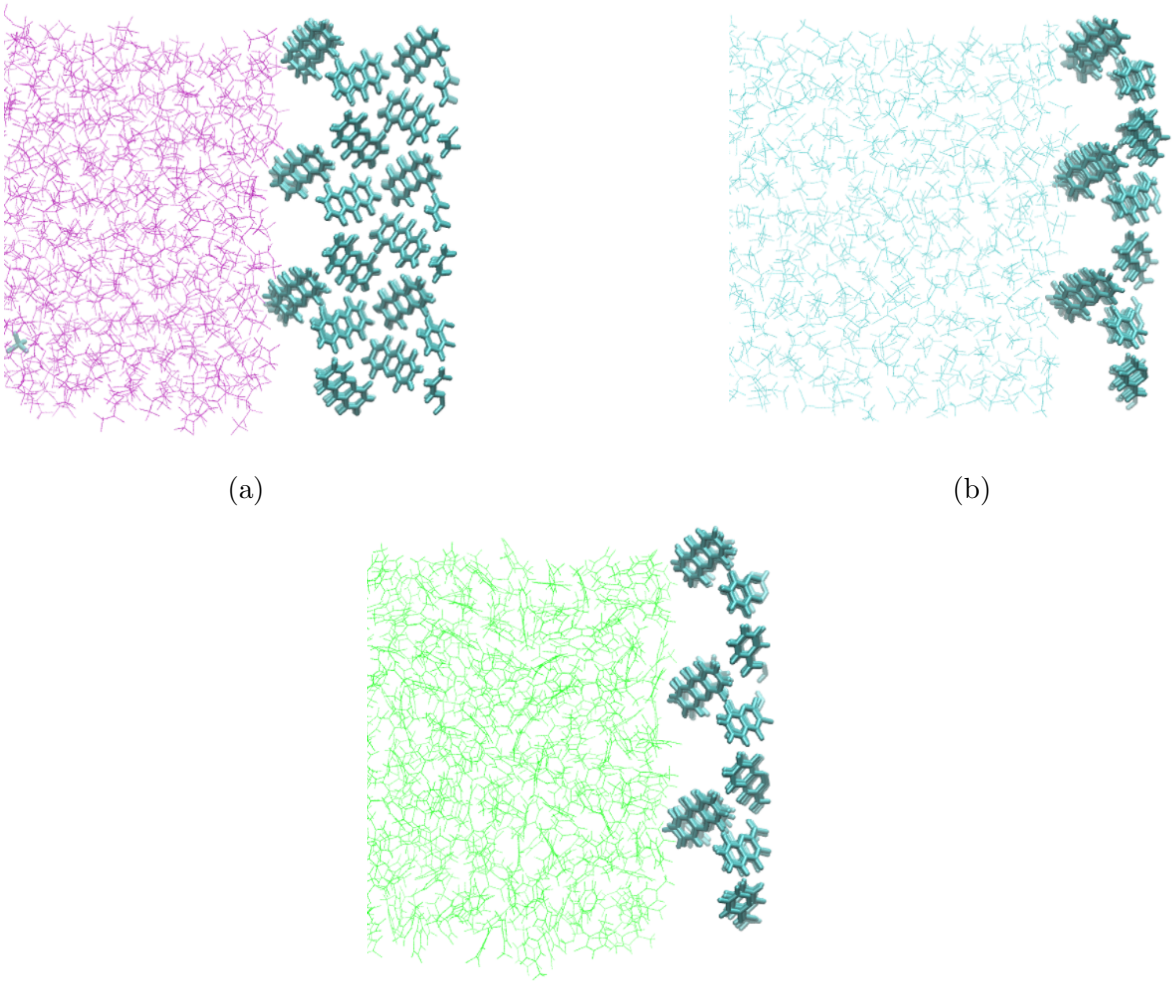
After the assembly of the combined xanthone–solvent simulation box, it was necessary to verify its structural and energetic stability, in order to ensure the reliability of subsequent

energy calculations.
The main critical aspect was the configuration of the solvent, which underwent significant geometric distortion as a result of coordinate translation during the box-merging process.

To address this, a series of computational protocols were tested with the objective of identifying a robust and reproducible procedure. The simulations, executed in chronological order, were as follows:

- Protocol A: Direct simulation of the merged system, without any pre-processing or minimization. The system was unstable, leading to early simulation failure.
- Protocol B: Application of LAMMPS' built-in energy minimization command (minimize) on the merged system. Instabilities persisted, and results were physically unreliable.
- Protocol C: Selective pre-minimization of the solvent component only (see Paragraph 3.2.1), maintaining the crystal fixed. This approach yielded a significantly improved configuration with lower internal energy and increased structural stability.
- Protocol D: Standard simulation using the configuration resulting from Protocol C, without additional minimization. The system remained stable and produced reliable energy data.

The combination of Protocol C and D was therefore adopted for all subsequent simulations involving the united box systems, ensuring both physical plausibility and computational consistency. An example of the three different systems at the interface is shown in Figure 3.12.



(a) (b)

(c)

Figure 3.12: Interface visualization in VMD of the crystal (right)-solvent (left) system for acetone (a), acetonitrile (b) and toluene (c) - face ac

After structural optimization of the combined crystal–solvent boxes (see Section 3.2.2), simulations were carried out to calculate the total internal energy of the interfaced system. These simulations were also performed in the NVT ensemble at 293.15 K and 1 atm, with a total runtime of 100,000 iterations (50 ps).

The value obtained represents the second term required for the evaluation of the U factor, whose complete definition and formulation is provided in Section 2.2.3.

Chapter 4

Result and Discussion

4.1 Experimental Results

4.1.1 Crystal growth analysis

The collected data enabled a detailed analysis of xanthone crystal growth, revealing clear differences in growth rates and morphological evolution across the three solvents. These results provide valuable guidance on which parameters should be considered when developing kinetic models of crystal growth. Moreover, higher growth rates correspond to shorter reaction times in batch processes, thereby increasing production efficiency. It is also important to emphasize that, since the solvent strongly influences crystal morphology, the findings support relevant qualitative considerations regarding the size and shape distribution of the final product, improving its suitability for different purposes and industrial applications.

In particular, the analysis of transmittance and absorbance values gives meaningful insight on the system's behavior. In the present experiments, the increase in absorbance is associated with the progressive light scattering caused by the formation and growth of suspended crystals: a sudden increase in absorbance may indicate a rapid rise in solid crystal concentration. This, in turn, can be associated with a common phenomenon in supersaturated agitated systems: secondary nucleation (see Section 2.1.3). Conversely, an abrupt decrease in transmittance can lead to the same interpretation, given the inverse relationship between these two parameters.

Furthermore, the direct link between absorbance and concentration can be employed to explain different asymptotic values reached by the experiments: in high-solubility solvents, the total amount of solute available for crystallization is greater, allowing for the formation of a larger crystal population and/or larger individual crystals. This results in more pronounced turbidity of the suspension and, consequently, higher absorbance val-

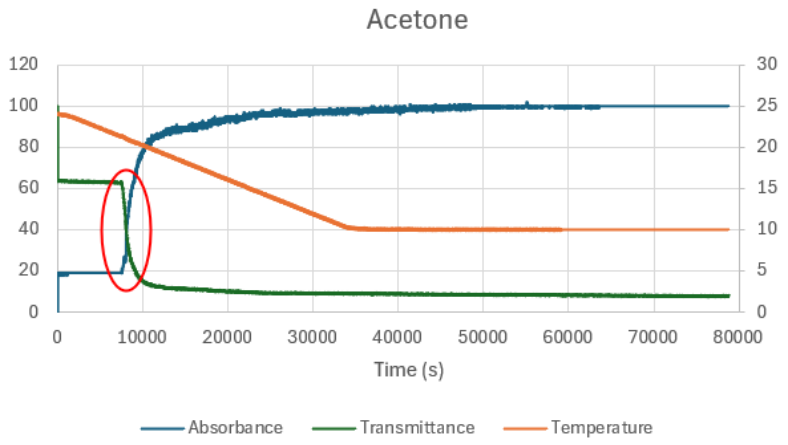
ues. Conversely, in low-solubility solvents, the limited availability of solute restricts the amount of suspended solid, leading to lower maximum absorbance values. This behavior can also be interpreted in light of the supersaturation theory: in highly soluble systems, the initial supersaturation level reached after cooling (or solvent evaporation) is higher, leading to a larger driving force for nucleation and growth. This results in the formation of a greater number of crystals and/or larger crystals, which remain suspended in the medium and increase light scattering. Conversely, in less soluble systems, the lower supersaturation limits both nucleation and growth, yielding fewer suspended particles and thus lower measured absorbance.

These trends are further validated from the results obtained from the solubility curves (see Section 2.1.3), which show clear differences in the concentration at the equilibrium, and its correlation with the solubility itself. As summarized in Table 4.1, toluene is the solvent in which xanthone is most soluble, followed by acetone and acetonitrile.

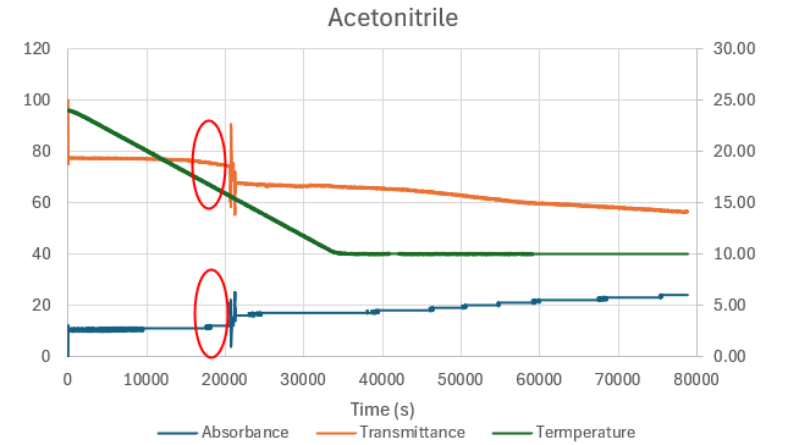
Table 4.1: Xanthone solubility in the different solvents at $24°C$

Solvent	**Xanthone Solubility at 24°C (mg/g)**
Acetone	25
Acetonitrile	14
Toluene	32

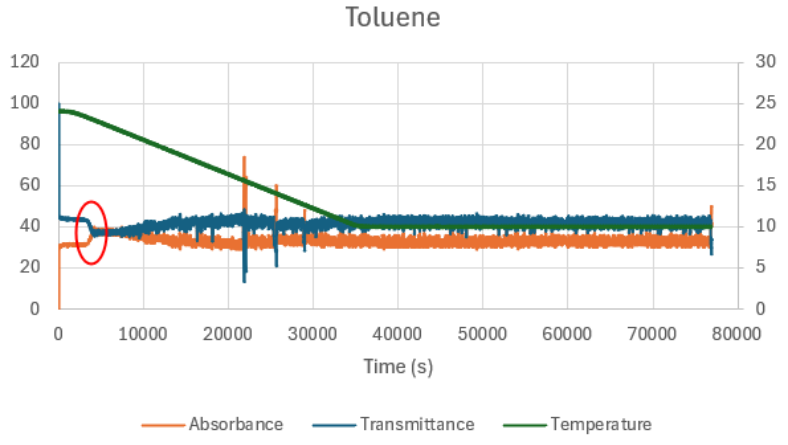
As such, the values of transmittance and absorbance measured during the experiments clearly indicate the occurrence of a second instance of secondary nucleation, as evidenced by the abrupt decrease in transmittance (and corresponding increase in absorbance) shown in Figure 4.1.



(a)



(b)



(c)

Figure 4.1: Absorbance, Transmittance and Temperature Trends in acetone (a), acetonitrile (b) and toluene (c), the red oval indicates where secondary nucleation occurred

During the experiment conducted with acetonitrile, a minor misalignment occurred due to incorrect assembly of the reactor's rotor, which required subsequent adjustment of both the thermocouple and the optical sensor. This procedural correction introduced a

transient irregularity in the signal acquisition process, resulting in the non-linear profile observed just after the red ovals in Figure 4.1 (b). In the toluene experiment, the peaks observed in Figure 4.1 (c) are attributed to a temporary increase in the rotor's speed. This adjustment was made to evaluate whether higher stirring rates could facilitate the suspension of larger crystals, which, due to their weight, tended to accumulate at the bottom of the reactor, preventing representative sampling.
The approach proved effective: increasing the rpm briefly allowed the detachment and subsequent collection of heavier crystals. Since these constituted the main crystal population, their inclusion in the sampling process was essential for reliable dimensional analysis.

To better appreciate the different shape development of the xanthone in the different solvents, particularly regarding the Aspect ratio, a comparison between this value's trend is shown in Figure 4.2. This allows to better understand the difference among solvents and the overall behavior of the crystal.

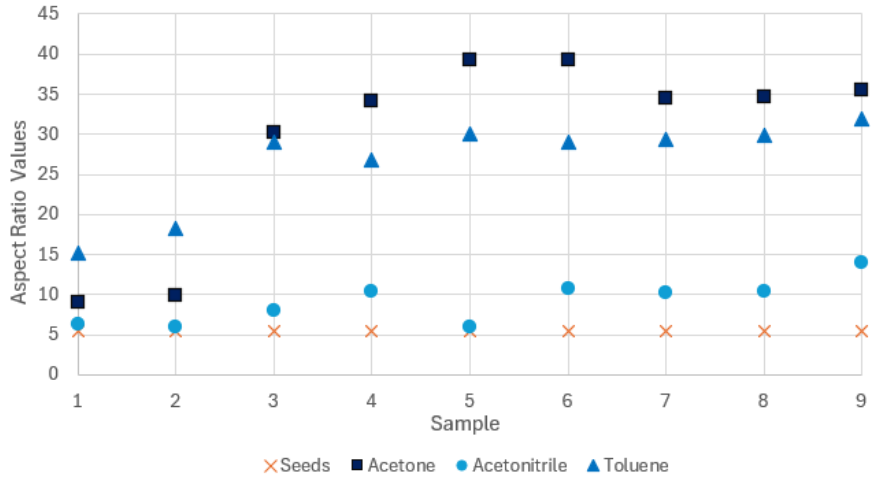
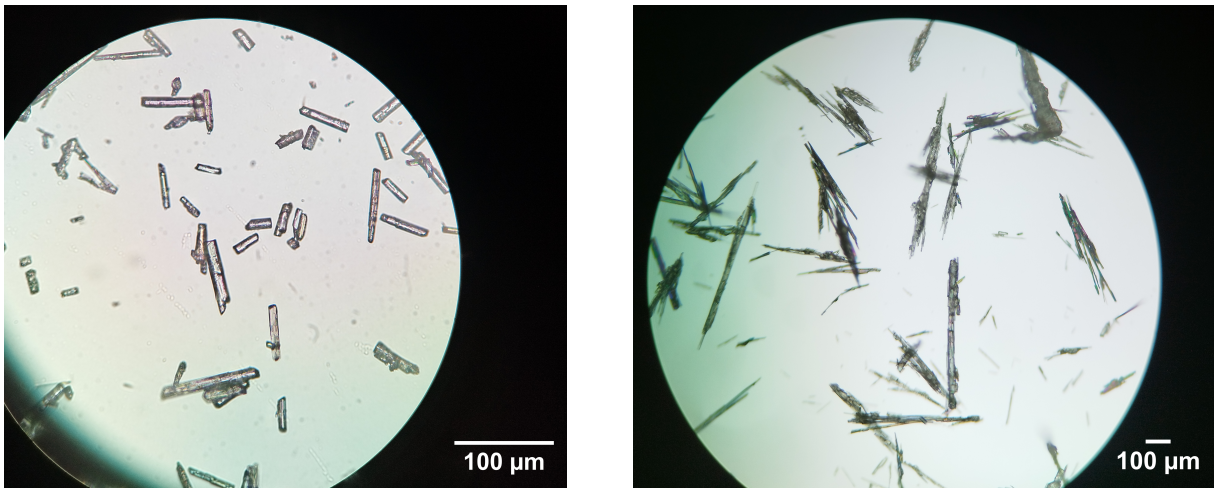


Figure 4.2: Comparison of the AR trends of xanthone in different solvents, the red line represents the AR of the seeds (AR = 5.5) as a baseline

The results specific to each solvent, along with their interpretation, are discussed below.

The experiment conducted in acetone revealed a moderate growth of xanthone, particularly when compared to its behavior in the other two solvents. Figure 4.3 illustrates the crystal evolution between the first and last samples, showing a preferential elongation along the longitudinal axis, while the width remains relatively stable throughout the experiment. Overall, the crystal remains needle-like in shape throughout the experiment, as depicted by the trend in Aspect Ratio (AR), which diverges from the unitary value (see Section 3.1.4), depicted in Figure 4.2.



(a) (b)

Figure 4.3: Xanthone in acetone, first sample (a, objective: 40x) and last (b, objective: 10x)

The progression of average size is summarized in Table 4.2. A gradual increase in both length and width is observed up to the sixth hour. After this point, a drop in average size is recorded, most likely due to secondary nucleation: the appearance of new, smaller crystals lowers the overall mean size of the sample, resulting in a non-linear growth trend, contrary to what would be expected if only growth from the original seeding were occurring. The subsequent increase in size observed in the final samples supports this hypothesis, suggesting continued growth of the overall crystal population despite the presence of secondary nuclei.

Table 4.2: Growth of xanthone in acetone, mean measurements

Sample	**Length (μm)**	**Width (μm)**
Sample 1 - 1 h	63.8	7.5
Sample 2 - 2 h	66.3	7.0
Sample 3 - 3 h	124.5	4.6
Sample 4 - 4 h	240.7	7.6
Sample 5 - 5 h	233.6	6.5
Sample 6 - 6 h	435.8	11.7
Sample 7 - 7 h	353.1	11.0
Sample 8 - 8 h	366.6	11.0
Sample 9 - 22 h	417.8	12.4

To better appreciate the size distributions in the two dimensions, which gets lost when observing only the average value, Figure 4.4 and 4.5 are provided, highlighting the evolution of crystal growth during the experiments. The choice of samples reflects the initial, intermediate, and final points of the process, thereby enabling a coherent evaluation of

the evolution of both size and shape distributions over time.
From the figures it is clear that at the initial stage (sample 1), the distribution is narrower and centered on smaller sizes, whereas at later stages (samples 5 and 9) the distribution progressively shifts towards larger sizes and broadens, reflecting ongoing crystal growth and the development of anisotropy in the population.

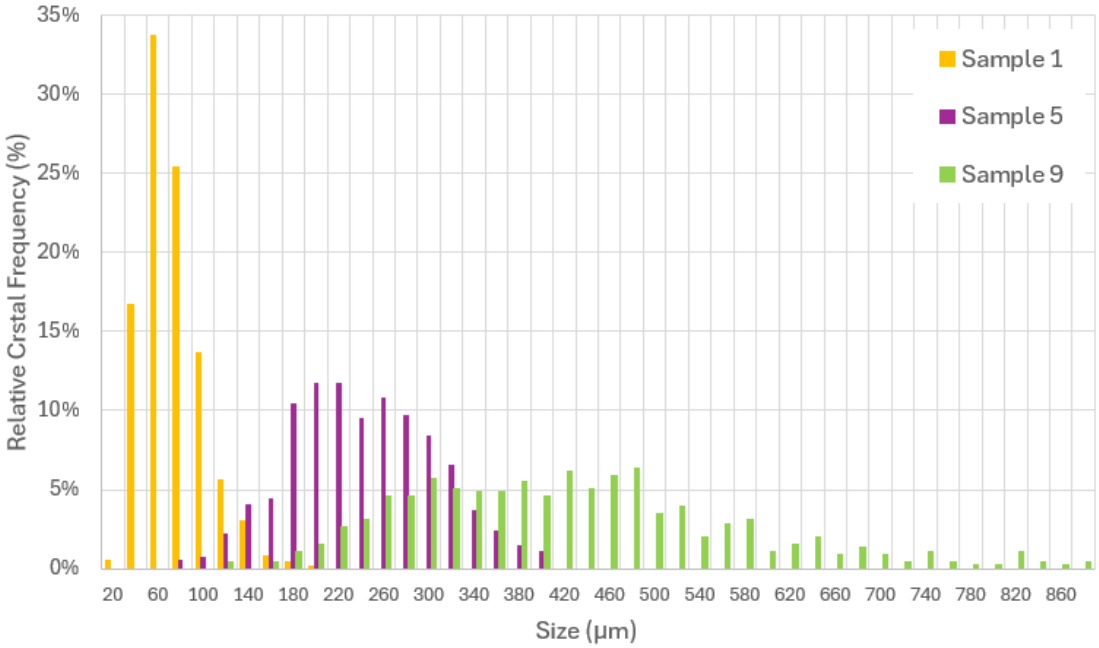


Figure 4.4: Evolution of Length (L_1) size distribution of xanthone crystals in acetone

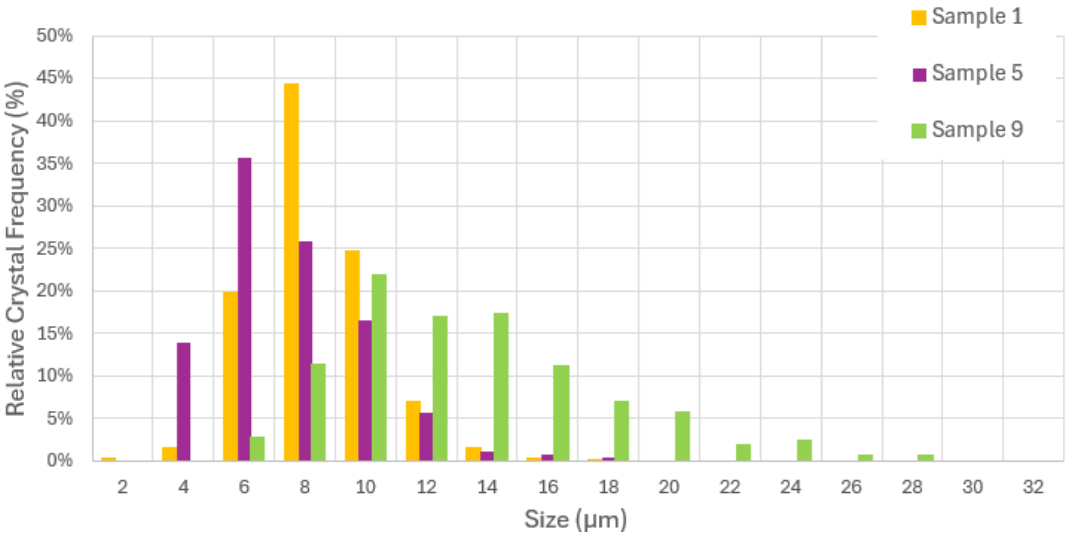
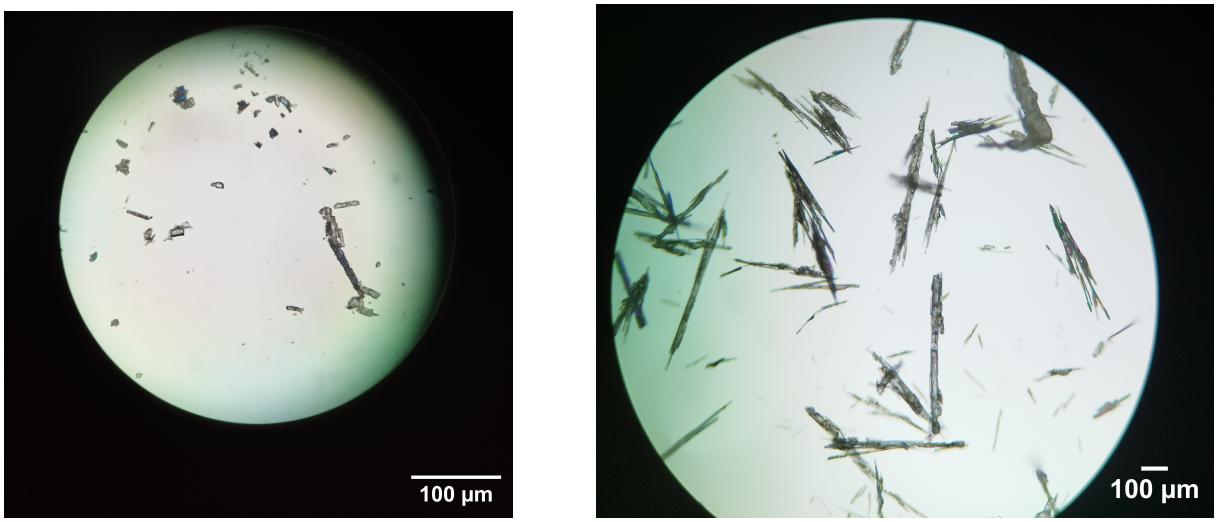


Figure 4.5: Evolution of Width (L_2) size distribution of xanthone crystals in acetone

Xanthone growth in acetonitrile is noticeably slower compared to its behavior in the other solvents. While the crystal maintains its characteristic needle-like morphology, a

more equivalent development along both the longitudinal and transverse axes is observed, resulting in a more squared overall shape. This morphological distinction is clearly visible in Figure 4.6, which also highlight the reduced dimensional growth, and further confirmed in the AR trend shown in Figure 4.2, which fluctuates less than the one in acetone or toluene.



