University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/ Specialization: Petroleum Technology / Production	Spring semester, 2014 Open
Writer: Sergei Tantciura	…… (Writer's signature)
Faculty supervisor: Thor Martin Svartås	
Titel of thesis: Hydrate nucleation kinetics and statistical analysis of experimental data	
Credits (ECTS): 30	
Key words: gas hydrate, nucleation, induction time, activation energy, critical radius, probability, Maximum Likelihood Estimator, permutation test, confidence interval	Pages: …83… + enclosure: …2… Stavanger, …16/06/2014… Date/year

ACKNOWLEDGEMENTS

First I would like to express my sincere thanks to my thesis supervisor Dr. Thor Martin Svartås for professional guidance, great interest and valuable feedback all along the work with the thesis.

The next, I thank Jan Terje Kvaløy for filling gaps in my knowledge concerning the mathematical statistics issues of the present work.

Also I am very grateful to Aina Undersrud Bratland for permission to use the data she produced as well as for interesting discussions during the work with the thesis.

And last but not least, I thank the UiS library and laboratory staff for fast and effective service provided for me.

ABSTRACT

The effect on accuracy of reducing the number of experiments in the series has been investigated on isochoric, constant temperature nucleation experiments. In addition, data from two other nucleation studies have been analyzed.

The experiments were conducted in two autoclave high pressure cells with similar construction but different overall dimensions. Two different gas mixtures were used: SNG-2 and SNG-7. Experimental conditions such as pressure (90 bar) and cooling rate (6.0 °C/h) were kept constant for all the experiments. Three experimental series at different temperatures have been conducted on each cell, which allowed estimation of activation energy of gas hydrate nucleation by means of the Arrhenius equation through the Arrhenius plot. The radius of the critical size nuclei was determined once the activation energy was estimated.

Two different techniques were utilized in order to estimate nucleation rate of hydrate nucleation, namely a best-fit method included in the graphical program tool KaleidaGraph® and penalized Maximum Likelihood Estimation (MLE) for two-parameter exponential distribution.

Two maximum deviation levels of 20% and 30% from estimated value obtained with the maximum number of parallels included were set as test criteria evaluating the minimum number of parallels required to determine particular parameter.

Based on the results of statistical analysis of the experimental data in the present study the minimum number of experiments to obtain a representative value of nucleation rate estimate can be proposed to be not less than several tens to provide accuracy within 20-30%. To make some conclusions about the exact number, analysis of a series with more parallels is required.

For the conducted experiments acceptable accuracies were obtained for estimation of nuclei size even though the number of parallels was reduced to 6 for the large cell experiments. However, there was not found clear relationship between number of parallels and fit of $Ln\ J$ as function of $1/T$ to a straight line.

The set accuracy limit that determines maximum acceptable deviation in a parameter estimate showed great influence on the required number of parallels.

NOMENCLATURE

ΔG_s – Gibbs free energy of the surface of a hydrate cluster, J

ΔG_v – Gibbs free energy of a hydrate cluster, J

E_a – activation energy of nucleation, J

J_0 – pre-exponential factor

P_{exp} – experimental pressure, bar

T_a – absolute temperature, K

T_{eq} – equilibrium temperature, °C

d_i – number of events occurred at time t_i

k_B – Boltzmann's constant, J/K

r_c – critical nucleus radius, Å

$x_{1:n}$ – minimum value in a series

τ_0 – lag time, min

Δg_v – Gibbs free energy change per unit volume of hydrate bulk phase, J/m^3

ΔT – subcooling, °C

Δt – time interval, min

AA – antiagglomerant

CDF – cumulative distribution function

HEN – heterogeneous nucleation

HON – homogeneous nucleation

KHI – kinetic hydrate inhibitor

MLE – Maximum Likelihood Estimation

PDF – probability density function

sH – structure H hydrate

sI – structure I hydrate

sII – structure II hydrate

SNG – synthetic natural gas

A – area, m^2

$GradP$ – pressure gradient, bar/°C

J – rate of nucleation, $m^{-3}s^{-1}$ or min^{-1}

M – total number of experiments

$P(t)$ – cumulative probability of detecting hydrate formation at or before time t

S – supersaturation ratio

T – temperature, °C

V – volume, m^3

m – number of formed nuclei

$n(t)$ – number of experiments in which nuclei were detected within time t

rpm – revolutions per minute

t – induction time, s or min

$\Delta T'$ – difference between the initial temperature and the temperature of the experiment

α – significance level

η – continuous location parameter for MLE a two-parameter exponential distribution

θ – continuous scale parameter of a two-parameter exponential distribution

σ – interfacial tension, N/m^2

TABLE OF CONTENTS

1 THESIS DEFINITION

The main goal of hydrate kinetics determination is to predict hydrate formation in order to prevent blockages of flow channels (e.g. deep-water pipelines, well choke and kill lines, etc.) by hydrate accumulations [1].