(a) (b)

Figure 4.6: Xanthone in acetonitrile, first sample (a, objective: 40x) and last (b, objective: 40x)

Similar to what was observed in acetone, a peak followed by a decline in the mean crystal size is evident in Table 4.3. This trend is again attributed to the onset of secondary nucleation: the appearance of new, smaller crystals reduces the average size measured during sampling.

However, in this case, the decline occurs earlier in the experiment (Sample 5 instead of Sample 6), suggesting a different nucleation timing. The relatively low solubility of xanthone in acetonitrile may also account for the overall modest growth: fewer solute molecules in solution likely lead to a reduced number of growth sites, and therefore, to slower overall crystal development (see Section 2.1.3).

Table 4.3: Growth of xanthone in acetonitrile, mean measurements

Sample	**Length (μm)**	**Width (μm)**
Sample 1 - 1 h	48.10	8.4
Sample 2 - 2 h	48.47	8.5
Sample 3 - 3 h	71.23	9.4
Sample 4 - 4 h	93.80	10.0
Sample 5 - 5 h	71.50	12.0
Sample 6 - 6 h	97.60	10.0
Sample 7 - 7 h	100.90	10.8
Sample 8 - 8 h	103.70	10.7
Sample 9 - 22 h	140.70	10.8

The particular behavior of xanthone in acetonitrile is also evident when analyzing the evolution of the size distributions for length (Figure 4.7) and width (Figure 4.8). Although the trend resembles that observed in acetone, where later samples show broader and flatter distributions, which is indicative of crystal growth, the overall crystal sizes in acetonitrile remain comparatively smaller than in the other solvents, suggesting a slower overall growth rate.

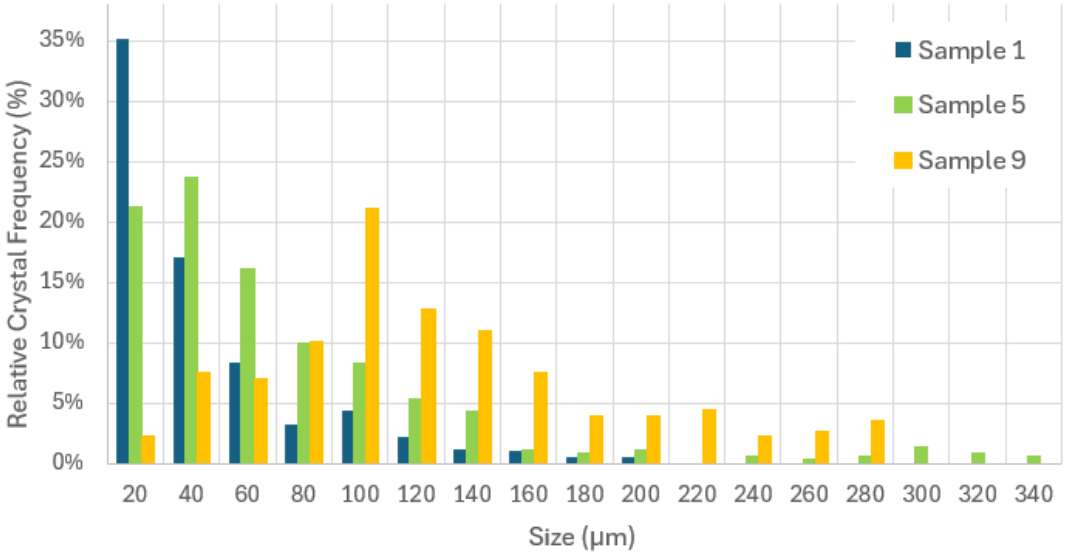


Figure 4.7: Evolution of Length (L_1) size distribution of xanthone crystals in acetonitrile

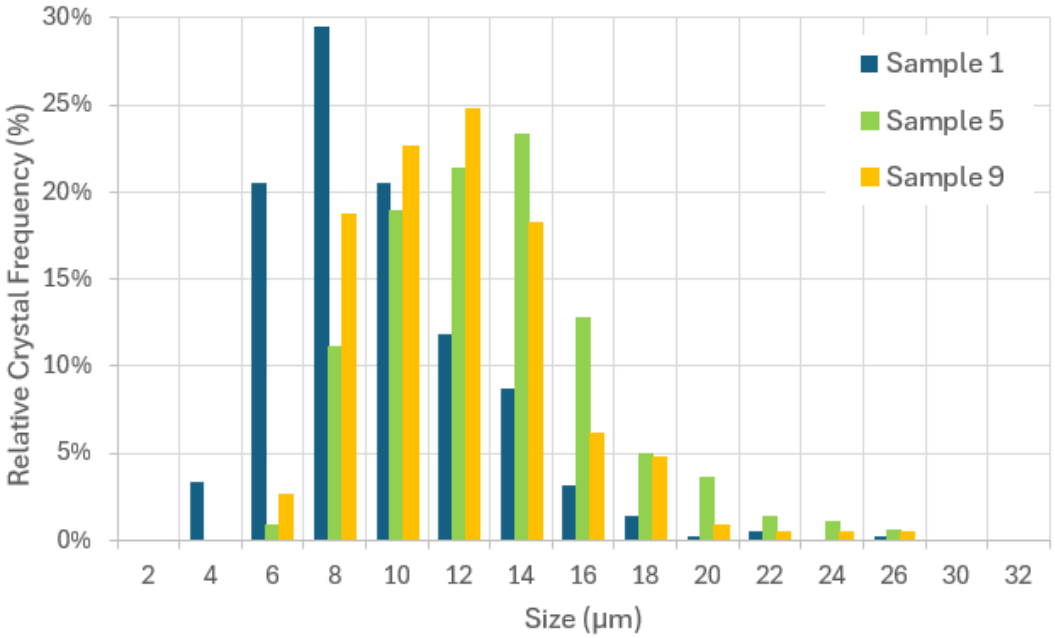
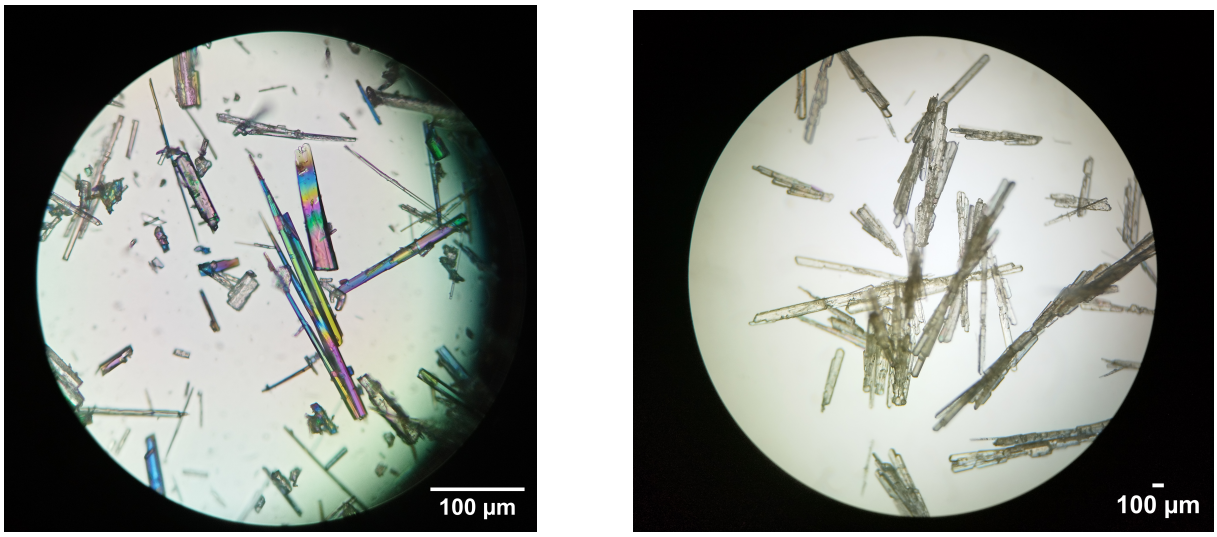


Figure 4.8: Evolution of Width (L_2) size distribution of xanthone crystals in acetonitrile

Among the three solvents analyzed, toluene produces the fastest crystal growth. The resulting particles show a marked elongation along the longitudinal axis, confirming a distinct needle-like morphology. This anisotropic growth is reflected in the Aspect Ratio trend reported in Figure 4.1, which further highlights the deviation from spherical symmetry over time.

Figure 4.9 visually demonstrates the dimensional progression of the crystals between the first and final samples, emphasizing the sharp increase in size and preferential elongation.



(a) (b)

Figure 4.9: Xanthone in toluene, first sample (a, objective: 40x) and last (b, objective: 5x)

As observed in the other two solvents, a decrease in mean size appears in Table 4.4 starting from Sample 6. This behavior is again attributed to the onset of secondary nucleation, which alters the distribution of particle sizes and introduces a deviation from a strictly increasing trend. Interestingly, this inflection point occurs at a comparable time to that observed in acetone.

Table 4.4: Growth of xanthone in toluene, mean measurements

Sample	**Length (μm)**	**Width (μm)**
Sample 1 - 1 h	92.90	7.35
Sample 2 - 2 h	279.30	18.2
Sample 3 - 3 h	1310.80	47.3
Sample 4 - 4 h	1348.60	53.3
Sample 5 - 5 h	1581.20	55.7
Sample 6 - 6 h	1741.40	62.0
Sample 7 - 7 h	1634.40	58.8
Sample 8 - 8 h	1689.54	59.4
Sample 9 - 22 h	1765.20	57.0

Furthermore, the evolution of the size distributions highlights the increased growth rate, particularly in length (Figure 4.10). The distributions not only broaden over time, but also shift markedly towards larger values: in the later samples (5 and 9), the crystals are drastically longer than in sample 1, with no overlap remaining between their size ranges. This behavior contrasts with the other solvents, where remnants of the initial distribution could still be observed in later samples. In terms of width (Figure 4.11), however, the distributions remain comparatively compact, underlining the strong anisotropy of xanthone crystallization in toluene.

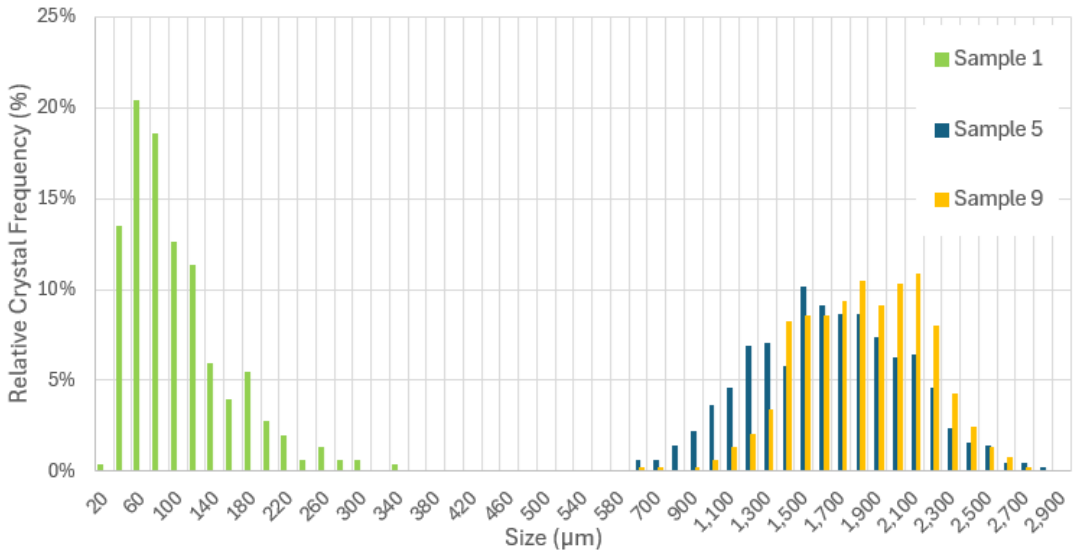


Figure 4.10: Evolution of Length (L_1) size distribution of xanthone crystals in toluene

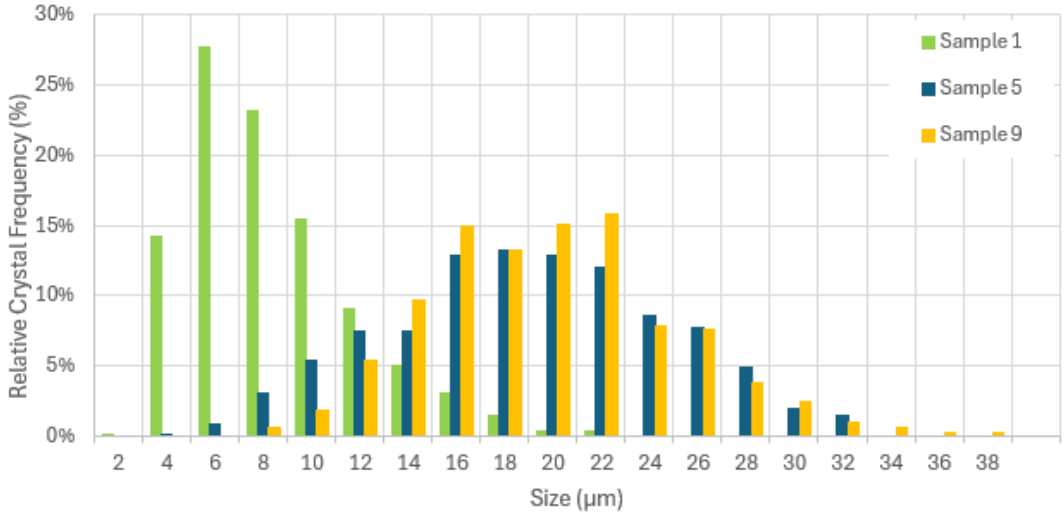


Figure 4.11: Evolution of Width (L_2) size distribution of xanthone crystals in toluene

4.1.2 Statistical Analysis

Comparison between calculations performed using moments and Cumulative Running Average

The results obtained by calculating the average geometrical properties, namely, the equivalent linear dimension and projected area, using both the cumulative approach and the moments show excellent agreement, as expected.
This consistency between the two approaches confirms the accuracy of the cumulative running average calculations and, at the same time, validates the implementation of the moments calculation for the dataset. It is worth noting that the moments, by definition, consider the entire crystal measured sample. Therefore, the observed match with the

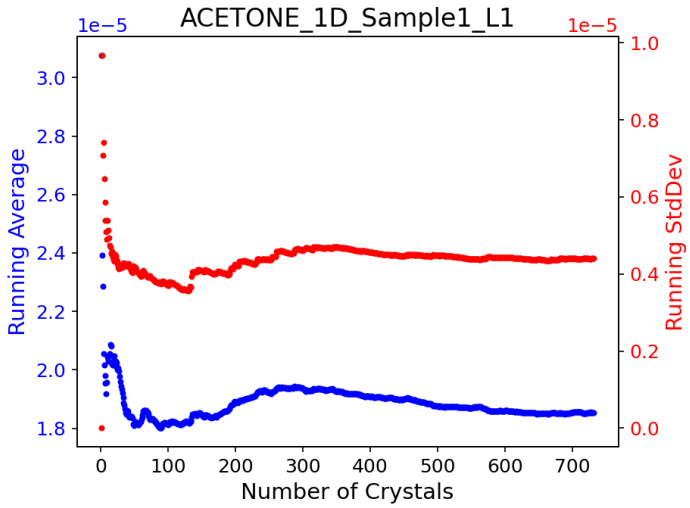
results from the cumulative approach, which directly averages the measured dimensions, indicates that no significant implementation errors were introduced by either method.

Statistical Relevance of the Properties for the individual Sample

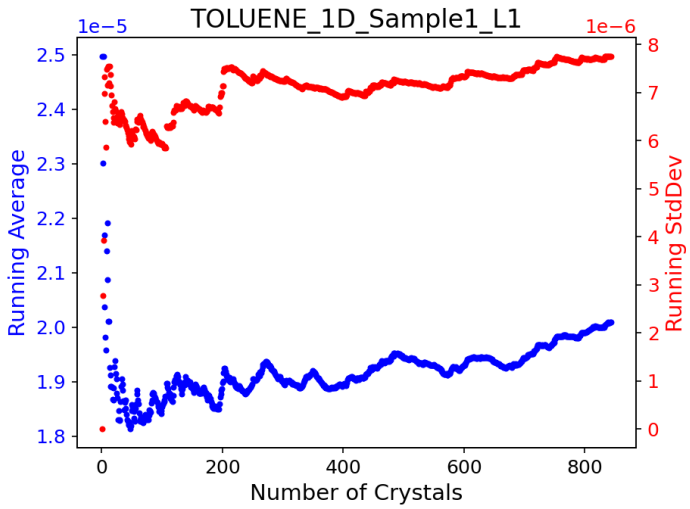
After validating the values obtained through the cumulative approach, the trend observed for each sample allows for classification into two categories, each with specific statistical implications:

- Convergence reached: If the functions representing the cumulative running average and the corresponding standard deviation stabilize at the end of the sample, that is, they cease to oscillate and remain approximately constant, this indicates that the number of crystals analyzed is sufficient for the statistical parameters to be considered reliable and representative of the underlying population.
- Convergence not reached: If the functions continue to fluctuate without stabilization, this suggests that the sample size is insufficient to reliably describe the population's characteristics in the investigated sample. In such cases, the extracted information may still offer insights, but must be interpreted cautiously, as it may not accurately reflect average population behavior.

Representative examples of both scenarios are depicted in Figure 4.12.



(a)



(b)

Figure 4.12: Plot of cumulative running average and its standard deviation with (a) and without (b) convergence

In most samples, convergence was successfully achieved, supporting the assumption that the analyzed crystals provide a representative snapshot of the entire crystal population at the sampling time. For clarity and conciseness, the plots of the cumulative running average and its standard deviation for all samples, in both 1D and 2D analyses, are reported in Appendix B. A collective summary is provided in Table 4.5, indicating which samples reached convergence. In this analysis, convergence was defined as a variation of no more than 5% in the mean value calculated over the last 100 crystals.

Table 4.5: Convergence analysis for 1D and 2D (L_1, L_2) analyses for each solvent and sample: the presence of the check mark ✓denotes that the calculations for the sample have reached convergence

Solvent	**Sample**	**1D - analysis**	**2D - analysis L_1**	**2D - analysis L_2**
	1	✓		✓
	2	✓	✓	
	3	✓	✓	✓
	4	✓	✓	✓
Acetone	5		✓	
	6	✓	✓	✓
	7	✓	✓	✓
	8	✓	✓	✓
	9	✓	✓	✓
	1	✓		✓
	2	✓		✓
	3	✓	✓	✓
	4	✓	✓	✓
Acetonitrile	5	✓		
	6	✓	✓	✓
	7			✓
	8			✓
	9			✓
	1			✓
	2	✓	✓	✓
	3	✓	✓	✓
	4	✓	✓	✓
Toluene	5	✓		✓
	6	✓	✓	
	7	✓		✓
	8	✓	✓	✓
	9	✓		

Statistical Relevance of the Sample

To assess whether the number of crystals analyzed in each sample could be deemed statistically representative of the total population, confidence intervals were calculated at two different confidence levels: 90% and 99%. This analysis estimated the reliability with

which the measured crystal properties reflect the actual values of the broader population, essentially quantifying how likely it is that a given measurement lies above a defined margin of error around the estimated parameter.
Confidence intervals were computed for all three properties investigated (average length, average area, and average volume) both for the one-dimensional analysis based on the equivalent diameter, and for the two-dimensional approach, which considered the two principal growth directions separately. However, it is important to note that if the number of crystals used to estimate the average length is insufficient to represent the population adequately, then any parameters derived from higher-dimensional properties (such as area or volume) will inevitably exhibit even greater uncertainty and error, making their interpretation less meaningful.
For this reason, although all calculations were carried out, only the results related to the average length are discussed in this chapter for the sake of clarity. Complete data for all the properties can be found in Appendix B.

The results of the confidence interval calculations (see Equation 2.36) and the corresponding estimates of the minimum number of crystals required (see Equation 2.37) are presented in the following tables: Table 4.6 reports the data for the equivalent diameter (1D analysis), while Tables 4.7 and 4.8 display the outcomes for the two-dimensional analysis, specifically for the L_1 and L_2 dimensions, respectively. The number of crystals used are found in Table 3.6.

Table 4.6: 1D Equivalent Diameter: mean length ($\bar{L}$), standard deviation of the mean (S), confidence intervals (CI_W) and relative confidence interval (RCI) for each solvent and sample

Solvent	**Sample**	$\bar{L}$ **(μm)**	S **(μm)**	CI_{W90}**(μm)**	CI_{W99}**(μm)**	RCI_{90}	RCI_{99}
Acetone	1 (1 h)	18.5	4.4	1.1	1.7	6%	9%
	2 (2 h)	19.2	4.8	1.0	1.6	5%	8%
	3 (3 h)	16.7	5.5	1.13	1.8	7%	11%
	4 (4 h)	29.2	9.8	1.9	3.0	7%	10%
	5 (5 h)	26.2	8.2	2.3	3.6	9%	14%
	6 (6 h)	48.0	14.5	4.1	6.5	9%	14%
	7 (7 h)	42.8	13.3	3.0	4.8	7%	11%
	8 (8 h)	43.7	14.0	3.3	5.11	7%	12%
	9 (22 h)	49.1	16.3	5.0	7.9	10%	16%
Acetonitrile	1 (1 h)	17.2	6.6	2.2	3.4	12%	20%
	2 (2 h)	17.6	7.5	2.0	3.1	11%	18%
	3 (3 h)	21.4	8.5	3.1	4.8	14%	23%

Solvent	**Sample**	$\bar{L}$ **(**μ**m)**	S **(**μ**m)**	CI_{W90}**(**μ**m)**	CI_{W99}**(**μ**m)**	RCI_{90}	RCI_{99}
	4 (4 h)	24.7	7.6	1.9	3.0	8%	12%
	5 (5 h)	26.0	8.0	2.8	4.4	11%	17%
	6 (6 h)	25.2	7.2	2.3	3.7	9%	15%
	7 (7 h)	27.0	6.9	2.3	3.6	9%	14%
	8 (8 h)	27.2	7.5	3.0	4.8	11%	18%
	9 (22 h)	30.5	8.3	3.6	5.6	12%	19%
	1 (1 h)	20.1	7.7	1.8	2.8	9%	14%
	2 (2 h)	51.1	28.5	6.2	9.8	12%	19%
	3 (3 h)	175.0	51.8	12.5	19.6	7%	11%
	4 (4 h)	192.0	40.0	12.9	20.2	7%	11%
Toluene	5 (5 h)	208.0	48.5	12.6	19.7	6%	9%
	6 (6 h)	231.0	46.1	13.4	21.0	6%	9%
	7 (7 h)	218.0	44.2	14.3	22.4	7%	10%
	8 (8 h)	223.0	45.5	15.8	24.7	7%	11%
	9 (22 h)	220.0	39.9	11.3	17.7	5%	8%

Table 4.7: 2D Length L_1: mean ($\bar{L_1}$), standard deviation of the mean (S), confidence intervals (CI_W) e RCI for each solvent and sample

Solvent	**Sample**	$\bar{L_1}$ **(**μ**m)**	S **(**μ**m)**	CI_{W90}**(**μ**m)**	CI_{W99}**(**μ**m)**	RCI_{90}	RCI_{99}
	1 (1 h)	63.9	26.9	6.6	10.3	10%	16%
	2 (2 h)	79.8	32.0	6.7	10.5	8%	13%
	3 (3 h)	124.0	47.9	9.9	15.5	8%	12%
	4 (4 h)	241.0	96.7	19.1	29.8	8%	12%
Acetone	5 (5 h)	234.0	67.8	19.2	30.1	8%	13%
	6 (6 h)	436.0	140.0	39.9	62.5	9%	14%
	7 (7 h)	353.0	121.0	27.9	43.7	8%	12%
	8 (8 h)	367.0	124.0	28.9	45.2	8%	12%
	9 (22 h)	418.0	150.0	46.0	72.1	11%	17%
	1 (1 h)	48.1	48.4	15.7	24.5	33%	51%
	2 (2 h)	48.5	45.7	12.0	18.8	25%	39%
	3 (3 h)	71.2	59.2	21.4	33.5	30%	47%
	4 (4 h)	93.9	64.2	16.0	25.0	17%	27%
Acetonitrile	5 (5 h)	71.6	45.0	15.6	24.4	22%	34%

Solvent	**Sample**	$\bar{L_1}$ **(μm)**	S **(μm)**	CI_{W90}**(μm)**	CI_{W99}**(μm)**	RCI_{90}	RCI_{99}
	6 (6 h)	97.7	57.2	18.7	29.3	19%	30%
	7 (7 h)	101.0	58.0	19.6	30.6	19%	30%
	8 (8 h)	104.0	57.7	23.3	36.5	22%	35%
	9 (22 h)	141.0	67.9	29.3	45.9	21%	33%
	1 (1 h)	92.9	57.2	13.0	20.3	14%	22%
	2 (2 h)	279.0	158.0	34.7	54.4	12%	19%
	3 (3 h)	1310.0	479.0	115.7	181.2	9%	14%
	4 (4 h)	1350.0	281.0	90.6	141.9	7%	11%
Toluene	5 (5 h)	1580.0	418.0	108.7	170.2	7%	11%
	6 (6 h)	1740.0	458.0	133.2	208.5	8%	12%
	7 (7 h)	1630.0	410.0	132.4	207.4	8%	13%
	8 (8 h)	1690.0	347.0	120.5	188.6	7%	11%
	9 (22 h)	1770.0	343.0	97.2	152.2	6%	9%

Table 4.8: 2D Length L_2: mean ($\bar{L_2}$), standard deviation of the mean (S), confidence intervals (CI_W) e RCI for each solvent and sample

Solvent	**Sample**	$\bar{L_2}$ **(μm)**	S **(μm)**	CI_{W90}**(μm)**	CI_{W99}**(μm)**	RCI_{90}	RCI_{99}
	1 (1 h)	7.5	1.9	0.5	0.7	6%	10%
	2 (2 h)	7.0	2.1	0.4	0.7	6%	10%
	3 (3 h)	4.6	2.0	0.4	0.6	9%	14%
	4 (4 h)	7.6	3.0	0.6	0.9	8%	12%
Acetone	5 (5 h)	6.5	2.6	0.7	1.2	12%	18%
	6 (6 h)	11.7	4.0	1.2	1.8	10%	15%
	7 (7 h)	11.0	4.0	0.9	1.5	8%	13%
	8 (8 h)	11.1	4.0	0.9	1.5	8%	13%
	9 (22 h)	12.4	4.6	1.41	2.2	11%	18%
	1 (1 h)	8.4	3.2	1.0	1.6	12%	19%
	2 (2 h)	8.5	3.3	0.9	1.3	10%	16%
	3 (3 h)	9.4	3.3	1.2	1.9	13%	20%
	4 (4 h)	10.0	3.4	0.8	1.3	8%	13%
Acetonitrile	5 (5 h)	12.1	3.6	1.3	2.0	10%	16%
	6 (6 h)	10.0	3.3	1.1	1.7	11%	17%
	7 (7 h)	10.8	3.4	1.1	1.8	10%	16%

Solvent	**Sample**	$\bar{L}_2$ **(**μ**m)**	S **(**μ**m)**	CI_{W90}**(**μ**m)**	CI_{W99}**(**μ**m)**	RCI_{90}	RCI_{99}
	8 (8 h)	10.7	3.5	1.4	2.2	13%	20%
	9 (22 h)	10.8	3.4	1.5	2.3	14%	21%
	1 (1 h)	7.4	3.5	0.8	1.2	11%	17%
	2 (2 h)	18.2	10.3	2.3	3.5	12%	19%
	3 (3 h)	47.3	14.9	3.6	5.7	8%	12%
	4 (4 h)	53.3	14.2	4.6	7.2	9%	13%
Toluene	5 (5 h)	55.7	15.5	4.0	6.3	7%	11%
	6 (6 h)	62.0	14.1	4.1	6.4	7%	10%
	7 (7 h)	58.8	14.7	4.7	7.4	8%	13%
	8 (8 h)	59.4	15.3	5.3	8.3	9%	14%
	9 (22 h)	56.9	12.7	3.6	5.7	6%	10%

Although the confidence interval on its own does not provide substantial inputs for further analysis, it is still meaningful to compare the CI values obtained from sample analysis with the optical microscope resolution, reported in Table 3.7. The two differ, as expected, since experimental measurements rarely achieve the same precision as the instrument resolution; however, the discrepancy remains limited. This indicates that, while sampling could be improved by repeating the experiments, it is nonetheless adequate for supporting the qualitative evaluations performed in this study.

In addition, a comparison between the confidence interval and the convergence of the cumulative running average was carried out. As stated previously (see Section 3.1.5), a confidence interval was classified as narrow when RCI$<$ 15%. By comparing the results reported in Tables 4.6–4.8 with the convergence analysis shown in Table 4.5, four different cases were identified, each leading to a specific interpretation of the representativeness and reliability of the analyzed sample:

- Case A - Convergence reached and RCI$<$ 15%: the sample can be considered representative, and the calculated mean value reliable;
- Case B - Convergence reached but RCI$>$ 15%: the sample requires a larger number of crystals to improve reliability, thus the quantitative analysis should be viewed as indicative rather than conclusive;
- Case C - Convergence not reached and RCI$<$ 15%: the number of collected data points is insufficient to draw fully reliable conclusions;
- Case D - Convergence not reached and RCI$>$ 15%: the sample cannot be considered representative.

It should be emphasized that representativeness here refers to the investigated sample only, and not to the entire crystal population.
To synthesize the previous considerations, Table 4.9 indicates the case associated with each sample, and Figure 4.13 illustrates the corresponding distribution in terms of relative frequency.

Table 4.9: Case analysis: results of comparison between convergence studies and cumulative running average calculations

Solvent	**Sample**	**1D - analysis**		**2D - analysis L_1**		**2D - analysis L_2**	
		$\alpha = 90\%$	$\alpha = 99\%$	$\alpha = 90\%$	$\alpha = 99\%$	$\alpha = 90\%$	$\alpha = 99\%$
Acetone	1	A	A	C	D	A	A
	2	A	A	A	A	C	C
	3	A	A	A	A	A	A
	4	A	A	A	A	A	A
	5	C	C	A	A	C	D
	6	A	A	A	A	A	B
	7	A	A	A	A	A	A
	8	A	A	A	A	A	A
	9	A	B	A	B	A	B
Acetonitrile	1	A	B	D	D	A	B
	2	A	B	D	D	A	B
	3	A	B	B	B	A	B
	4	A	A	B	B	A	A
	5	A	B	D	D	C	D
	6	A	B	B	B	A	B
	7	C	A	D	D	A	B
	8	C	B	D	D	A	B
	9	C	B	D	D	A	B
Toluene	1	C	C	D	D	A	B
	2	A	B	A	B	A	B
	3	A	A	A	A	A	A
	4	A	A	A	A	A	A
	5	A	A	C	C	A	A
	6	A	A	A	A	C	D
	7	A	A	C	C	A	A
	8	A	A	A	A	A	A
	9	A	A	C	C	C	C

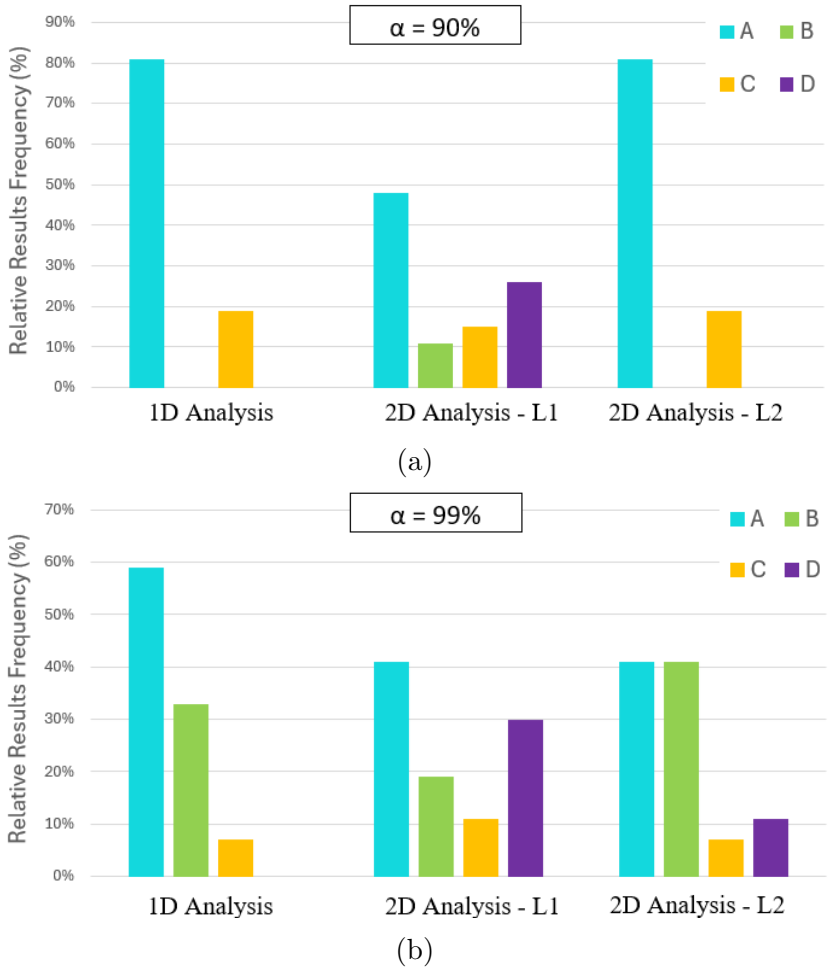


Figure 4.13: Relative Frequency of the results with $\alpha = 90\%$ (a) and $\alpha = 99\%$ (b)

As shown above, the number of crystals collected can, in most cases, be considered reliable. For the instances in which this is not valid, while this does not invalidate the qualitative or quantitative analyses performed, it does call for careful interpretation: the calculated confidence intervals are based on high confidence levels (90% and 99%) applied to single-sample experiments without replicates, an experimental limitation that inherently reduces the robustness of statistical inference and the ability to generalize results.

Therefore, the statistical findings presented here do not undermine the validity of the analyses performed, but rather help contextualize their scope. The measured values should not be interpreted as absolute descriptors of the entire crystal population, but as representative of the specific samples analyzed. These data remain sufficiently reliable for the purposes of this study, while reflecting real crystallization behavior with an acknowledged degree of partiality and statistical limitation.

4.2 Computational Results

4.2.1 Property validation

The analysis of the computed properties shows good agreement with established models and reference data from literature, supporting the validity of the simulation setup. In particular, the properties of the two orthogonal faces, AC and BC, were examined under the assumption that, owing to their structural and morphological similarities, the third orthogonal face would exhibit analogous behavior. For clarity, only the results for the AC face are presented in this work, since in the liquid phase the properties are independent from the box geometry. The detailed results for each property are discussed below.

The density estimates, reported in Table 4.10, show good accuracy for the analyzed faces, with a small relative error compared to literature values[26][27][28]. This agreement supports the reliability of subsequent property analyses and allows to consider as accurate the initial disposition of the models in terms of box dimension and volume.

Table 4.10: Comparison between solvents' density results obtained from the simulations (right) and literature (left) for acetone, acetonitrile and toluene, face AC

Density (g/cm^3)			
Solvent	AC		err (%)
	literature	LAMMPS	
Acetone	0.789[26]	0.787	0.20
Acetonitrile	0.786[27]	0.785	0.19
Toluene	0.867[28]	0.866	0.16

To obtain these results, the mean value of the density evolution among each iteration was calculated, as explained in detail in Section 3.2.2.

The analysis of the Radial Distribution Function (RDF) trends for the relevant atom–pairs within each solvent molecule is a key step in validating both the selected force field and the accuracy of interatomic distances. The RDF enables the comparison between the simulated molecular arrangement and literature-reported reference values, providing a basis for considering the subsequent property evaluations as reliable.

For each solvent, the most representative atom–pair interactions were selected and are listed in Table 4.11. The nomenclature adopted is as follows:

- C1: carbon atom belonging to the main molecular chain;

- C2: carbon atom belonging to a methyl group;
- H: hydrogen atom;
- N: nitrogen atom (present only in acetonitrile);
- O: oxygen atom (present only in acetone).

Table 4.11: List of the atom-pair interactions considered during RDF evaluation for acetone, acetonitrile and toluene

Atom-pair	**Acetone**	**Acetonitrile**	**Toluene**
C1C1	yes	yes	yes
C2C2	yes	yes	yes
C1H			yes
HH			yes
NC1		yes	
NC2		yes	
NH		yes	
NN		yes	
OC2	yes		
OH	yes		
OO	yes		

In evaluating each RDF curve, the following aspects were considered:

- Position of the first peak, as it provides the most probable interatomic distance between the two atom types and allows direct comparison with experimental and literature data;
- Width of the first peak, as it indicates the degree of structural order. A narrow, high peak corresponds to a well-defined coordination shell, whereas a broader peak suggests greater positional disorder;
- Overall trend of the function: it reveals the level of short- and medium-range order in the system, useful for distinguishing between liquid-like and solid-like behavior;
- Asymptotic value at large distances: in liquids, the RDF should approach a value of 1 as $r \to \infty$, indicating a homogeneous particle distribution at long range; deviations from this behavior can signal artifacts or non-equilibrated simulations.

The complete set of RDF curves for all atom–pair interactions listed in Table 4.11 is reported in Appendix C. Here, only representative trends are presented in a condensed form (Figure 4.14, 4.15 and 4.16), with each plot showing the RDFs for all selected atom–pairs within a given solvent.

In all cases, the simulated RDFs show:

- First peak positions in agreement with literature data, confirming that the interatomic distances predicted by the selected force field are physically consistent[29][30][31].
- Peak widths and amplitudes compatible with the liquid state of the systems, with sharper peaks for nearest-neighbor shells and progressively smoother features at larger r.
- Correct asymptotic convergence to $g(r) = 1$, indicating the expected long-range homogeneity of the liquid phase.

These results confirm the validity of the constructed simulation boxes and justify the reliability of the subsequent property evaluations. Detailed, solvent-specific RDF plots and quantitative peak analyses are provided in Appendix C.

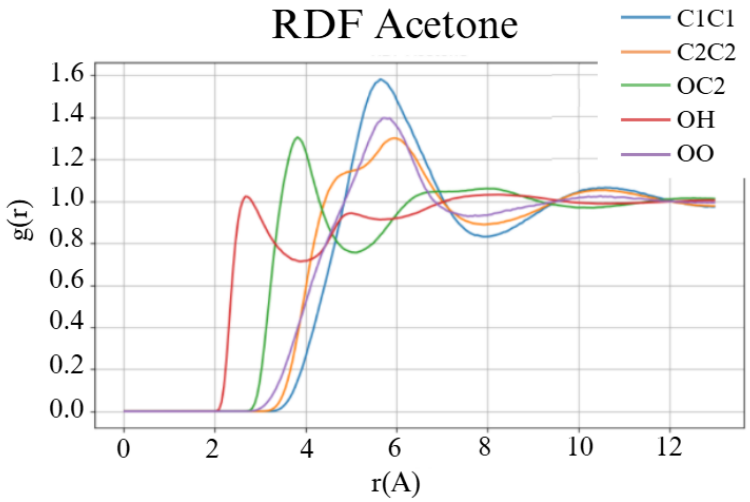


Figure 4.14: Calculated RDF trends for chosen atom-pair in acetone

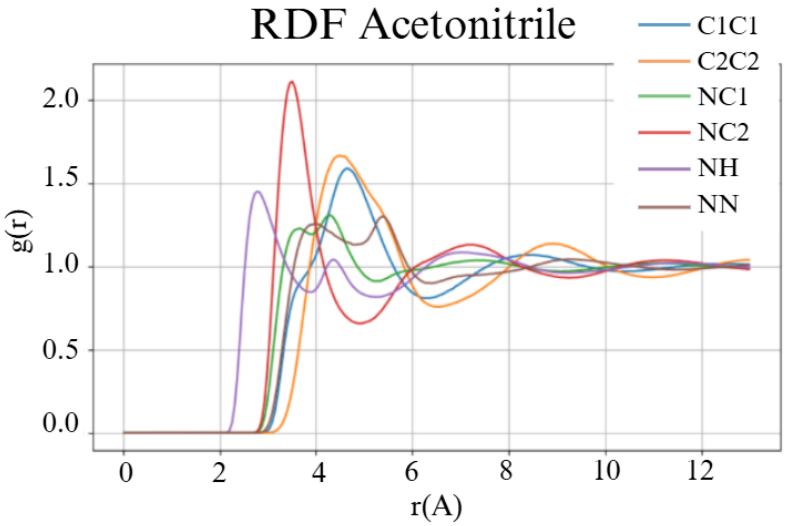


Figure 4.15: Calculated RDF trends for chosen atom-pair in acetonitrile

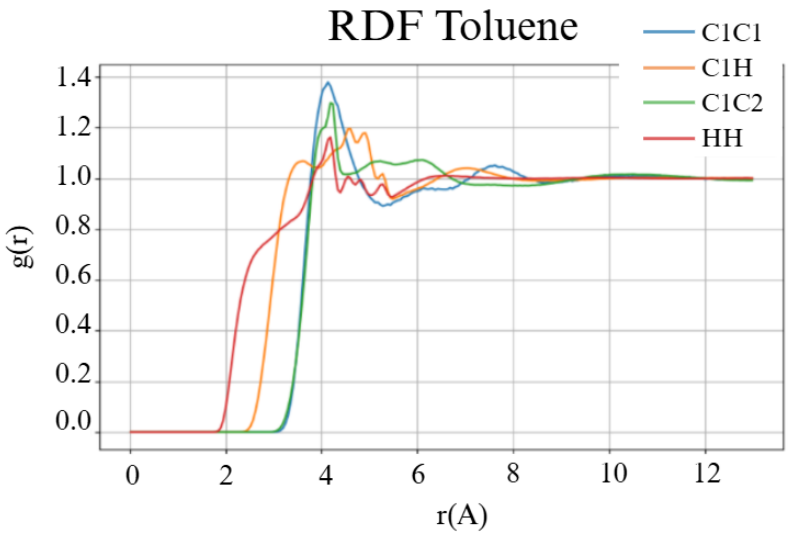


Figure 4.16: Calculated RDF trends for chosen atom-pair in toluene

As the final property, diffusivity was evaluated to verify that the spatial interactions between particles in the system were consistent with literature data[32][33][34]. The results, reported in Table 4.12, show a relative error below 10%in all cases, which is within the accepted range for experimental validation. This agreement further confirmed the reliability of the chosen force field and provided a solid basis for the subsequent energy calculations.

Table 4.12: Comparison between solvents' diffusivity results obtained from the simulations (right) and literature (left) for acetone, acetonitrile and toluene, face AC

Diffusivity $\times 10^{-9}$ **(m^2/s)**			
Solvent	AC		err (%)
	literature	LAMMPS	
Acetone	4.80[32]	4.70	2.11
Acetonitrile	4.37[33]	4.66	6.74
Toluene	2.66[34]	2.87	7.89

4.2.2 Energy calculations

The complete set of results, for both the individual monocomponent boxes and the merged systems, is reported in Table 4.13. For clarity, the different energy components listed in the table are defined as follows:

- Potential: the total potential energy of the system, obtained as the sum of all

contributions (Van der Waals, Coulomb, bond, angle, and dihedral energies), as detailed in Section 2.2.2;

- Van der Waals: the energy associated with Van der Waals interactions, modeled through the Lennard–Jones potential;
- Coulomb: the total Coulomb potential energy, given by the sum of short- and long-range electrostatic interactions;
- E_{long}: the long-range Coulomb potential, accounting for electrostatic interactions between molecules at larger distances;
- $E_{Coulomb}$: the short-range Coulomb potential, describing the energy contribution from interactions between molecules in close proximity.

Table 4.13: Results of energy calculation for single-component and merged system

BOX		**ENERGIES (cal/mol)**				
Group	**System**	**Potential**	**Van der Waals**	**Coulomb**	**E_{long}**	**$E_{Coulomb}$**
PURE	Xanthone	2865.1	-4577.9	1893.2	-8358.2	10251.4
PURE	Acetone AC	2715.8	-1327.3	318.5	-11081.5	11400.0
PURE	Acetone BC	3689.9	-1810.2	435.6	-15562.2	15997.8
PURE	Acetone AB	2747.1	-1298.2	329.3	-11180.9	11510.2
PURE	Acetonitrile AC	-2121.3	-1990.1	-3261.1	-17448.8	14187.7
PURE	Acetonitrile BC	-2967.63	-2764.72	-4502.0	-23394.68	18892.64
PURE	Acetonitrile AB	3867.38	-1463.31	-4757.9	-23449.36	18691.50
PURE	Toluene AC	23248.4	-2053.8	918.9	-2104.6	3023.6
PURE	Toluene BC	27982.5	-4027.4	1248.0	-2941.2	4189.2
PURE	Toluene AB	95585.9	17333.4	1409.3	-2911.4	4320.8
MERGED	Acetone AC	-11675.5	-854.9	-20600.7	-19712.0	-888.7
MERGED	Acetone BC	-13816.2	-884.0	-25257.2	-23472.6	-1784.6
MERGED	Acetone AB	-12088.6	-1045.3	-20787.7	-19480.0	-1307.7
MERGED	Acetonitrile AC	-11196.5	-626.0	-19695.2	-24961.1	5265.9
MERGED	Acetonitrile BC	-13226.8	-621.8	-24120.7	-31403.3	7282.7
MERGED	Acetonitrile AB	-10066.8	-579.5	-25216.0	-31409.1	6193.1
MERGED	Toluene AC	5760.3	3043.3	-9263.1	-10189.5	926.4
MERGED	Toluene BC	7079.9	3476.7	-9852.5	-10852.5	1000.0
MERGED	Toluene AB	71735.6	23533.6	-10098.7	-10985.0	886.4

As such, following equation 2.57, it was possible to evaluate the first parameter for the calculation of Helmholtz free energy, with the equation adapted to the elements of the simulation following equation 4.1: the results are shown in Table 4.14 and, for better

understanding of the individual trends in different solvents, represented in Figure 4.17.

$$U = U_{X+S} - (U_{BX} + U_{BS}) \tag{4.1}$$

where U_{X+S} is the internal potential energy of the merged systems, while U_{BX} and U_{BS} are the potential energy of the isolated xanthone bulk system and isolated Solvent bulk system respectively.