Gas hydrate formation is a complex phenomenon. Hydrate formation is crystallization process that consists of two distinct regions: nucleation period and growth stage. The nucleation region is the process wherein a subcritically sized crystal may either grow or shrink before becoming a stable crystal. The growth stage begins once the nuclei have surpassed its critical size and larger crystals are being formed [1, 2].

Like other crystallization processes hydrate formation is not governed by thermodynamic laws but is rather stochastic [3-8]. There exist many different studies that investigate the nucleation phenomenon. One may notice a big discrepancy in a proposed number of experiments that required to produce acceptable limit of accuracy for a given parameter estimation. Considering stochastic behavior of nucleation process, some authors propose several tens of experiments, other ones state that the amount should be in order of hundreds.

A technique for estimation of the nucleation parameters by means of statistical methods was initially developed by Toschev et al. [9]. They stated that reliability of the method used in their theory can only be provided by large number of experiments. They performed 500 uniform experiments with electrolytically deposited mercury droplets on platinum in order to estimate parameters of highly stochastic phenomenon of nucleation.

As for other studies, Jiang and ter Horst [4] to estimate nucleation rates for m-aminobenzoic acid and l-histidine used at least 80 induction times per supersaturation. Yang and Rasmunson when studied nucleation of butyl paraben performed 100 experiments at each condition. Kulkarni et al. [5] studied crystal nucleation rate by using isonicotinamide in ethanol and performed 144 experiments per supersaturation ratio. Wilson et al. [10] assumed 300 parallels as sufficient to provide accurate statistics for their nucleation studies. Ohmura et al. [7] claimed that “the number of data items should typically be several tens or larger to enable application of any statistical data processing procedure that can reveal the stochastic nature”. They used thirty to fifty samples of water/hydrochlorofluorocarbon system in each thermal

history program under atmospheric pressure to collect a sufficient amount of data for statistical treatment when searched for the nature of the “memory effect”.

When it comes to last works done at the University of Stavanger, Magnus Palm [11] expressed his concern about insufficiency of 6 experiments in a baseline in his MSc thesis. Eirik Høvring [12] in his MSc work related to estimation of the activation energy of hydrate formation, stated that “in some cases six experiments are too few to cover the time window where nuclei formation would occur most frequently at given experimental conditions”. He suggested that ideally the number of experiments within a set of parallel experiments should probably be around 15 to 20 or even greater.

The induction time in gas hydrate crystallization is an important characteristic of the kinetics of the process. Long induction period allows transport of fluids through the production facilities to the processing plants without crystallization of hydrates in the system [13]. Therefore it is a parameter of interest in development of new inhibition technologies.

Since the experiments for determination of induction time are often time consuming, it is of a great importance to know exactly the number of experiments that is needed to obtain results with acceptable level of accuracy.

Accurately estimated nucleation rate in turn will affect evaluation of activation energy of hydrate nucleation and a critical nuclei radius, which are also key parameters in development of low dosage hydrate inhibitors.

The main objective of the present MSc thesis is to determine whether it is acceptable to conduct just few experiments in order to estimate appropriate parameters of highly stochastic phenomenon of gas hydrate nucleation.

2 THEORETICAL PART

2.1 General overview of gas hydrates phenomena

Gas hydrates are a unique class of chemical compounds that physically resemble ice or wet snow [14]. Figure 1 shows photos of gas hydrates formed in transparent sapphire cell in the hydrate laboratory at Department of Petroleum Engineering, UiS during the present study. But apart from appearance, structures of gas hydrate and ice are significantly different. In gas hydrates molecules of one compound (the guest material) are enclosed without bonding chemically within an open solid lattice composed of another compound (the host material). These types of configurations are known as clathrates [15]. The term clathrate was originated from the Latin *clatratus* meaning “with bars” or a “lattice”. The lattice structure formed from the water molecules is thermodynamically unstable and is always stabilized by the incorporation of gas molecules [16]. Without support of the trapped molecules, the lattice of hydrate clathrates would collapse into conventional ice crystal structure or liquid water. The gas molecule is kept inside the lattice by weak van der Waals forces [17].