Table 4.14: Internal energy value calculation, first term of Helmholtz energy equation

$U_{\text{interface}}$ **(cal/mol)**			
Face	**Acetone**	**Acetonitrile**	**Toluene**
AC	-17256.4	-11940.3	-20353.3
BC	-20371.2	-13124.3	-23767.7
AB	-17700.8	-16799.3	-26715.4

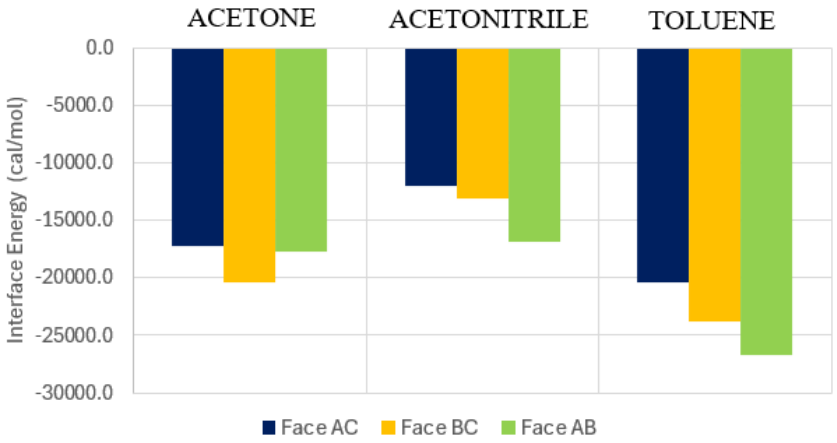


Figure 4.17: Visual representation of the different energy values for faces AB (001),AC (010),BC (100) in acetone, acetonitrile and toluene

Through the analysis of the data reported in Tables 4.14 and 4.15, and Figures 4.17 and 4.18, several key findings can be highlighted.
First, when comparing the three solvents, acetonitrile exhibits the highest interface energy values with the xanthone crystal, followed by acetone, while toluene consistently shows the lowest values. This means that, from a thermodynamic perspective, the crystal–solvent interface in toluene is the most favorable, facilitating crystal growth. As the system tends to minimize energetic expenditure, it can therefore be expected that xanthone crystals will grow more rapidly in toluene, while acetonitrile will promote the slowest growth, with

acetone behaving as an intermediate case.

Looking more closely at the individual faces, in acetone the face (100) displays the lowest interface energy, suggesting that growth will preferentially occur along the direction perpendicular to this surface. In acetonitrile, face (001) is the most stabilized, whereas in toluene face (001) shows by far the lowest energy, strongly favoring its growth compared to the other two faces. Conversely, face (010) is consistently the least stable across all solvents, presenting the highest interface energy values. This consistency indicates that face (010) is generally the most inhibited in growth, independently of the solvent considered.

The differences in interface energies between faces, summarized in Table 4.13 and Figure 4.11, further support these considerations. Toluene shows the largest relative differences, suggesting a strongly anisotropic growth, with a clear preference for specific directions. Acetonitrile, instead, presents the smallest differences, pointing toward a more homogeneous, less directional growth. Acetone again lies in between, with moderate differences that suggest some degree of preferential growth, but not as marked as in toluene.

These results generally outline a coherent picture: the solvent not only affects the overall growth rate (with toluene > acetone > acetonitrile), but also strongly influences the anisotropy of growth, determining whether the crystal will preferentially extend along one axis or more evenly across its structure.

Table 4.15: Interface energy differences between the three investigated faces for the three solvents

Energy differences Δ (cal/mol)			
Face comparison	**Acetone**	**Acetonitrile**	**Toluene**
Δ(AC–BC)	3114.8	1184.1	3414.5
Δ(AB–BC)	2670.4	-3674.9	-2947.7
Δ(AC–AB)	444.4	4859.0	6362.2

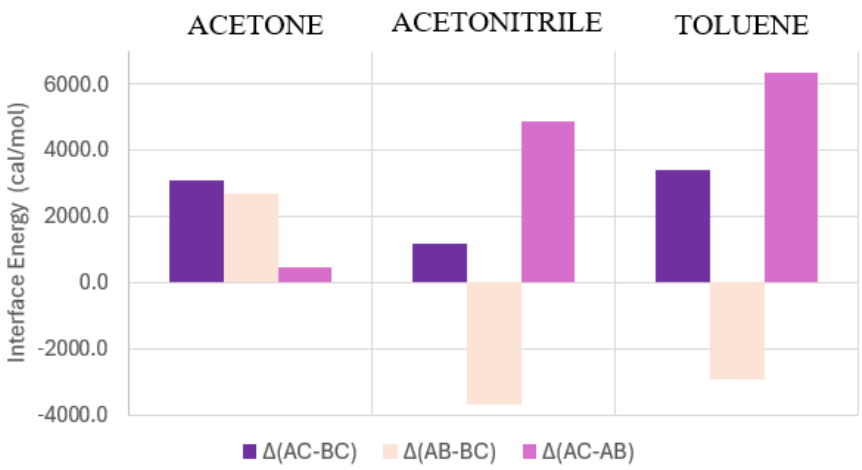


Figure 4.18: Visual representation of the interface energy differences between the three investigated faces for the three solvents

4.3 Result Discussion through Data Comparison

In order to validate the central hypothesis of the present work, namely, to understand if the evaluation of only the internal free energy is sufficient to describe the property of the crystal interface, a comparison between the experimental observations and the computational data is required.

Before proceeding with the integrated analysis, it is first necessary to summarize the results obtained from both approaches. This step also serves to align the different nomenclatures used to describe the crystal faces, thereby ensuring greater clarity in the comparative discussion that follows.

From an experimental perspective, the crystallization of xanthone in the three solvents revealed a distinct preferential behavior, both in terms of growth rate and anisotropy. Toluene exhibited the fastest crystal growth, followed by acetone and finally acetonitrile. This same trend was reflected in the degree of anisotropy observed: toluene crystals showed the most pronounced directional growth, while acetonitrile favored more homogeneous expansion.

From a computational standpoint, significant differences were observed among the solvents with respect to the internal energy values. Specifically, toluene showed the lowest interfacial energy across all simulated faces, followed by acetone and, lastly, acetonitrile. Furthermore, the simulations identified the energetically most favorable faces as follows: face AB (corresponding to the crystallographic face (001)) for both toluene and acetonitrile, and face BC (100) for acetone.

Comparison between the experimental and computational analyses reveals a partial mis-

match in the predicted growth directions. According to the simulations, the crystal faces with the lowest internal energy values (face AB (001) in toluene and acetonitrile, and face BC (100) in acetone) would be expected to grow preferentially, consistent with the thermodynamic principle that systems evolve toward lower-energy configurations. This trend, however, is not fully reflected in the experimental observations. In acetone, for example, the AB face (001), which is the widest and thus more prone to strong solvent interactions, was expected to grow more slowly; yet computational results assign it a lower energy, implying a tendency toward growth. Conversely, in toluene and acetonitrile, face BC (100) does not exhibit the lowest internal energy but nonetheless shows substantial growth in the experiments.

On a broader scale, however, the general growth trends of xanthone across the three solvents are consistently confirmed. Both the experimental and computational results agree that the crystal grows fastest in toluene, followed by acetone and then acetonitrile. This agreement supports the idea that internal energy is a good predictor of the overall growth tendency in different solvents.

From these observations, two key conclusions can be drawn:

- The comparative analysis indicates that internal energy, considered without the second entropic term, is insufficient to provide an accurate description of crystal face behavior. As a result, this limitation could explain the inconsistencies observed with experimental data.
- While internal energy alone does not allow accurate predictions of morphological development at the molecular level, it is nonetheless a reliable indicator for assessing general growth kinetics across different solvents. Though it cannot predict the dominant face or final morphology of the crystal, it can serve as a useful screening parameter in early-stage analyses, particularly in batch crystallization processes, where solvent selection, process timeframes, and expected growth rates are critical.

Chapter 5

Conclusions

This work investigated the polymorphic behavior of xanthone in three different solvents (acetone, acetonitrile, and toluene) through a dual approach combining laboratory experiments and molecular dynamics simulations.

The primary objective was to correlate the experimental results with computational estimates of interfacial energies, aiming to eventually use simulation as a predictive tool for crystallization behavior. In particular, the internal energy of each solvent–crystal interface was computed under the assumption that it represents the dominant term of the Helmholtz free energy equation, allowing the entropic component to be neglected. If valid, this simplification would imply that the crystal faces with the lowest internal energy should exhibit the fastest growth rates, as thermodynamically favored pathways.

Experimentally, xanthone demonstrated the highest and most anisotropic growth in toluene, followed by acetone, and finally acetonitrile, where both the growth rate and anisotropy were reduced. Detailed image and statistical analysis, employing both the calculation of moments and the cumulative running average approach, show that, although sampling could be improved, the discrepancy remains limited, supporting the validity of the qualitative evaluations carried out in this study.

Cross-validation with molecular simulations revealed both consistencies and discrepancies. While overall solvent-dependent trends aligned well, with toluene promoting faster growth and acetonitrile the slowest, the analysis of individual crystal faces did not fully support the initial assumption. In several instances, the faces that were predicted to grow fastest (due to low interfacial energy) did not do so experimentally, suggesting that internal energy alone is not a sufficient descriptor of face-specific growth behavior.

These findings highlight a key conclusion: while internal energy calculations can provide reliable comparative insight into overall crystal growth kinetics across different solvents, they are insufficient for accurate prediction of crystal morphology, particularly in terms

of face-specific development and anisotropy.

In conclusion, the integration of experimental and computational analyses has led to a more comprehensive understanding of xanthone's crystallization dynamics. Although some inconsistencies emerged, the broader alignment of growth tendencies confirms the usefulness of molecular dynamics simulations as a screening tool for the design of crystallization processes. Future work should focus on incorporating the entropic component of the Helmholtz free energy to improve the predictive accuracy of morphological outcomes. Additionally, calculating face-specific growth rates and replicates of experiments will offer a deeper insight into the underlying mechanisms. This combined approach sets a solid foundation for more rational crystallization design, both in academic research and industrial applications.

Appendix A

Python Codes

This appendix contains the Python codes developed and implemented in the course of this work, each preceded by a brief description of its purpose and functionality.

Code A.1 was employed to analyze the measures taken for the crystals and provide the values of cumulative running average and cumulative running standard deviation both for the one dimensional and two-dimensional analysis, completed with plots.

Listing A.1: Data processing and plotting for crystal samples 1D analysis

```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

#File input
input_excel = rExcel_file
output_excel = Excel_file_output # path to create the new excel work sheet with
    outputs
plots_dir = graphs_samples1  # new folder for graphs
dot_size = 8  # personalized dimensions for scatter plot
# Insert personalized dictionary with the name of the Excel
#from which the data is recollected from, specifying the columns that contain the
    dataset
#example given is for the 1D analysis,
#the 2D case will have two columns ('L1' and 'L2' for each sample)
experiments_config = {
    'ACETONE_1D': {
        'sheet': 'ACETONE_1D',
        'time_columns': {
            1: 'F',   # one dimensions = only one column considered
            2: 'U',
            3: 'AJ',
            4: 'AY',
```

```
            5: 'BN',
            6: 'CC',
            7: 'CR',
            8: 'DG',
            22: 'DV'
        }
    },
    'TOLUENE_1D': {
        'sheet': 'TOLUENE_1D',
        'time_columns': {
            1: 'F',
            2: 'U',
            3: 'AJ',
            4: 'AY',
            5: 'BN',
            6: 'CC',
            7: 'CR',
            8: 'DG',
            22: 'DV'
        }
    },
    'ACN_1D': {
        'sheet': 'ACN_1D',
        'time_columns': {
            1: 'F',
            2: 'U',
            3: 'AJ',
            4: 'AY',
            5: 'BN',
            6: 'CC',
            7: 'CR',
            8: 'DG',
            22: 'DV'
        }
    }
}


# excel to python utility
def col_letter_to_index(col_letter: str) -> int:
    #Converts 'A'->0, 'B'->1, ..., 'Z'->25, 'AA'->26, etc for python coding
    col_letter = str(col_letter).strip().upper()
    num = 0
    for ch in col_letter:
        if not ('A' <= ch <= 'Z'):
            raise ValueError(fNon valid column letter: {col_letter})
```

```
        num = num * 26 + (ord(ch) - ord('A') + 1)
    return num - 1

def to_numeric_series(series: pd.Series) -> pd.Series:

    #Robust conversion of the letters in numerical series:
    #no spaces, comma->dot, NaN if not number

    s = series.astype(str).str.strip().str.replace(,, ., regex=False)
    return pd.to_numeric(s, errors=coerce)

def running_stats(values: np.ndarray):
    #Returns (running_mean, running_std) cumulativi, std campionaria (ddof=1)
    running_mean = []
    running_std = []
    for i in range(1, len(values)+1):
        chunk = values[:i]
        m = float(np.mean(chunk))
        s = float(np.std(chunk, ddof=1)) if i > 1 else 0.0
        running_mean.append(m)
        running_std.append(s)
    return running_mean, running_std

def safe_sheet_name(name: str) -> str:
    #clean excel sheet's name or not accepted
    invalid = set(r'[]:*?/\\')
    cleaned = .join(_ if c in invalid else c for c in name)
    return cleaned[:31]

def iter_time_columns(time_columns):
    #Iterator on time_columns:
    #if value= string, then single column (label 'L1')
    #if value = dict with 'column' key, then single column (label 'L1')
    #if value = dict with 'column, column' key, then iterates all of them (2D case)
    #Returns tuple (label, col_letter)

    for t, v in time_columns.items():
        if isinstance(v, str):
            yield t, L1, v
        elif isinstance(v, dict):
            # single column case {'column': 'F'}
            if 'column' in v and isinstance(v['column'], str):
                yield t, L1, v['column']
            else:
                # double column case {'L1':'C','L2':'D'}
                for label, col in v.items():
                    yield t, str(label), str(col)
```

```
        else:
            raise ValueError(fFormat not supported for time_columns[{t}]: {v})


#main
os.makedirs(plots_dir, exist_ok=True)

all_sheets_cache = {}

with pd.ExcelWriter(output_excel, engine=openpyxl) as writer:
    for exp_name, exp_cfg in experiments_config.items():
        sheet_name = exp_cfg[sheet]

        # load file if not yet in cache
        if sheet_name not in all_sheets_cache:
            try:
                df_full = pd.read_excel(input_excel, sheet_name=sheet_name, header=
    None)
                all_sheets_cache[sheet_name] = df_full
            except Exception as e:
                print(f Can't read sheet '{sheet_name}' ({exp_name}): {e})
                continue
        else:
            df_full = all_sheets_cache[sheet_name]

        # for each time/column (double or single)
        for time_key, label, col_letter in iter_time_columns(exp_cfg[time_columns]):
            try:
                col_idx = col_letter_to_index(col_letter)
                #takes column, converts to numeric and drops NaN
                col_series = to_numeric_series(df_full.iloc[:, col_idx]).dropna()
                values = col_series.to_numpy(dtype=float)

                if values.size == 0:
                    print(f!! {exp_name} Sample{time_key} {label} ({col_letter}): no
    value
                    continue

                # running stats
                r_mean, r_std = running_stats(values)

                # dataframe risultati
                out_df = pd.DataFrame({
                    Index: np.arange(1, values.size + 1, dtype=int),
                    Original values: values,
                    Running Average: r_mean,
                    Running StdDev: r_std
```

```
                })

                # sheet name info
                sheet_title = f{exp_name}_Sample{time_key}_{label}
                out_sheet = safe_sheet_name(sheet_title)
                out_df.to_excel(writer, sheet_name=out_sheet, index=False)

                # grafico scatter: blue (mean) and red (std)
                fig, ax1 = plt.subplots()
                x = out_df[Index].values

                ax1.scatter(x, out_df[Running Average].values, color=blue, s=dot_size,
     label=Running Average)
                ax1.set_xlabel(Number of Crystals)
                ax1.set_ylabel(Running Average, color=blue)
                ax1.tick_params(axis='y', labelcolor=blue)

                ax2 = ax1.twinx()
                ax2.scatter(x, out_df[Running StdDev].values, color=red, s=dot_size,
    label=Running StdDev)
                ax2.set_ylabel(Running StdDev, color=red)
                ax2.tick_params(axis='y', labelcolor=red)

                # Legend
                handles1, labels1 = ax1.get_legend_handles_labels()
                handles2, labels2 = ax2.get_legend_handles_labels()
                fig.legend(handles1 + handles2, labels1 + labels2,
                           loc=upper center, bbox_to_anchor=(0.5, 1.05), ncol=2)

                plt.title(sheet_title)
                fig.tight_layout()

                # save image
                safe_png = safe_sheet_name(sheet_title) + .png
                png_path = os.path.join(plots_dir, safe_png)
                plt.savefig(png_path, dpi=150)
                plt.close()

                print(f{sheet_title}: {values.size} values. Sheet: {out_sheet} | Graph
    : {png_path})

            except Exception as e:
                print(f Error on {exp_name} Sample{time_key} {label} ({col_letter}): {
    e})

print(f\n Stats saved in: {output_excel})
print(f Graphs saved in : {plots_dir})
```

Code A.2 produces the RDF graphs with and without error bars visible in Appendix C.

Listing A.2: RDF Solvent with and without error bars

```
#Path for RDF Files
filepath = rFileX

group_size = 250

# Initializations of lists to collect all the samples
r_data  = {i: [] for i in range(1, group_size+1)}
gr_data = {i: [] for i in range(1, group_size+1)}

# File reading
with open(filepath, r) as f:
    for line in f:
        line = line.strip()
        if not line or line.startswith(#):
            continue
        parts = line.split()
        if len(parts) < 3:
            continue
        try:
            iter_id = int(float(parts[0]))
            if 1 <= iter_id <= group_size:
                r_data[iter_id].append(float(parts[1]))
                gr_data[iter_id].append(float(parts[2]))
        except ValueError:
            continue

# Evaluations of mean values and standard deviation
r_mean  = np.zeros(group_size)
gr_mean = np.zeros(group_size)
r_std   = np.zeros(group_size)
gr_std  = np.zeros(group_size)

for i in range(1, group_size+1):
    arr_r  = np.array(r_data[i])
    arr_gr = np.array(gr_data[i])
    if arr_r.size == 0 or arr_gr.size == 0:
        raise RuntimeError(fNo data for iteration {i})
    r_mean[i-1]  = arr_r.mean()
    gr_mean[i-1] = arr_gr.mean()
    r_std[i-1]   = arr_r.std(ddof=1)
    gr_std[i-1]  = arr_gr.std(ddof=1)

```

```
# Clean graph with line
plt.figure(figsize=(8, 5))
plt.plot(r_mean, gr_mean, '-', lw=1.5)
plt.xlabel(r ())
plt.ylabel(g(r))
plt.title(RDF Acetone OObc) #Label needs to be changed depending on the dataset
plt.grid(True)
plt.tight_layout()

# 2) Graph with errorbar
plt.figure(figsize=(8, 5))
plt.errorbar(
    r_mean, gr_mean,
    xerr=r_std, yerr=gr_std,
    fmt='-o', ecolor='gray', elinewidth=1, capsize=2,
    markersize=4, lw=1
)
plt.xlabel(r ())
plt.ylabel(g(r))
plt.title(RDF Acetone OO bc with error bars)  #Label needs to be changed depending on
    the #dataset
plt.grid(True)
plt.tight_layout()

plt.show()
```

Code A.3 was employed to plot all the RDF of a given solvent in a single graph, which is visible in Section 4.2.1.

Listing A.3: RDF Solvent with multiple input files

```
#The code uses 6 RDF input files
filepaths = [
    rFile1,
    rFile2,
    rFile3,
    rFile4,
    rFile5,
   rFile6,
]

# Peronalized label for the legend, (same order as the input files)
labels = [
    C1C1,
    C2C2,
    NC1,
    NC2,
```

```
    NH,
    NN,
]

#check if number of label == number inputfiles
if len(labels) != len(filepaths):
    raise ValueError(The label number needs to be the same as the number of files)

group_size = 250

def compute_rdf_means(filepath, group_size):

    Reads one RDF file and evaluates the average r and g(r) for each iteration
    Returns two lists: r_mean and gr_mean

    sum_r  = [0.0] * group_size
    sum_gr = [0.0] * group_size
    count  = [0] * group_size

    with open(filepath, 'r') as f:
        for line in f:
            line = line.strip()
            if not line or line.startswith('#'):
                continue
            parts = line.split()
            if len(parts) < 3:
                continue
            try:
                iter_id = int(float(parts[0]))
                if 1 <= iter_id <= group_size:
                    idx    = iter_id - 1
                    r_val  = float(parts[1])
                    gr_val = float(parts[2])
                    sum_r[idx]  += r_val
                    sum_gr[idx] += gr_val
                    count[idx]  += 1
            except ValueError:
                continue

    r_mean  = []
    gr_mean = []
    for i in range(group_size):
        if count[i] == 0:
            raise RuntimeError(fNo data found for iteration {i+1} in file {os.path.
    basename(filepath)})
        r_mean.append(sum_r[i] / count[i])
        gr_mean.append(sum_gr[i] / count[i])
```

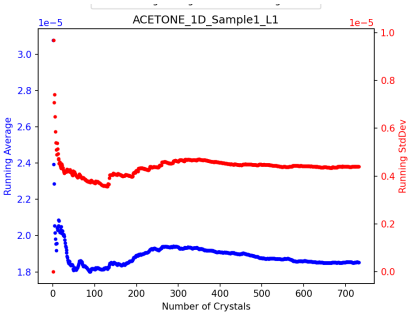
```
63
64      return r_mean, gr_mean
65
66  # Plot on the same graph all the files
67  plt.figure(figsize=(8, 5))
68  for filepath, label in zip(filepaths, labels):
69      r_mean, gr_mean = compute_rdf_means(filepath, group_size)
70      plt.plot(r_mean, gr_mean, lw=1.5, label=label)
71
72  plt.xlabel(r ())
73  plt.ylabel(g(r))
74  plt.title(RDF Acetonitrile) #Label needs to be changed depending on the dataset
75  plt.legend(title=)
76  plt.grid(True)
77  plt.tight_layout()
78  plt.show()
```

Appendix B

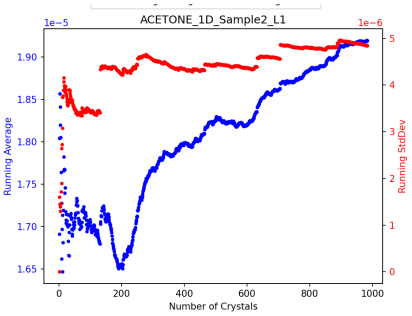
Statistical Analysis

B.1 Plots

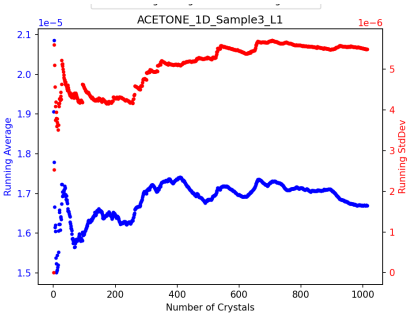
This section of Appendix B presents the plots of the running time averages and running standard deviations for each sample in the three solvents. Both one-dimensional analyses, based on the equivalent diameter (L_1 in the 1D plots), and two-dimensional analyses, distinguishing between L_1 and L_2, are reported.



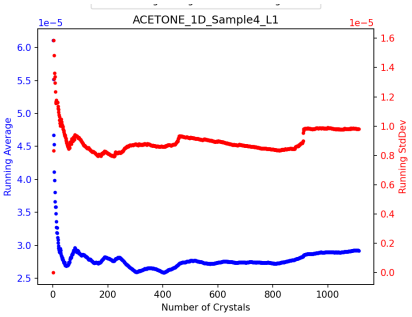
(a) Sample 1 - 1 hour



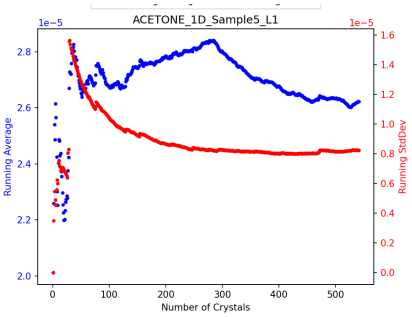
(b) Sample 2 - 2 hours



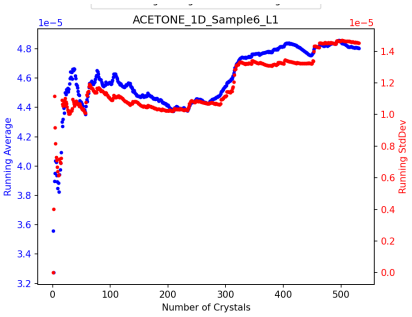
(c) Sample 3 - 3 hours



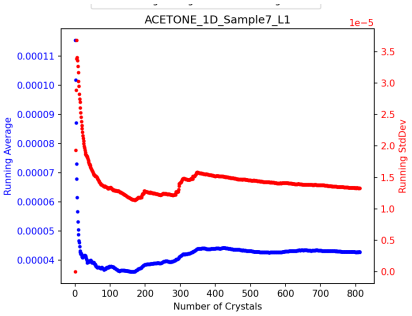
(d) Sample 4 - 4 hours



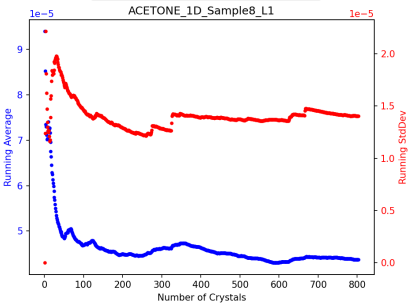
(e) Sample 5 - 5 hours



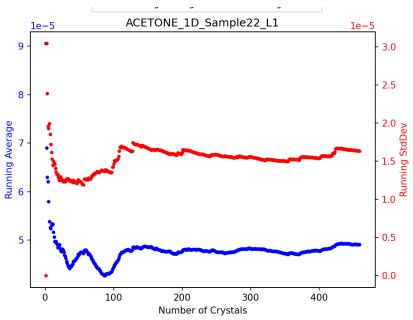
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

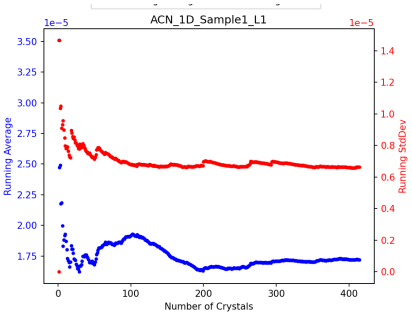


(h) Sample 8 - 8 hours

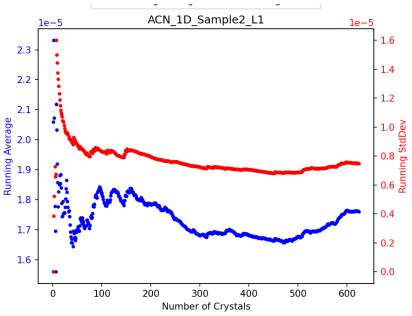


(i) Sample 9 - 22 hours

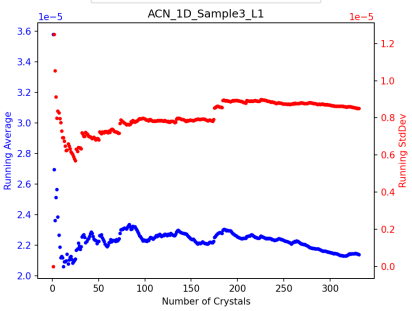
Figure B.1: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONE, 1D - analysis



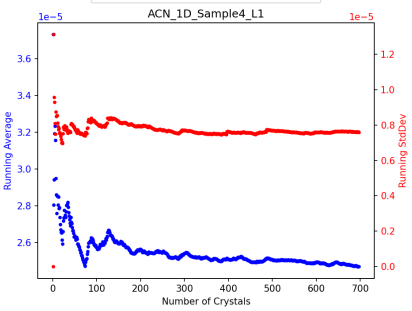
(a) Sample 1 - 1 hour



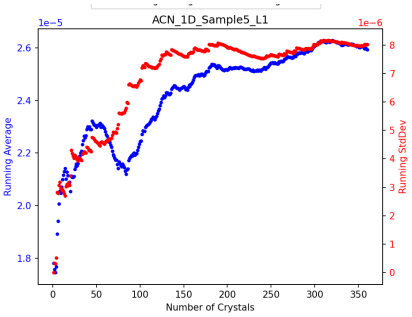
(b) Sample 2 - 2 hours



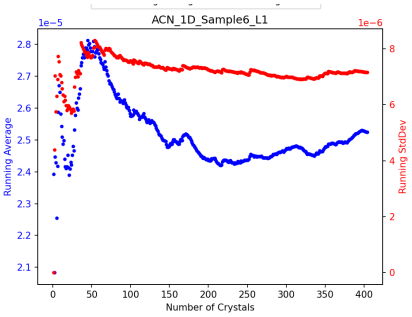
(c) Sample 3 - 3 hours



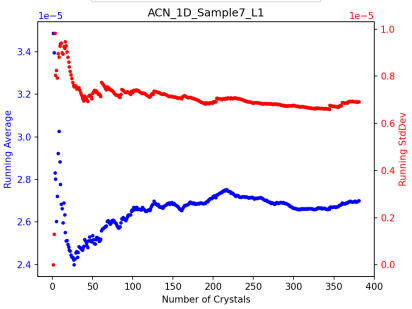
(d) Sample 4 - 4 hours



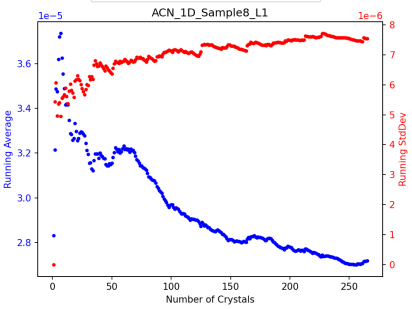
(e) Sample 5 - 5 hours



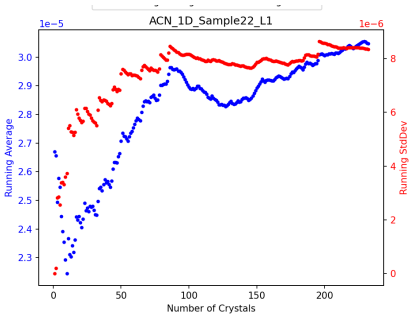
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

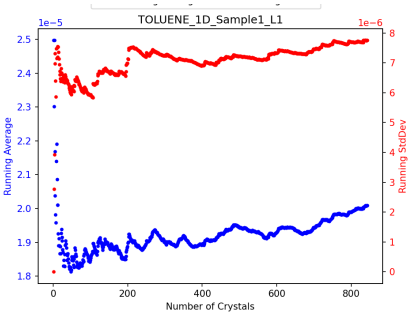


(h) Sample 8 - 8 hours

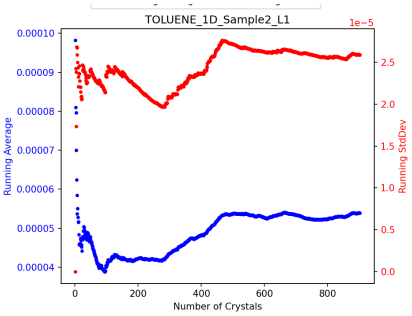


(i) Sample 9 - 22 hours

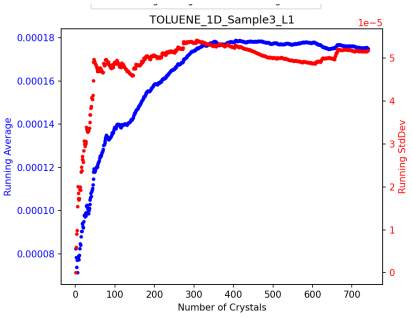
Figure B.2: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONITRILE, 1D - analysis



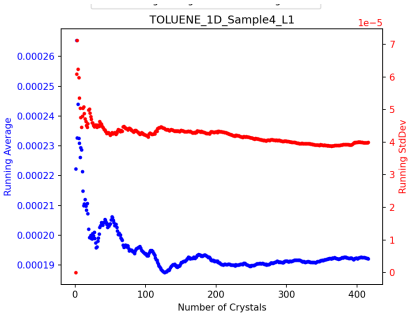
(a) Sample 1 - 1 hour



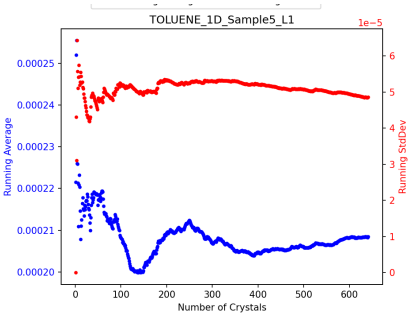
(b) Sample 2 - 2 hours



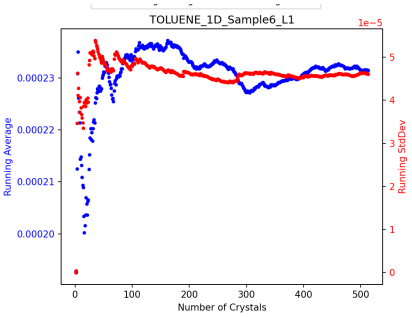
(c) Sample 3 - 3 hours



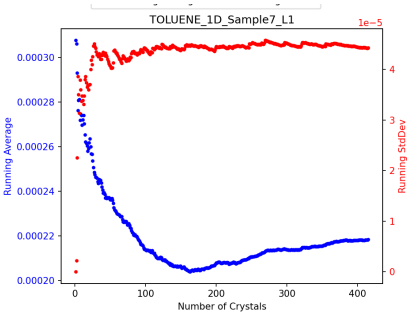
(d) Sample 4 - 4 hours



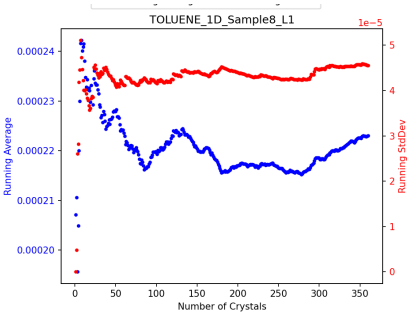
(e) Sample 5 - 5 hours



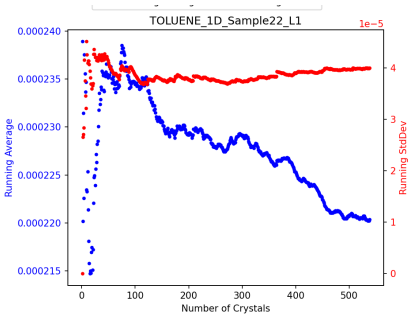
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

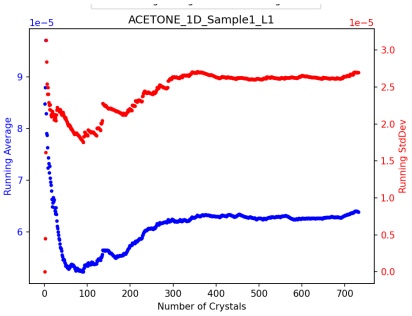


(h) Sample 8 - 8 hours

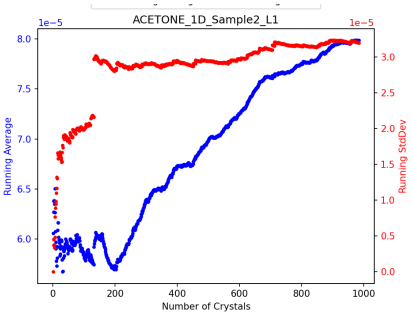


(i) Sample 9 - 22 hours

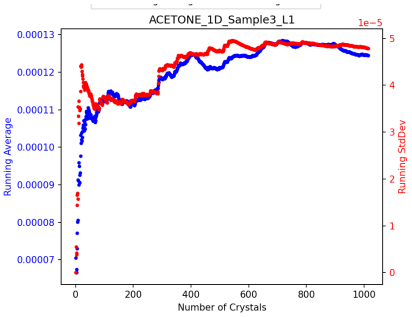
Figure B.3: Plots of Running Average Time and Standard deviation of xanthone's crystals size in TOLUENE, 1D - analysis



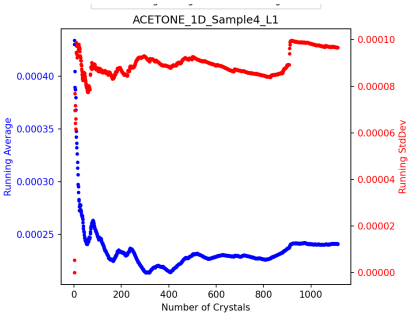
(a) Sample 1 - 1 hour



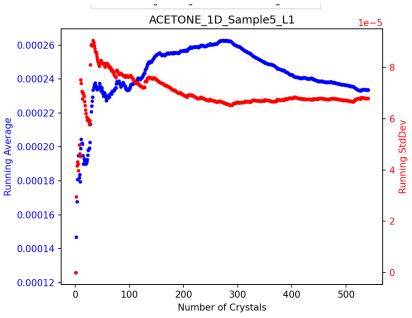
(b) Sample 2 - 2 hours



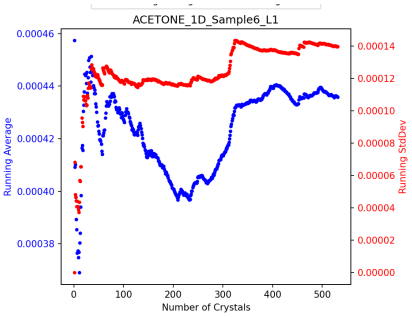
(c) Sample 3 - 3 hours



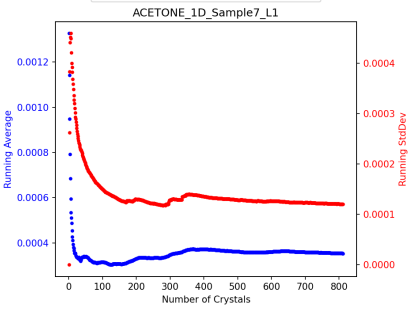
(d) Sample 4 - 4 hours



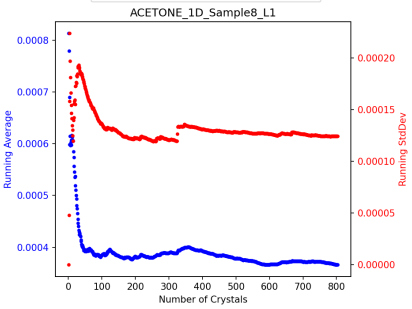
(e) Sample 5 - 5 hours



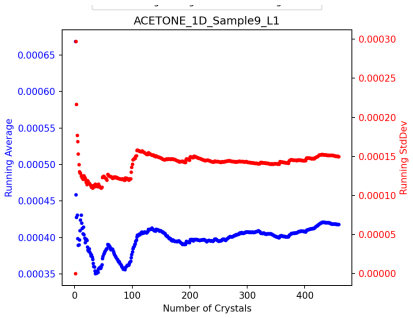
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

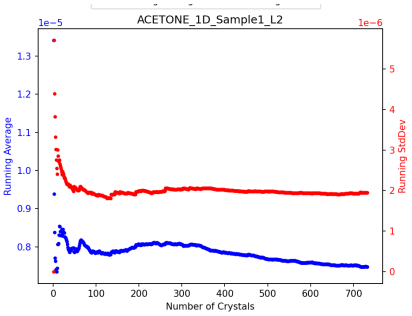


(h) Sample 8 - 8 hours

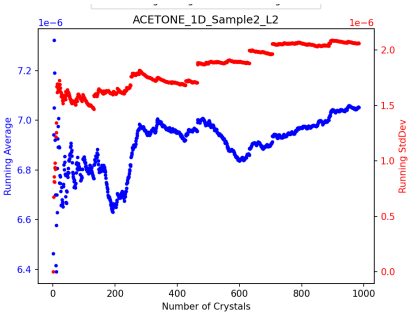


(i) Sample 9 - 22 hours

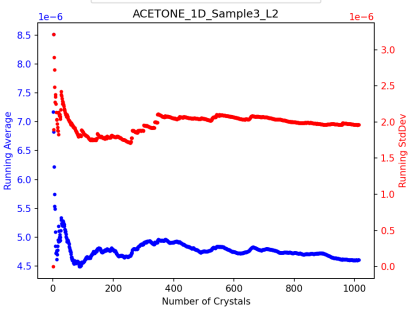
Figure B.4: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONE L_1, 2D - analysis



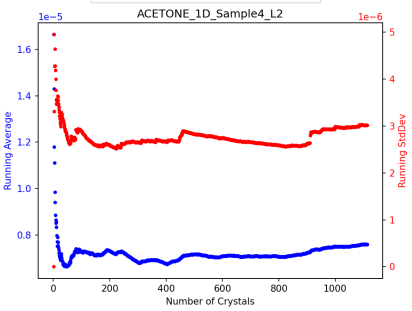
(a) Sample 1 - 1 hour



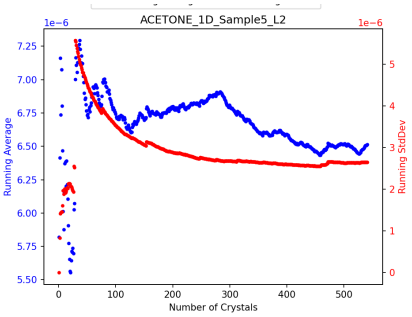
(b) Sample 2 - 2 hours



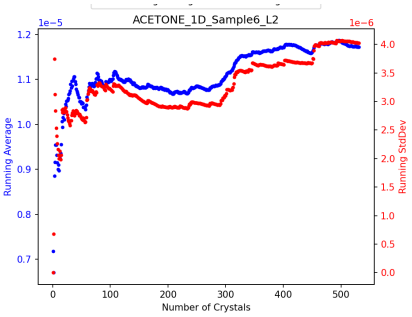
(c) Sample 3 - 3 hours



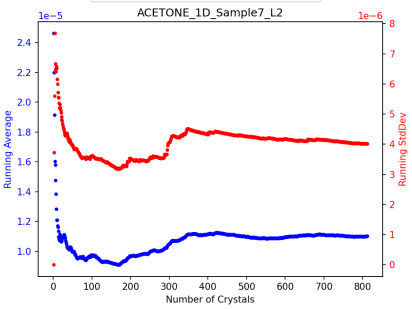
(d) Sample 4 - 4 hours



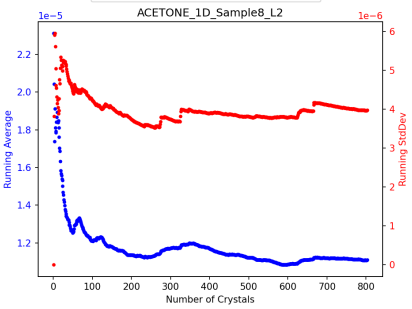
(e) Sample 5 - 5 hours



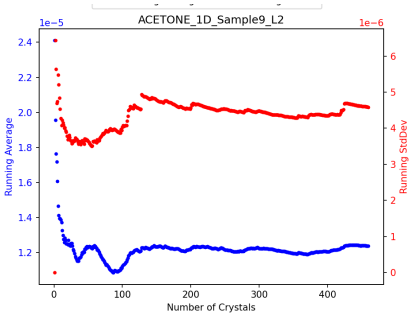
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

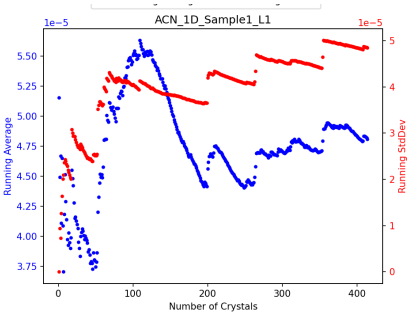


(h) Sample 8 - 8 hours

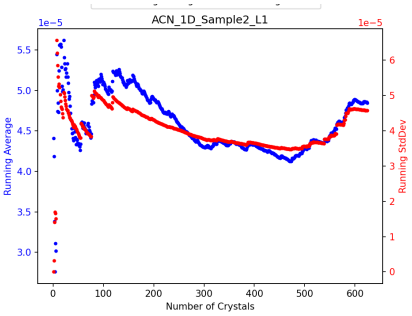


(i) Sample 9 - 22 hours

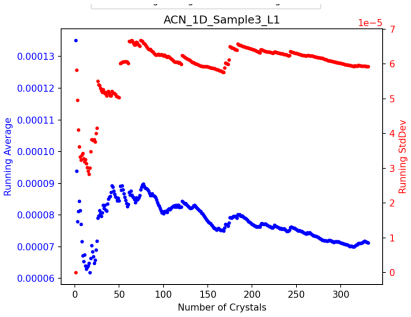
Figure B.5: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONE L_2, 2D - analysis



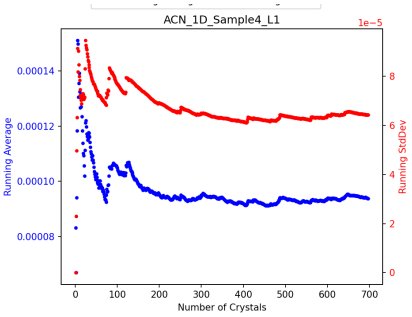
(a) Sample 1 - 1 hour



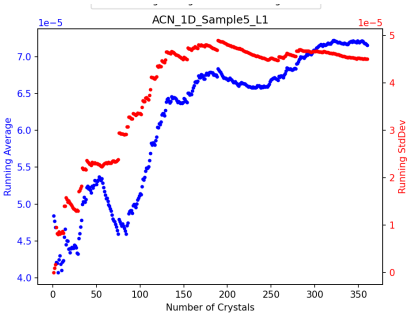
(b) Sample 2 - 2 hours



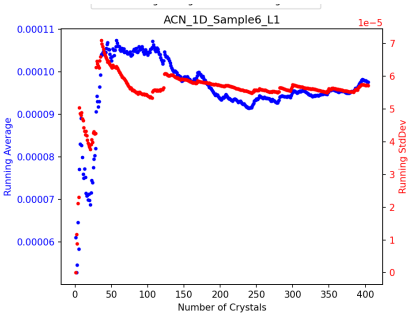
(c) Sample 3 - 3 hours



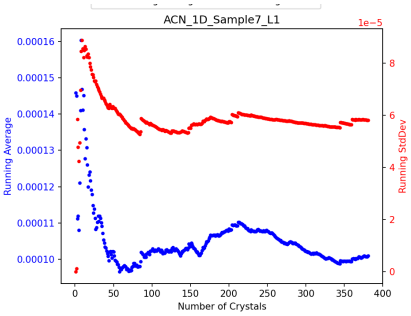
(d) Sample 4 - 4 hours



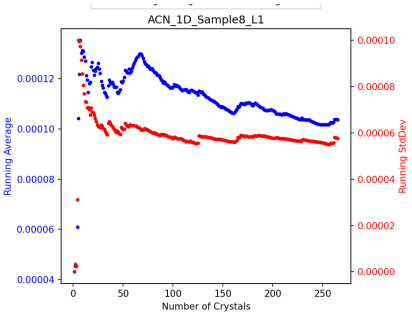
(e) Sample 5 - 5 hours



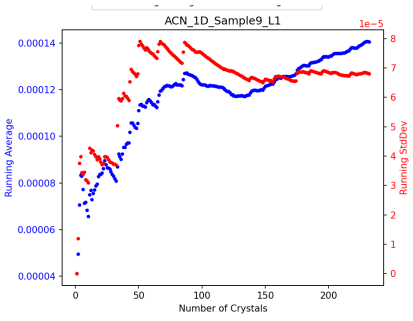
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

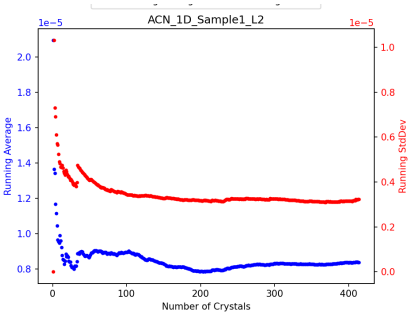


(h) Sample 8 - 8 hours

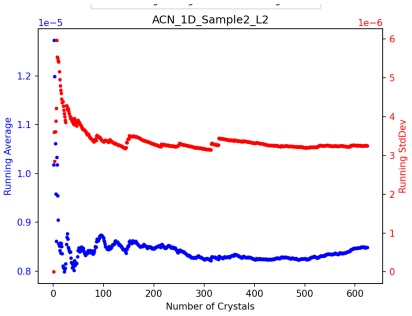


(i) Sample 9 - 22 hours

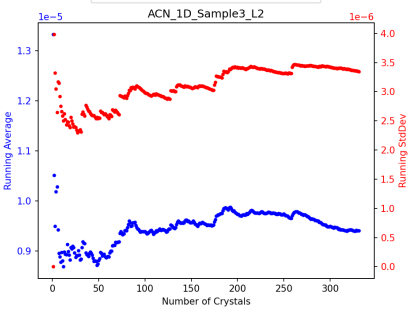
Figure B.6: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONITRILE L_1, 2D - analysis[106]



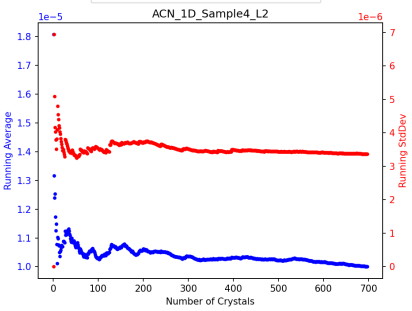
(a) Sample 1 - 1 hour



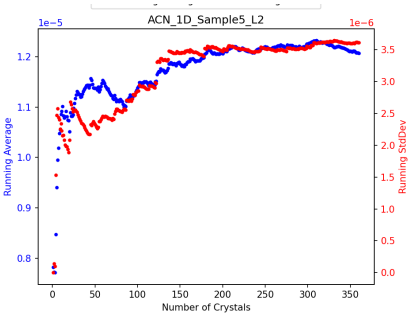
(b) Sample 2 - 2 hours



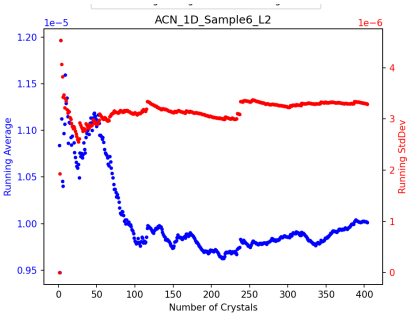
(c) Sample 3 - 3 hours



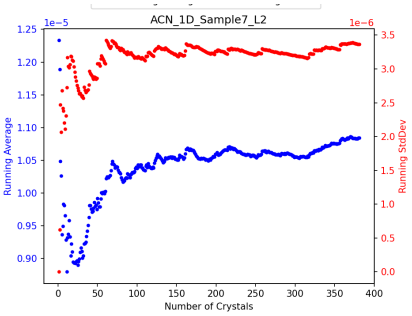
(d) Sample 4 - 4 hours



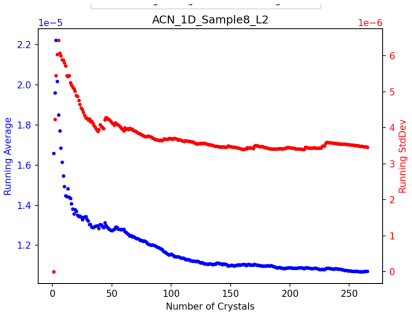
(e) Sample 5 - 5 hours



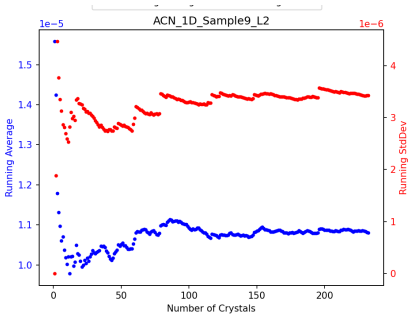
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

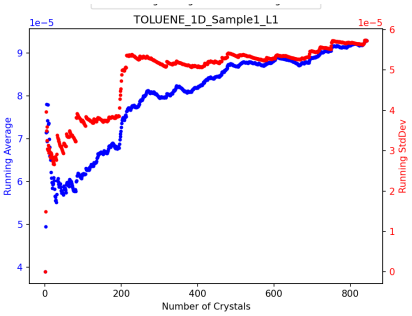


(h) Sample 8 - 8 hours

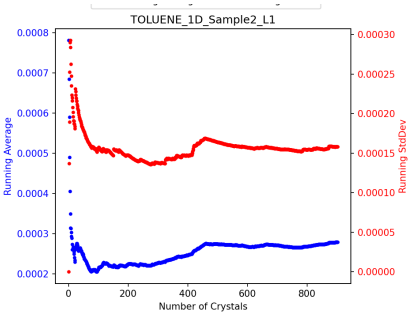


(i) Sample 9 - 22 hours

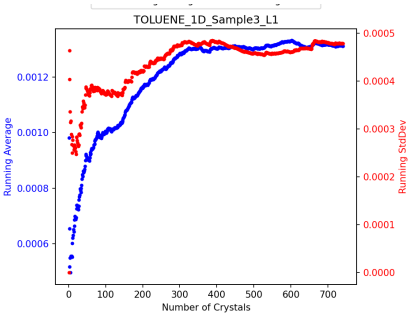
Figure B.7: Plots of Running Average Time and Standard deviation of xanthone's crystals size in ACETONITRILE L_2, 2D - analysis[107]



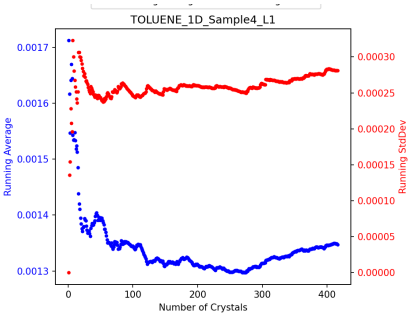
(a) Sample 1 - 1 hour



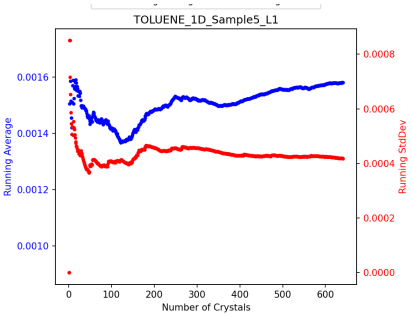
(b) Sample 2 - 2 hours



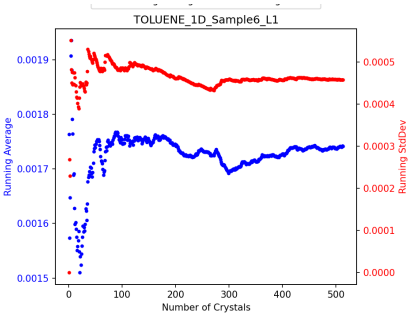
(c) Sample 3 - 3 hours



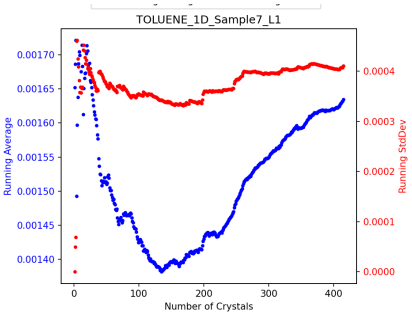
(d) Sample 4 - 4 hours



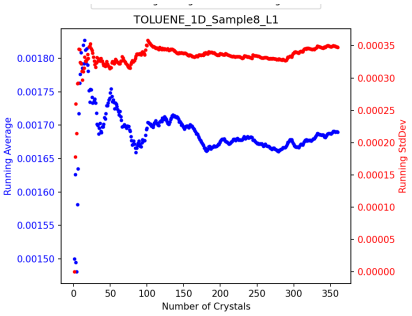
(e) Sample 5 - 5 hours



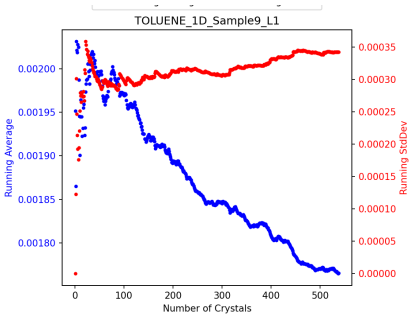
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours

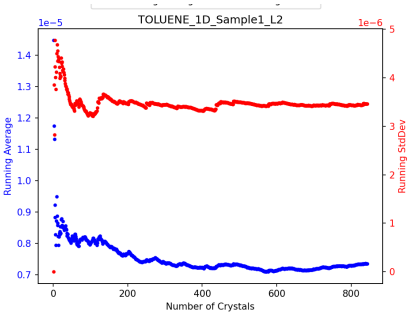


(h) Sample 8 - 8 hours

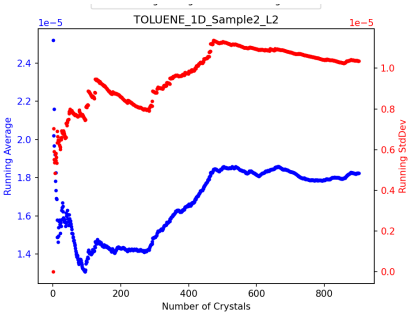


(i) Sample 9 - 22 hours

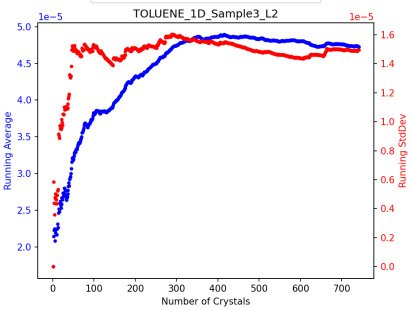
Figure B.8: Plots of Running Average Time and Standard deviation of xanthone's crystals size in TOLUENE L_1, 2D - analysis



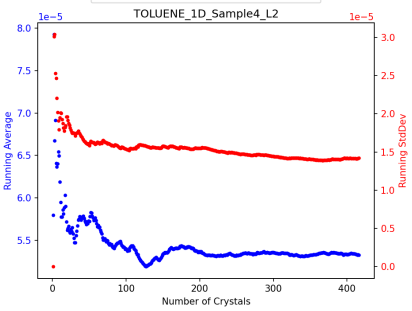
(a) Sample 1 - 1 hour



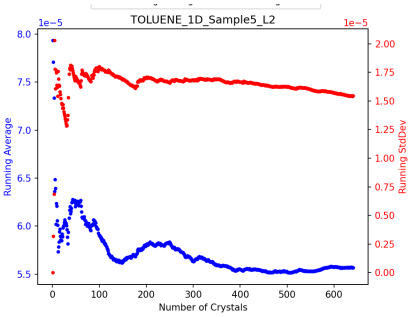
(b) Sample 2 - 2 hours



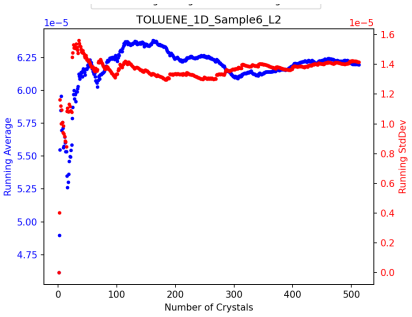
(c) Sample 3 - 3 hours



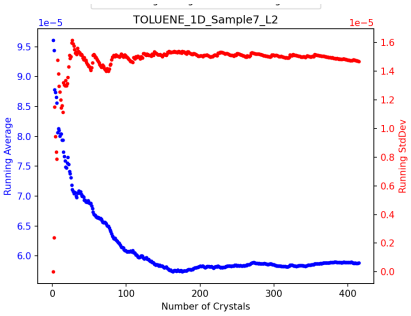
(d) Sample 4 - 4 hours



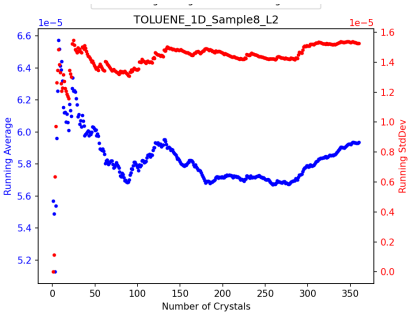
(e) Sample 5 - 5 hours



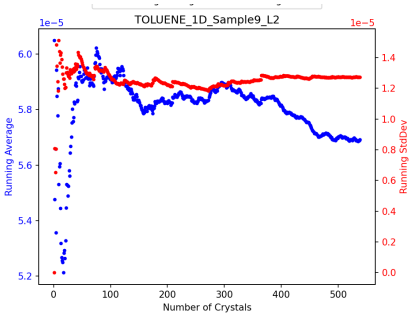
(f) Sample 6 - 6 hours



(g) Sample 7 - 7 hours



(h) Sample 8 - 8 hours



(i) Sample 9 - 22 hours

Figure B.9: Plots of Running Average Time and Standard deviation of xanthone's crystals size in TOLUENE L_2, 2D - analysis

B.2 Tables

This section of the appendix contains the tables relative to the average projected area, average projected volume and their standard deviation for xanthone's crystals in the three solvents, both mono dimensional and bi-dimensional studies, completed with confidence interval (CI) and relative confidence interval (RCI) estimation.

Table B.1: 1D Equivalent Area: mean ($\bar{A}$), standard deviation (S), confidence intervals (CI_W) for each solvent and sample

Solvent	**Sample**	$\bar{A}$ (μ**m**^{2})	S (μ**m**^{2})	CI_{W90} (μ**m**^{2})	CI_{W99} (μ**m**^{2})	RCI_{90}	RCI_{99}
	1 (1 h)	1.97E+02	1.30E+02	3.06E-05	4.78E-05	16%	24%
	2 (2 h)	3.62E+02	2.10E+02	4.35E-05	6.81E-05	12%	19%
	3 (3 h)	3.92E+02	4.00E+01	8.32E-06	1.30E-05	2%	3%
	4 (4 h)	3.09E+02	1.10E+03	2.21E-04	3.47E-04	72%	112%
Acetone	5 (5 h)	9.50E+02	5.20E+02	1.48E-04	2.32E-04	16%	24%
	6 (6 h)	7.56E+02	3.40E+03	9.72E-04	1.52E-03	129%	202%
	7 (7 h)	2.52E+03	1.20E+03	2.85E-04	4.46E-04	11%	18%
	8 (8 h)	2.01E+03	1.50E+03	3.55E-04	5.55E-04	18%	28%
	9 (22 h)	2.11E+03	2.20E+03	6.79E-04	1.06E-03	32%	50%
	1 (1 h)	1.97E+02	6.10E+01	1.98E-05	3.11E-05	10%	16%
	2 (2 h)	3.39E+02	3.00E+02	7.80E-05	1.22E-04	23%	36%
	3 (3 h)	3.66E+02	5.80E+02	2.09E-04	3.28E-04	57%	90%
	4 (4 h)	5.31E+02	6.10E+02	1.52E-04	2.37E-04	29%	45%
Acetonitrile	5 (5 h)	6.67E+02	4.60E+02	1.59E-04	2.49E-04	24%	37%
	6 (6 h)	7.37E+02	3.10E+02	1.01E-04	1.59E-04	14%	22%
	7 (7 h)	6.88E+02	4.40E+02	1.50E-04	2.35E-04	22%	34%
	8 (8 h)	7.77E+02	4.40E+02	1.76E-04	2.75E-04	23%	35%
	9 (22 h)	7.95E+02	4.60E+02	1.98E-04	3.10E-04	25%	39%
	1 (1 h)	1.97E+02	9.30E+01	2.12E-05	3.32E-05	11%	17%
	2 (2 h)	4.64E+02	7.80E+03	1.71E-03	2.68E-03	369%	577%
	3 (3 h)	3.42E+03	1.00E+05	2.51E-02	3.93E-02	734%	1149%
	4 (4 h)	3.33E+04	2.20E+04	7.12E-03	1.11E-02	21%	33%
Toluene	5 (5 h)	3.85E+04	2.10E+04	5.33E-03	8.35E-03	14%	22%
	6 (6 h)	4.58E+04	2.70E+04	7.77E-03	1.22E-02	17%	27%
	7 (7 h)	5.57E+04	1.80E+04	5.95E-03	9.32E-03	11%	17%
	8 (8 h)	4.96E+04	2.00E+04	6.77E-03	1.06E-02	14%	21%

Solvent	**Sample**	$\bar{A}$ (μ**m**^{2})	S (μ**m**^{2})	CI_{W90} (μ**m**^{2})	CI_{W99} (μ**m**^{2})	RCI_{90}	RCI_{99}
	9 (22 h)	5.18E+04	1.50E+04	4.39E-03	6.87E-03	8%	13%

Table B.2: 2D Area L_1: mean ($\bar{A}_1$), standard deviation (S), confidence intervals (CI_W) e RCI for each solvent and sample

Solvent	**Sample**	$\bar{A}$ (μ**m**^{2})	S (μ**m**^{2})	CI_{W90} (μ**m**^{2})	CI_{W99} (μ**m**^{2})	RCI_{90}	RCI_{99}
	1 (1 h)	4.81E+03	4.32E+03	1.05E-03	1.65E-03	22%	34%
	2 (2 h)	7.40E+03	6.58E+03	1.38E-03	2.16E-03	19%	29%
	3 (3 h)	1.78E+04	1.38E+04	2.85E-03	4.46E-03	16%	25%
	4 (4 h)	6.73E+04	6.68E+04	1.32E-02	2.06E-02	20%	31%
Acetone	5 (5 h)	5.92E+04	3.39E+04	9.58E-03	1.50E-02	16%	25%
	6 (6 h)	2.09E+05	1.49E+05	4.26E-02	6.67E-02	20%	32%
	7 (7 h)	1.39E+05	1.16E+05	2.68E-02	4.19E-02	19%	30%
	8 (8 h)	1.50E+05	1.15E+05	2.68E-02	4.20E-02	18%	28%
	9 (22 h)	1.97E+05	1.46E+05	4.50E-02	7.04E-02	23%	36%
	1 (1 h)	4.66E+03	1.40E+04	4.52E-03	7.08E-03	97%	152%
	2 (2 h)	4.44E+03	1.12E+04	2.96E-03	4.64E-03	67%	105%
	3 (3 h)	8.58E+03	1.63E+04	5.88E-03	9.20E-03	68%	107%
	4 (4 h)	1.29E+04	2.01E+04	5.00E-03	7.83E-03	39%	61%
Acetonitrile	5 (5 h)	7.14E+03	9.62E+03	3.34E-03	5.22E-03	47%	73%
	6 (6 h)	1.28E+04	1.59E+04	5.22E-03	8.17E-03	41%	64%
	7 (7 h)	1.36E+04	1.88E+04	6.33E-03	9.91E-03	47%	73%
	8 (8 h)	1.41E+04	1.75E+04	7.07E-03	1.11E-02	50%	79%
	9 (22 h)	2.44E+04	2.46E+04	1.06E-02	1.66E-02	44%	68%
	1 (1 h)	1.19E+04	1.63E+04	3.68E-03	5.77E-03	31%	48%
	2 (2 h)	1.03E+05	1.27E+05	2.77E-02	4.35E-02	27%	42%
	3 (3 h)	1.95E+06	1.37E+06	3.31E-01	5.18E-01	17%	27%
	4 (4 h)	1.89E+06	7.90E+05	2.55E-01	3.99E-01	13%	21%
Toluene	5 (5 h)	2.67E+06	1.35E+06	3.52E-01	5.51E-01	13%	21%
	6 (6 h)	3.24E+06	1.63E+06	4.73E-01	7.40E-01	15%	23%
	7 (7 h)	2.84E+06	1.37E+06	4.42E-01	6.92E-01	16%	24%
	8 (8 h)	2.98E+06	1.20E+06	4.16E-01	6.51E-01	14%	22%
	9 (22 h)	3.23E+06	1.21E+06	3.42E-01	5.35E-01	11%	17%

Table B.3: 2D Area L_2: mean ($\bar{A}_2$), standard deviation (S), confidence intervals (CI_W) e RCI pfor each solvent and sample

Solvent	**Sample**	$\bar{A}$ (μ**m**^{2})	S (μ**m**^{2})	CI_{W90} (μ**m**^{2})	CI_{W99} (μ**m**^{2})	RCI_{90}	RCI_{99}
	1 (1 h)	5.97E+01	3.24E+01	7.88E-06	1.23E-05	13%	21%
	2 (2 h)	5.40E+01	4.03E+01	8.45E-06	1.32E-05	16%	25%
	3 (3 h)	2.51E+01	2.56E+01	5.29E-06	8.28E-06	21%	33%
	4 (4 h)	6.67E+01	6.75E+01	1.33E-05	2.08E-05	20%	31%
Acetone	5 (5 h)	4.94E+01	5.87E+01	1.66E-05	2.60E-05	34%	53%
	6 (6 h)	1.54E+02	1.22E+02	3.48E-05	5.45E-05	23%	36%
	7 (7 h)	1.37E+02	1.11E+02	2.55E-05	4.00E-05	19%	29%
	8 (8 h)	1.39E+02	1.20E+02	2.78E-05	4.36E-05	20%	31%
	9 (22 h)	1.75E+02	1.45E+02	4.45E-05	6.96E-05	25%	40%
	1 (1 h)	8.05E+01	6.81E+01	2.20E-05	3.45E-05	27%	43%
	2 (2 h)	8.25E+01	7.21E+01	1.90E-05	2.97E-05	23%	36%
	3 (3 h)	9.97E+01	7.66E+01	2.77E-05	4.34E-05	28%	44%
	4 (4 h)	1.11E+02	8.09E+01	2.02E-05	3.16E-05	18%	28%
Acetonitrile	5 (5 h)	1.59E+02	1.00E+02	3.47E-05	5.44E-05	22%	34%
	6 (6 h)	1.11E+02	8.19E+01	2.68E-05	4.20E-05	24%	38%
	7 (7 h)	1.29E+02	8.47E+01	2.86E-05	4.47E-05	22%	35%
	8 (8 h)	1.27E+02	9.28E+01	3.75E-05	5.87E-05	30%	46%
	9 (22 h)	1.28E+02	9.04E+01	3.91E-05	6.12E-05	30%	48%
	1 (1 h)	6.61E+01	6.73E+01	1.52E-05	2.39E-05	23%	36%
	2 (2 h)	4.40E+02	5.55E+02	1.22E-04	1.90E-04	28%	43%
	3 (3 h)	2.46E+03	1.50E+03	3.63E-04	5.68E-04	15%	23%
	4 (4 h)	3.04E+03	1.57E+03	5.07E-04	7.95E-04	17%	26%
Toluene	5 (5 h)	3.34E+03	1.82E+03	4.74E-04	7.42E-04	14%	22%
	6 (6 h)	4.04E+03	1.85E+03	5.36E-04	8.39E-04	13%	21%
	7 (7 h)	3.67E+03	1.78E+03	5.76E-04	9.02E-04	16%	25%
	8 (8 h)	3.76E+03	1.96E+03	6.80E-04	1.07E-03	18%	28%
	9 (22 h)	3.40E+03	1.55E+03	4.39E-04	6.87E-04	13%	20%

Table B.4: 1D Equivalent Volume: mean ($\bar{V}$), standard deviation (S), confidence intervals (CI_W) e RCI for each solvent and sample

Solvent	**Sample**	$\bar{V}$ (μ**m**^{3})	S (μ**m**^{3})	CI_{W90} (μ**m**^{3})	CI_{W99} (μ**m**^{3})	RCI_{90}	RCI_{99}
	1 (1 h)	7.44E+03	6.10E+03	1.47E-09	2.31E-09	20%	31%
	2 (2 h)	7.49E+03	1.20E+04	2.43E-09	3.80E-09	32%	51%
	3 (3 h)	3.77E+04	7.50E+03	1.55E-09	2.43E-09	4%	6%
	4 (4 h)	6.79E+03	6.80E+04	1.33E-08	2.09E-08	196%	308%
Acetone	5 (5 h)	3.61E+04	4.30E+04	1.22E-08	1.91E-08	34%	53%
	6 (6 h)	2.48E+04	1.90E+05	5.47E-08	8.57E-08	220%	345%
	7 (7 h)	1.49E+05	1.30E+05	2.97E-08	4.65E-08	20%	31%
	8 (8 h)	1.04E+05	1.70E+05	4.04E-08	6.33E-08	39%	61%
	9 (22 h)	1.19E+05	2.10E+05	6.31E-08	9.88E-08	53%	83%
	1 (1 h)	1.55E+04	1.00E+04	3.35E-09	5.25E-09	22%	34%
	2 (2 h)	8.39E+03	1.40E+04	3.64E-09	5.70E-09	43%	68%
	3 (3 h)	8.92E+03	2.40E+04	8.83E-09	1.38E-08	99%	155%
	4 (4 h)	1.52E+04	2.10E+04	5.13E-09	8.04E-09	34%	53%
Acetonitrile	5 (5 h)	2.15E+04	2.30E+04	7.81E-09	1.22E-08	36%	57%
	6 (6 h)	3.04E+04	1.80E+04	6.05E-09	9.48E-09	20%	31%
	7 (7 h)	2.15E+04	2.00E+04	6.84E-09	1.07E-08	32%	50%
	8 (8 h)	2.38E+04	2.30E+04	9.25E-09	1.45E-08	39%	61%
	9 (22 h)	3.53E+04	3.40E+04	1.46E-08	2.29E-08	41%	65%
	1 (1 h)	1.63E+04	1.60E+04	3.72E-09	5.83E-09	23%	36%
	2 (2 h)	1.34E+04	4.40E+05	9.70E-08	1.52E-07	725%	1136%
	3 (3 h)	2.36E+05	6.10E+06	1.48E-06	2.31E-06	625%	978%
	4 (4 h)	6.97E+06	4.80E+06	1.55E-06	2.43E-06	22%	35%
Toluene	5 (5 h)	8.02E+06	7.20E+06	1.86E-06	2.91E-06	23%	36%
	6 (6 h)	1.08E+07	8.20E+06	2.40E-06	3.75E-06	22%	35%
	7 (7 h)	1.39E+07	6.90E+06	2.21E-06	3.47E-06	16%	25%
	8 (8 h)	1.25E+07	7.80E+06	2.71E-06	4.24E-06	22%	34%
	9 (22 h)	1.61E+07	6.60E+06	1.87E-06	2.92E-06	12%	18%

Table B.5: 2D Volume L_1: mean ($\bar{V_1}$), standard deviation (S), confidence intervals (CI_W) e RCI per solvente e campione

Solvent	**Sample**	$\bar{V_1}$ (μ**m**^{3})	S (μ**m**^{3})	CI_{W90} (μ**m**^{3})	CI_{W99} (μ**m**^{3})	RCI_{90}	RCI_{99}
	1 (1 h)	4.20E+05	6.26E+05	1.52E-07	2.38E-07	36%	57%
	2 (2 h)	7.91E+05	1.36E+06	2.85E-07	4.46E-07	36%	56%
	3 (3 h)	2.86E+06	3.48E+06	7.18E-07	1.12E-06	25%	39%
	4 (4 h)	2.21E+07	5.44E+07	1.07E-05	1.68E-05	49%	76%
Acetone	5 (5 h)	1.61E+07	1.42E+07	4.01E-06	6.28E-06	25%	39%
	6 (6 h)	1.12E+08	1.39E+08	3.98E-05	6.23E-05	36%	56%
	7 (7 h)	6.24E+07	1.14E+08	2.64E-05	4.13E-05	42%	66%
	8 (8 h)	6.87E+07	1.02E+08	2.38E-05	3.72E-05	35%	54%
	9 (22 h)	1.04E+08	1.23E+08	3.78E-05	5.92E-05	36%	57%
	1 (1 h)	8.35E+05	4.71E+06	1.52E-06	2.39E-06	182%	286%
	2 (2 h)	6.81E+05	3.31E+06	8.73E-07	1.37E-06	128%	201%
	3 (3 h)	1.50E+06	4.85E+06	1.75E-06	2.75E-06	117%	183%
	4 (4 h)	2.42E+06	6.49E+06	1.62E-06	2.53E-06	67%	104%
Acetonitrile	5 (5 h)	9.21E+05	2.06E+06	7.14E-07	1.12E-06	77%	121%
	6 (6 h)	2.12E+06	4.40E+06	1.44E-06	2.26E-06	68%	107%
	7 (7 h)	2.39E+06	6.25E+06	2.11E-06	3.30E-06	88%	138%
	8 (8 h)	2.42E+06	5.14E+06	2.08E-06	3.25E-06	86%	134%
	9 (22 h)	5.04E+06	8.11E+06	3.51E-06	5.49E-06	70%	109%
	1 (1 h)	1.99E+06	4.57E+06	1.04E-06	1.62E-06	52%	82%
	2 (2 h)	4.79E+07	9.82E+07	2.15E-05	3.37E-05	45%	70%
	3 (3 h)	3.19E+09	3.41E+09	8.25E-04	1.29E-03	26%	41%
	4 (4 h)	2.77E+09	1.77E+09	5.71E-04	8.94E-04	21%	32%
Toluene	5 (5 h)	4.79E+09	3.57E+09	9.28E-04	1.45E-03	19%	30%
	6 (6 h)	6.38E+09	4.69E+09	1.36E-03	2.13E-03	21%	33%
	7 (7 h)	5.19E+09	3.71E+09	1.20E-03	1.88E-03	23%	36%
	8 (8 h)	5.44E+09	3.25E+09	1.13E-03	1.77E-03	21%	32%
	9 (22 h)	6.12E+09	3.34E+09	9.48E-04	1.48E-03	15%	24%

Table B.6: 2D Volume L_2: mean ($\bar{V_2}$), standard deviation (S), confidence intervals (CI_W) e RCI per solvente e campione

Solvent	**Sample**	$\bar{V}$ (μ**m**^{3})	S (μ**m**^{3})	CI_{W90} (μ**m**^{3})	CI_{W99} (μ**m**^{3})	RCI_{90}	RCI_{99}
	1 (1 h)	5.08E+02	4.53E+02	4.53E-16	1.73E-10	22%	34%
	2 (2 h)	4.59E+02	8.02E+02	3.75E-16	2.63E-10	37%	57%
	3 (3 h)	1.63E+02	3.29E+02	1.29E-16	1.07E-10	42%	65%
	4 (4 h)	6.96E+02	1.68E+03	5.31E-16	5.18E-10	48%	74%
Acetone	5 (5 h)	4.61E+02	1.62E+03	2.32E-16	7.17E-10	99%	155%
	6 (6 h)	2.27E+03	3.33E+03	1.80E-15	1.49E-09	42%	65%
	7 (7 h)	1.94E+03	2.68E+03	1.64E-15	9.70E-10	32%	50%
	8 (8 h)	2.00E+03	3.54E+03	1.58E-15	1.29E-09	41%	64%
	9 (22 h)	2.80E+03	4.23E+03	2.15E-15	2.04E-09	46%	73%
	1 (1 h)	8.86E+02	1.29E+03	6.77E-16	6.55E-10	47%	74%
	2 (2 h)	9.24E+02	1.53E+03	7.22E-16	6.33E-10	44%	69%
	3 (3 h)	1.19E+03	1.55E+03	9.07E-16	8.76E-10	47%	74%
	4 (4 h)	1.38E+03	1.67E+03	1.17E-15	6.51E-10	30%	47%
Acetonitrile	5 (5 h)	2.27E+03	2.36E+03	1.86E-15	1.28E-09	36%	56%
	6 (6 h)	1.37E+03	1.85E+03	1.07E-15	9.49E-10	44%	69%
	7 (7 h)	1.68E+03	1.81E+03	1.37E-15	9.58E-10	36%	57%
	8 (8 h)	1.67E+03	2.17E+03	1.23E-15	1.37E-09	52%	82%
	9 (22 h)	1.69E+03	2.12E+03	1.23E-15	1.44E-09	54%	85%
	1 (1 h)	7.11E+02	1.22E+03	5.73E-16	4.33E-10	39%	61%
	2 (2 h)	1.34E+04	3.08E+04	1.01E-14	1.06E-08	50%	79%
	3 (3 h)	1.38E+05	1.30E+05	1.22E-13	4.92E-08	23%	36%
	4 (4 h)	1.84E+05	1.41E+05	1.61E-13	7.14E-08	25%	39%
Toluene	5 (5 h)	2.14E+05	1.78E+05	1.91E-13	7.25E-08	22%	34%
	6 (6 h)	2.76E+05	1.95E+05	2.48E-13	8.87E-08	21%	32%
	7 (7 h)	2.42E+05	1.76E+05	2.13E-13	8.92E-08	24%	37%
	8 (8 h)	2.53E+05	2.04E+05	2.17E-13	1.11E-07	28%	44%
	9 (22 h)	2.13E+05	1.52E+05	1.92E-13	6.74E-08	20%	32%

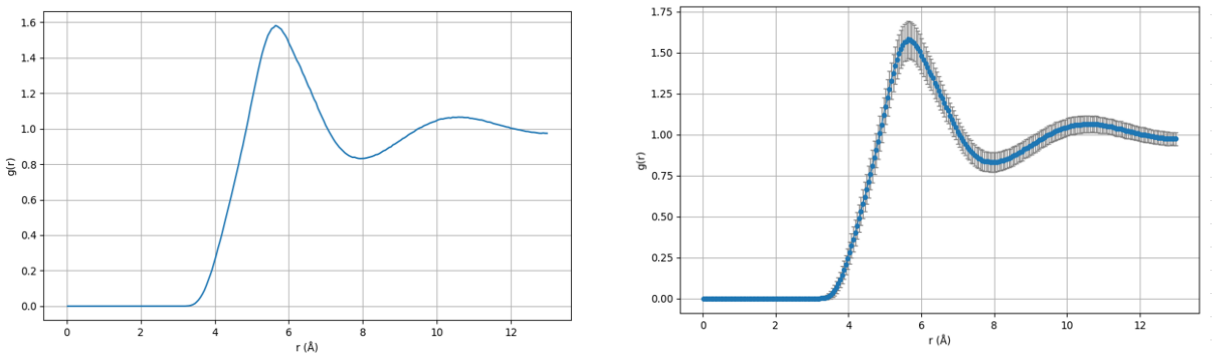
Appendix C

RDF Analysis of the three solvents

This appendix reports the RDF functions for the main atom–pairs of each solvent, alongside the corresponding error–bar plots. For each case, the first peak position, peak width, and the relevant literature reference are provided, followed by an explicit evaluation of the agreement between the simulated results and the reference data.

Since the RDF analysis was performed on the solvent in bulk and is not specific to a single crystal face, additional analyses showed that the interactions for the same atom-pair functions exhibited identical trends across the different configurations. Therefore, the graphs reported here refer to the simulations conducted on the AC face, with the understanding that those for the BC and AB faces yielded equivalent results. This approach was adopted to streamline the presentation of the results while avoiding unnecessary repetition of identical data.
The analysis presented here supports the validation of the chosen force fields and the structural accuracy of the simulated systems, providing a solid basis for the subsequent computational evaluations discussed in the main text.

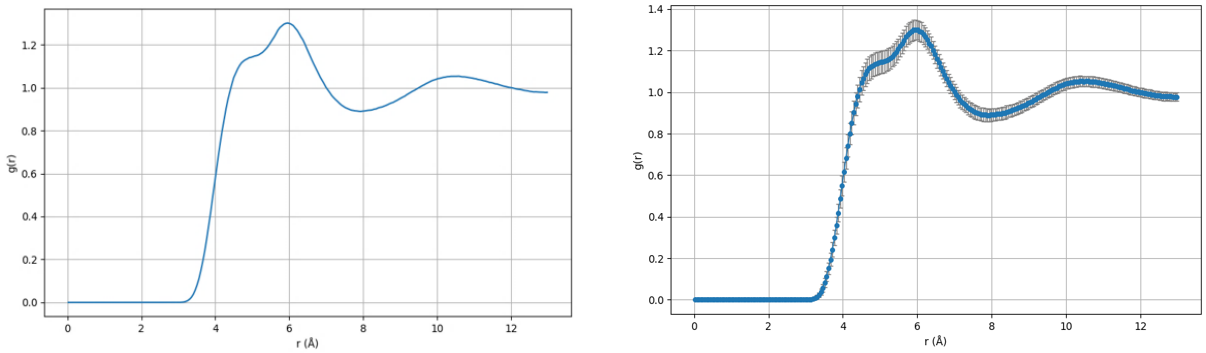
Acetone[29]



(a) (b)

Figure C.1: RDF for C1C1 in acetone, without (a) and with (b) error bars

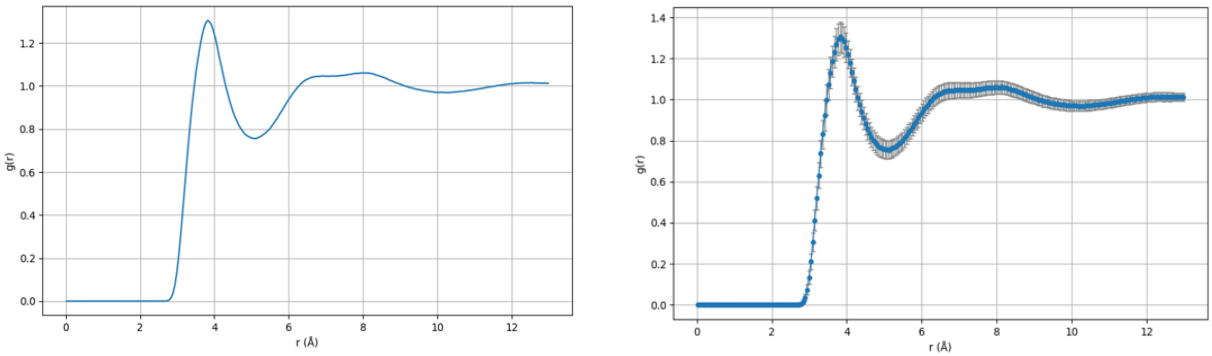
The graph shows a main peak at 5.8 Å with a width of 1.5 Å, this shows conformity with the literature data, as normally the expected first peak is between 5.8 Å - 6.0 Å.



(a) (b)

Figure C.2: RDF for C2C2 in acetone, without (a) and with (b) error bars

The graph shows a main peak at 6.0 Å with a width of 1.5 Å, this shows conformity with the literature data. However, normally the expected first peak is between 5.8 Å - 6.0 Å: the slightly bigger distance in comparison to the RDF for C1C1 could be attributed to thermal fluctuation at the simulated temperature.

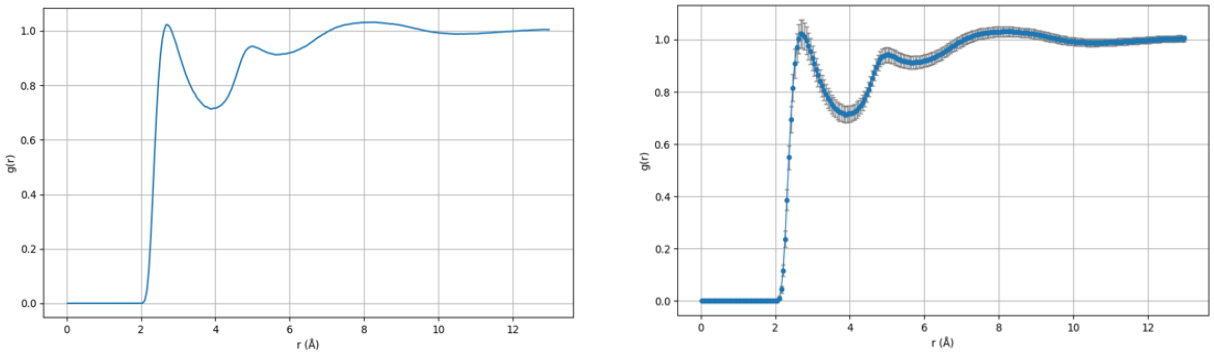


(a) (b)

Figure C.3: RDF for OC2 in acetone, without (a) and with (b) error bars

The graph shows a main peak at 3.9 Å with a width of 1.2 Å, this shows conformity with the literature data, since normally the expected first peak is registered between 3.8 Å-

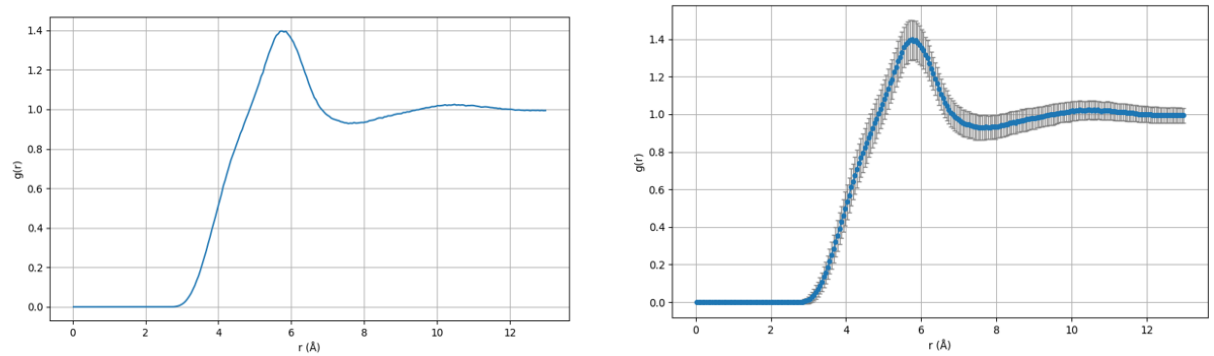
4.0 Å.



(a) (b)

Figure C.4: RDF for OH in acetone, without (a) and with (b) error bars

The graph shows a main peak at 2.8 Å with a width of 1.0 Å, this shows conformity with the literature data, since normally the expected first peak is registered between 2.7 Å - 2.9 Å.



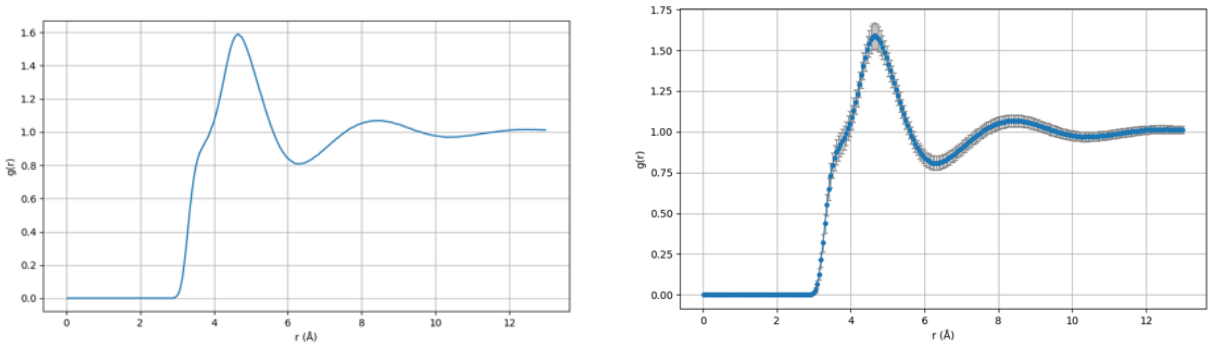
(a) (b)

Figure C.5: RDF for OO in acetone, without (a) and with (b) error bars

The graph shows a main peak at 5.9 Å with a width of 1.5 Å, this shows conformity with

the literature data, as normally the expected first peak is registered between 5.8 Å - 6.0 Å.

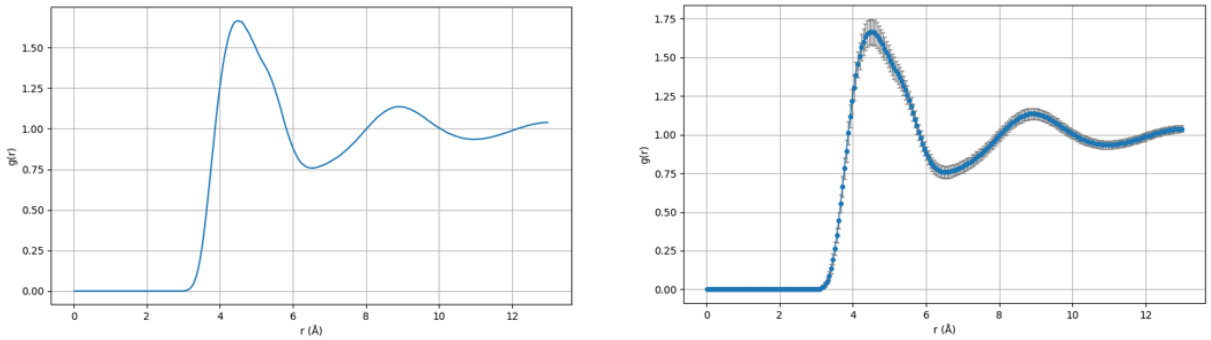
Acetonitrile[30]



(a) (b)

Figure C.6: RDF for C1C1 in acetonitrile, without (a) and with (b) error bars

The graph shows a main peak at 4.8 Å with a width of 1.3 Å, this shows conformity with the literature data, as normally the expected first peak is registered between 4.7 Å - 4.9 Å.

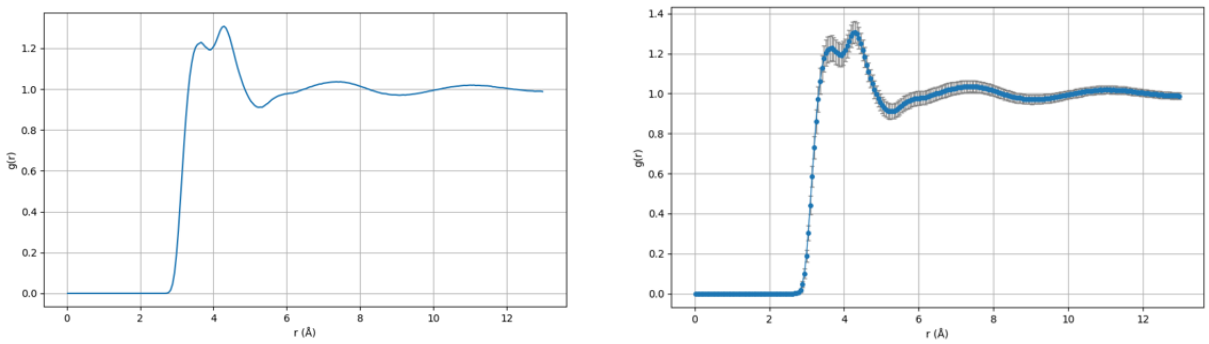


(a) (b)

Figure C.7: RDF for C2C2 in acetonitrile, without (a) and with (b) error bars

The graph shows a main peak at 4.3 Å with a width of 1.3 Å, this shows conformity with the literature data, since normally the expected first peak is registered between 4.0 Å -

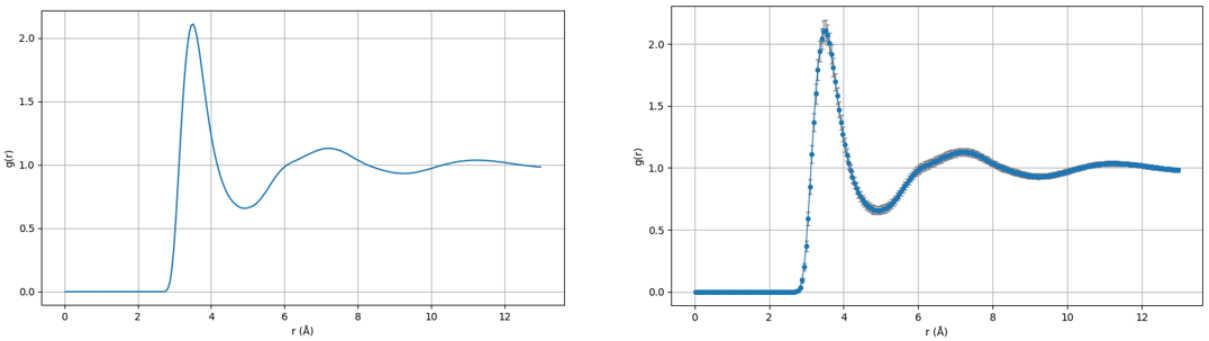
4.4 Å.



(a) (b)

Figure C.8: RDF for NC1 in acetonitrile, without (a) and with (b) error bars

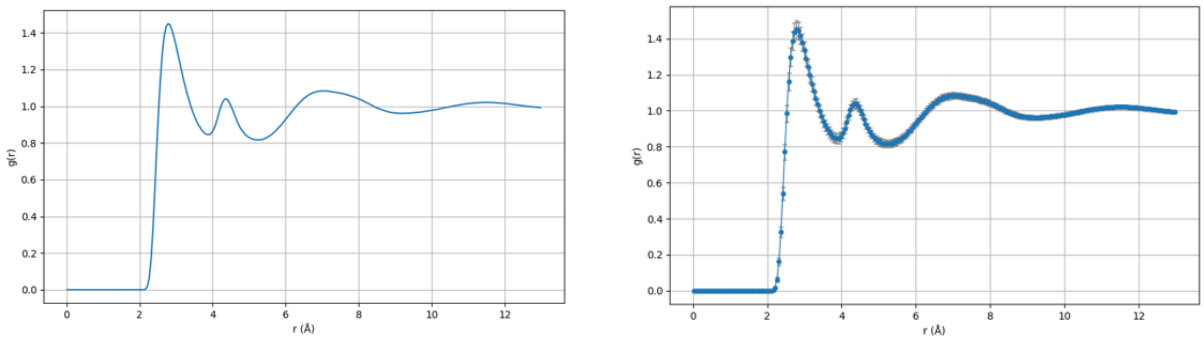
The graph shows a main peak at 4.2 Å with a width of 1.1 Å, this shows conformity with the literature data, in fact, normally the expected first peak is registered between 4.1 Å - 4.3 Å.



(a) (b)

Figure C.9: RDF for NC2 in acetonitrile, without (a) and with (b) error bars

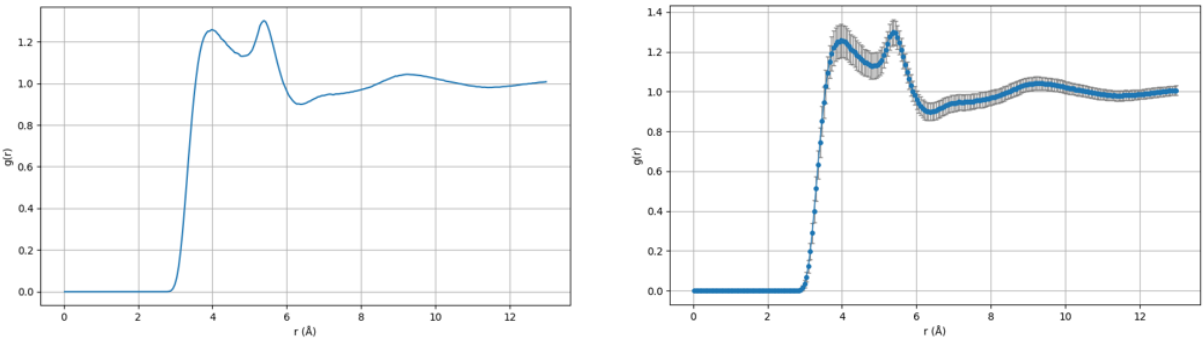
The graph shows a main peak at 3.7 Å with a width of 1.0 Å, this shows conformity with the literature data, as normally the expected first peak is registered between 3.6 Å - 3.8 Å.



(a) (b)

Figure C.10: RDF for NH in acetonitrile, without (a) and with (b) error bars

The graph shows a main peak at 2.9 Å with a width of 0.9 Å, this shows conformity with the literature data, since normally the expected first peak is registered between 2.8 Å - 3.0 Å.



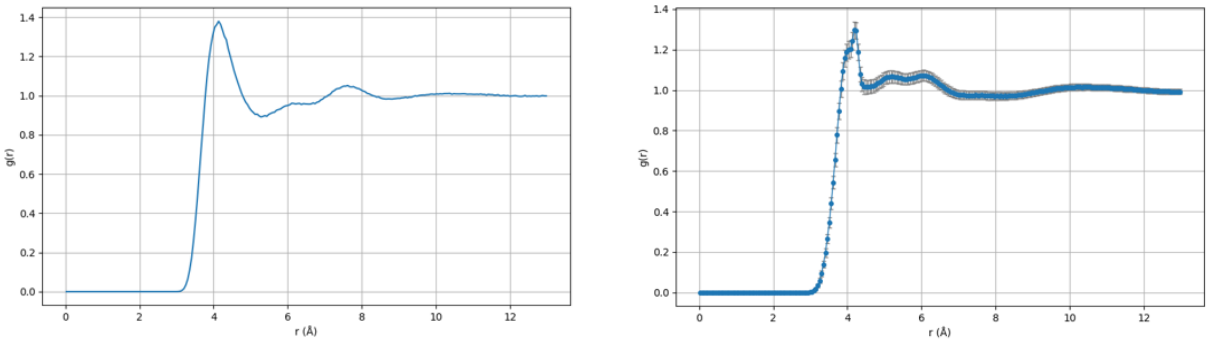
(a) (b)

Figure C.11: RDF for NN in acetonitrile, without (a) and with (b) error bars

The graph shows a main peak at 4.0 Å with a width of 1.2 Å, this shows conformity with

the literature data, as normally the expected first peak is registered between 4.0 Å - 4.4 Å.

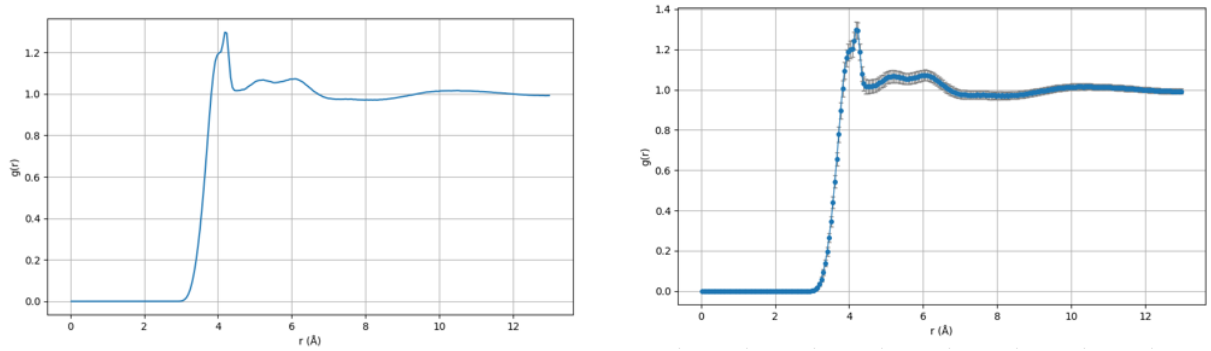
Toluene[31]



(a) (b)

Figure C.12: RDF for C1C1 in toluene, without (a) and with (b) error bars

The graph shows a main peak at 4.1 Å with a width of 1.0 Å, this shows conformity with the literature data, as normally the expected first peak for carbons in aromatic rings is registered between 4.0 Å - 4.1 Å.

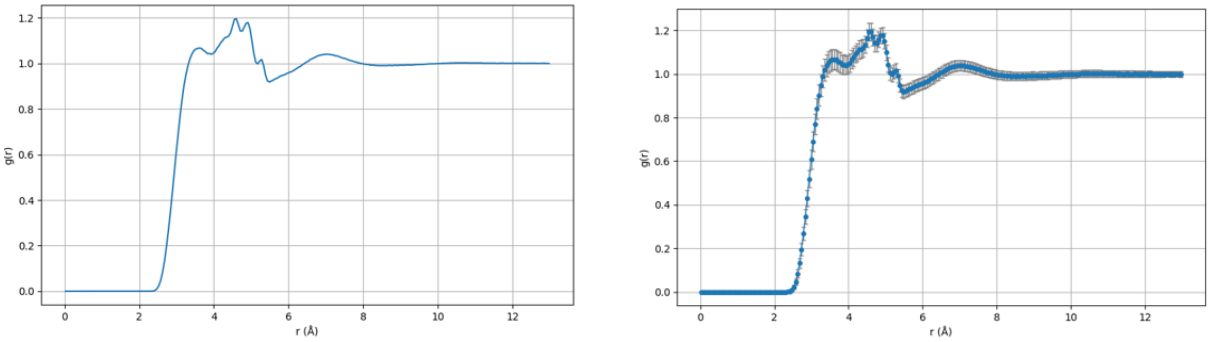


(a) (b)

Figure C.13: RDF for C1C2 in toluene, without (a) and with (b) error bars

The graph shows a main peak at 4.3 Å with a width of 0.9 Å, this shows conformity with the literature data, since normally the expected first peak for carbons in aromatic rings

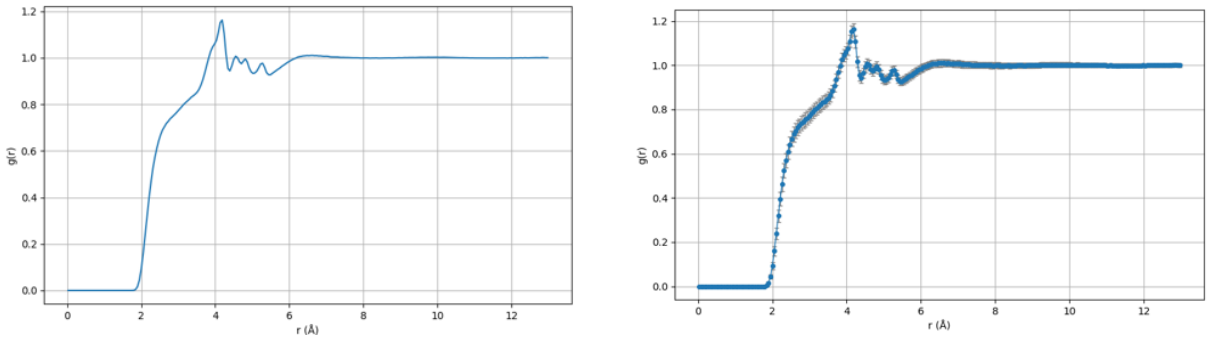
is registered between 4.0 Å - 4.4 Å.



(a) (b)

Figure C.14: RDF for C1H in toluene, without (a) and with (b) error bars

The graph shows a main peak at 5.0 Å with a width of 1.0 Å, this shows conformity with the literature data, in fact, normally the expected first peak for carbons in aromatic rings is registered between 4.8 Å - 5.1 Å.



(a) (b)

Figure C.15: RDF for HH in toluene, without (a) and with (b) error bars

The graph shows a main peak at 4.1 Å with a width of 1.0 Å, this shows conformity with the literature data, as normally the expected first peak for carbons in aromatic rings is around 4.1 Å.

Bibliography

[1] Janine Andrea Preston et al. "Elucidating the Polymorphism of Xanthone: A Crystallization and Characterization Study". en. In: *Crystal Growth & Design* 24.8 (Apr. 2024), pp. 3256–3268. ISSN: 1528-7483, 1528-7505. DOI: 10.1021/acs.cgd.3c01506. URL: https://pubs.acs.org/doi/10.1021/acs.cgd.3c01506 (visited on 09/04/2025).

[2] *Influenza della temperatura sulla solubilità di un sale*. URL: https://www.chimica-online.it/download/effetto-temperatura-solubilita.htm (visited on 07/14/2025).

[3] John William Mullin. *Crystallization*. en. 4th ed. Oxford Boston: Butterworth-Heinemann, 2001. ISBN: 978-0-7506-4833-2.

[4] Peter Rudolph, Thomas F. Kuech, and Tatau Nishinga. *Handbook of crystal growth*. en. 2nd rev. ed. Amsterdam: Elsevier, 2015. ISBN: 978-0-444-56369-9 978-0-444-63303-3 978-0-444-63304-0.

[5] Joop H. Ter Horst, Christiane Schmidt, and Joachim Ulrich. "Fundamentals of Industrial Crystallization". en. In: *Handbook of Crystal Growth*. Elsevier, 2015, pp. 1317–1349. ISBN: 978-0-444-63303-3. DOI: 10.1016/B978-0-444-63303-3.00032-8. URL: https://linkinghub.elsevier.com/retrieve/pii/B9780444633033000328 (visited on 09/04/2025).

[6] "The growth of crystals and the equilibrium structure of their surfaces". en. In: *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 243.866 (June 1951), pp. 299–358. ISSN: 0080-4614, 2054-0272. DOI: 10.1098/rsta.1951.0006. URL: https://royalsocietypublishing.org/doi/10.1098/rsta.1951.0006 (visited on 09/04/2025).

[7] M. Elwenspoek. "Comment on the α-factor of Jackson for crystal growth from solution". In: *Journal of Crystal Growth* 78 (1986), pp. 353–356. ISSN: 0022-0248. DOI: 10.1016/0022-0248(86)90071-0.

[8] G. Wulff. "XXV. Zur Frage der Geschwindigkeit des Wachsthums und der Auflösung der Krystallflächen". en. In: *Zeitschrift für Kristallographie - Crystalline Materials* 34.1-6 (Dec. 1901), pp. 449–530. ISSN: 2196-7105, 2194-4946. DOI: 10.1524/zkri.

1901.34.1.449. URL: https://www.degruyter.com/document/doi/10.1524/zkri.1901.34.1.449/html (visited on 09/04/2025).

[9] I. Chorkendorff and J. W. Niemantsverdriet. *Concepts of Modern Catalysis and Kinetics*. en. 1st ed. Wiley, Oct. 2003. ISBN: 978-3-527-30574-2 978-3-527-60265-0. DOI: 10.1002/3527602658. URL: https://onlinelibrary.wiley.com/doi/book/10.1002/3527602658 (visited on 09/04/2025).

[10] Hecham M. Omar and Sohrab Rohani. "Crystal Population Balance Formulation and Solution Methods: A Review". en. In: *Crystal Growth & Design* 17.7 (July 2017), pp. 4028–4041. ISSN: 1528-7483, 1528-7505. DOI: 10.1021/acs.cgd.7b00645. URL: https://pubs.acs.org/doi/10.1021/acs.cgd.7b00645 (visited on 09/04/2025).

[11] *Polarized Light Microscopy*. en. URL: https://www.microscopyu.com/techniques/polarized-light/polarized-light-microscopy (visited on 06/10/2025).

[12] optek-Danulat GmbH. *Control 4000 e Control 8000 – Analizzatori Universali e Fotometrici*. Brochure tecnica, versione italiana. Essen, Germania, 2011. URL: https://www.optek.com.

[13] A Buffo and V Alopaeus. "Experimental determination of size distributions: analyzing proper sample sizes". en. In: *Measurement Science and Technology* 27.4 (Apr. 2016), p. 045301. ISSN: 0957-0233, 1361-6501. DOI: 10.1088/0957-0233/27/4/045301. URL: https://iopscience.iop.org/article/10.1088/0957-0233/27/4/045301 (visited on 09/04/2025).

[14] Mathieu Rouaud. "Probability, Statistics and Estimation". en. In: ().

[15] Andrew R. Leach. *Molecular Modelling: Principles and Applications*. 2nd ed. Harlow, England: Prentice Hall, 2001. ISBN: 9780582382107. URL: http://inis.jinr.ru/sl/Simulation/Leach%20A.,%20Molecular%20Modelling.%20Principles%20and%20Applications,%202001.pdf.

[16] M.A. González. "Force fields and molecular dynamics simulations". en. In: *École thématique de la Société Française de la Neutronique* 12 (2011), pp. 169–200. ISSN: 2107-7223, 2107-7231. DOI: 10.1051/sfn/201112009. URL: http://www.neutron-sciences.org/10.1051/sfn/201112009 (visited on 09/04/2025).

[17] Hasitha Muthumala Waidyasooriya, Masanori Hariyama, and Kota Kasahara. "An FPGA Accelerator for Molecular Dynamics Simulation Using OpenCL:" en. In: *International Journal of Networked and Distributed Computing* 5.1 (2017), p. 52. ISSN: 2211-7946. DOI: 10.2991/ijndc.2017.5.1.6. URL: https://link.springer.com/10.2991/ijndc.2017.5.1.6 (visited on 09/04/2025).

[18] Neeraj Rai and J. Ilja Siepmann. "Transferable Potentials for Phase Equilibria. 10. Explicit-Hydrogen Description of Substituted Benzenes and Polycyclic Aromatic Compounds". en. In: *The Journal of Physical Chemistry B* 117.1 (Jan. 2013), pp. 273–288. ISSN: 1520-6106, 1520-5207. DOI: 10.1021/jp307328x. URL: https://pubs.acs.org/doi/10.1021/jp307328x (visited on 09/12/2025).

[19] Brian Doherty et al. "Revisiting OPLS Force Field Parameters for Ionic Liquid Simulations". en. In: *Journal of Chemical Theory and Computation* 13.12 (Dec. 2017), pp. 6131–6145. ISSN: 1549-9618, 1549-9626. DOI: 10.1021/acs.jctc.7b00520. URL: https://pubs.acs.org/doi/10.1021/acs.jctc.7b00520 (visited on 09/12/2025).

[20] Wendy D. Cornell et al. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". en. In: *Journal of the American Chemical Society* 117.19 (May 1995), pp. 5179–5197. ISSN: 0002-7863, 1520-5126. DOI: 10.1021/ja00124a002. URL: https://pubs.acs.org/doi/abs/10.1021/ja00124a002 (visited on 09/12/2025).

[21] A. D. MacKerell et al. "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins". en. In: *The Journal of Physical Chemistry B* 102.18 (Apr. 1998), pp. 3586–3616. ISSN: 1520-6106, 1520-5207. DOI: 10.1021/jp973084f. URL: https://pubs.acs.org/doi/10.1021/jp973084f (visited on 09/12/2025).

[22] Marcin Budka et al. "Molecular model of dynamic social network based on e-mail communication". en. In: *Social Network Analysis and Mining* 3.3 (Sept. 2013), pp. 543–563. ISSN: 1869-5450, 1869-5469. DOI: 10.1007/s13278-013-0101-4. URL: http://link.springer.com/10.1007/s13278-013-0101-4 (visited on 09/04/2025).

[23] A.K. Soper. "The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa". en. In: *Chemical Physics* 258.2-3 (Aug. 2000), pp. 121–137. ISSN: 03010104. DOI: 10.1016/S0301-0104(00)00179-8. URL: https://linkinghub.elsevier.com/retrieve/pii/S0301010400001798 (visited on 09/04/2025).

[24] *L'acqua e il suo comportamento anomalo: nuovi studi sui suoi passaggi di stato.* it. Nov. 2022. URL: https://www.geopop.it/comportamento-anomalo-acqua-nuovi-studi/ (visited on 08/12/2025).

[25] Johannes Schindelin et al. "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9.7 (July 2012), pp. 676–682. ISSN: 1548-7105. DOI: 10.1038/nmeth.2019. URL: https://doi.org/10.1038/nmeth.2019.

[26] IFA - Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung. *Acetone – GESTIS Substance Database*. https://gestis.dguv.de/data?name=011230. (Visited on 03/12/2025).

[27] IFA - Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung. *Acetonitrile – GESTIS Substance Database*. https://gestis.dguv.de/data?name=013660. (Visited on 03/12/2025).

[28] IFA - Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung. *Toluene – GESTIS Substance Database*. https://gestis.dguv.de/data?name=010070. (Visited on 03/12/2025).

[29] Sylvia E. McLain, Alan K. Soper, and Alenka Luzar. "Orientational correlations in liquid acetone and dimethyl sulfoxide: A comparative study". en. In: *The Journal of Chemical Physics* 124.7 (Feb. 2006), p. 074502. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.2170077. URL: https://pubs.aip.org/jcp/article/124/7/074502/910128/Orientational-correlations-in-liquid-acetone-and (visited on 09/04/2025).

[30] Szilvia Pothoczki and László Pusztai. "Intermolecular orientations in liquid acetonitrile: New insights based on diffraction measurements and all-atom simulations". en. In: *Journal of Molecular Liquids* 225 (Jan. 2017), pp. 160–166. ISSN: 01677322. DOI: 10.1016/j.molliq.2016.11.056. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167732216331701 (visited on 09/04/2025).

[31] Marco Fioroni and Dieter Vogt. "Toluene Model for Molecular Dynamics Simulations in the Ranges 298 ¡ T (K) ¡ 350 and 0.1 ¡ P (MPa) ¡ 10". en. In: *The Journal of Physical Chemistry B* 108.31 (Aug. 2004), pp. 11774–11781. ISSN: 1520-6106, 1520-5207. DOI: 10.1021/jp037605c. URL: https://pubs.acs.org/doi/10.1021/jp037605c (visited on 09/04/2025).

[32] Nadir Edith De Jesús-González et al. "Effect of the Dielectric Constant on the Solubility of Acetone in Water". en. In: *Journal of Chemical & Engineering Data* 63.5 (May 2018), pp. 1170–1179. ISSN: 0021-9568, 1520-5134. DOI: 10.1021/acs.jced.7b00573. URL: https://pubs.acs.org/doi/10.1021/acs.jced.7b00573 (visited on 09/04/2025).

[33] Robert L. Hurle and Lawrence A. Woolf. "Self-diffusion in liquid acetonitrile under pressure". en. In: *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* 78.7 (1982), p. 2233. ISSN: 0300-9599. DOI: 10.1039/f19827802233. URL: https://xlink.rsc.org/?DOI=f19827802233 (visited on 09/04/2025).

[34] Octavio Suárez-Iglesias et al. “Self-Diffusion in Molecular Fluids and Noble Gases: Available Data”. en. In: *Journal of Chemical & Engineering Data* 60.10 (Oct. 2015), pp. 2757–2817. ISSN: 0021-9568, 1520-5134. DOI: 10.1021/acs.jced.5b00323. URL: https://pubs.acs.org/doi/10.1021/acs.jced.5b00323 (visited on 09/04/2025).