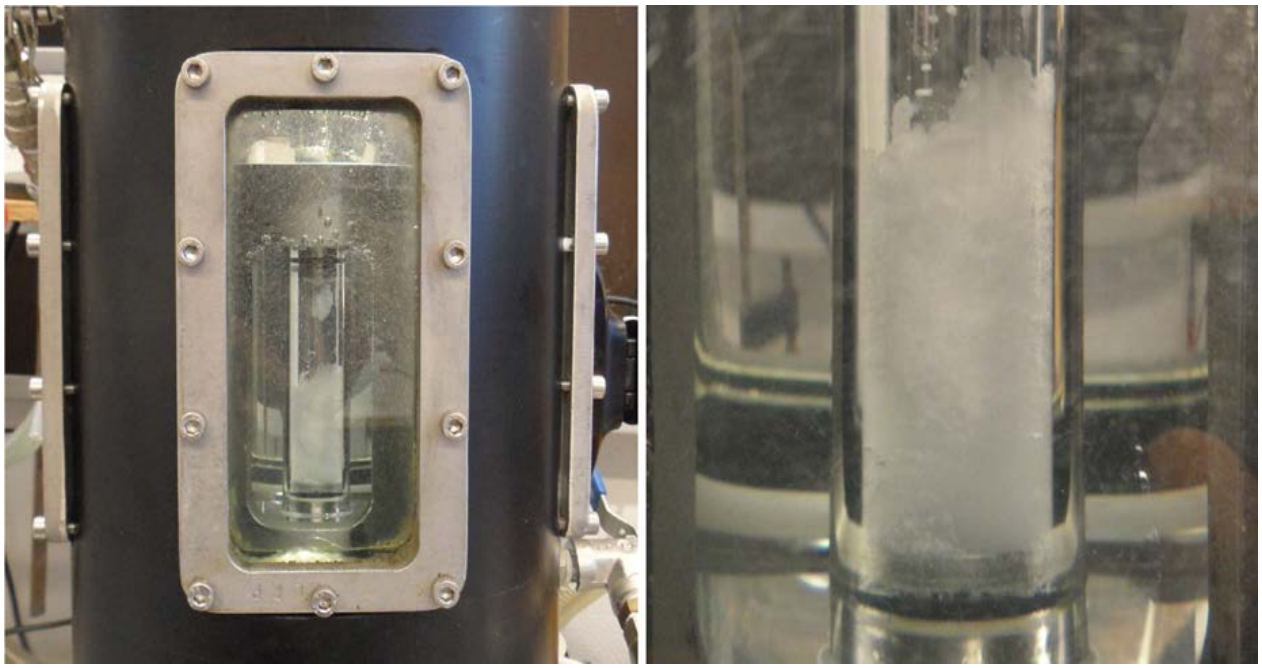

Figure 1 – Hydrates formed from water and the seven component gas mixture (SNG-7) in the sapphire cell used for experiments in the hydrate laboratory at Department of Petroleum Engineering, University of Stavanger

Some of gas hydrates properties are unique. For instance, one volume of water is bonding 160 volumes of methane. And the specific volume of water thereby increases by 26% in comparison with 9% during freezing [18].

Many different gases and organic liquids show an ability to form hydrates. In the oil and gas industry great significance attaches to the hydrates, which form from water and different hydrocarbon components such as methane, ethane, propane, isobutane, n-butane, nitrogen, carbon dioxide, or hydrogen sulfide [16].

Depending on gas composition clathrate hydrates can exist at temperatures far above freezing point of water (up to 29°C) and the pressures comparable to that of industrial production and transportation [19, 20]. Because of that, these compounds became a nuisance in high pressure gas operations where water is present, since their formation causes partial or complete plugging of valves and pipes [21].

Gas hydrates can be divided into two distinct groups according to their genesis: 1) natural, which form and exist in nature without human intervention; methane in natural gas hydrates is dominantly generated by bacterial degradation of organic matter in low oxygen environments [22]; 2) technogenic, which form in technological systems created and controlled by man.

Although natural and technogenic gas hydrates are identical in their essence, problems generated by both groups, methods of solution, and research goals and consequences of the solutions are very difficult. The study of natural gas hydrates is currently at a state where new problems and finding methods of their solution appear continuously [23].

Recently there were some discussions about environmental aspects of hydrate clathrates. The proximity of natural gas hydrate reservoirs to the seafloor has motivated speculations about release of methane in response to climate variability and as playing a role in large undersea slumps which could result in disastrous tsunamis [23-25].

The importance of natural gas hydrates phenomena is confirmed by their active study in the world by laboratories having the latest equipment as well as by sharp increasing in amount of publications. The findings of research published in numerous works indicate interesting results on the conditions of gas hydrates formation in earth, their spreading on land and under sea, and others were obtained [23].

2.2 Historical review

Gas hydrates first documentation and description were given in 1811 by Sir Humphrey Davy with brief comments in the Bakerian lecture to the Royal Society in 1810 [3, 14, 17, 26]. He performed laboratory experiments with chlorine and water. Villard (1888) was the first who reported hydrates of hydrocarbons such as methane, ethane, propane, acetylene and ethylene [3, 14].

For more than one century after the discovery hydrates were considered as scientific curiosity before E.G. Hammershmidt in 1934 determined that methane hydrate was responsible for plugging natural gas pipelines [14]. At that time, with the growth of the natural gas industry, the production, processing and distribution became high-pressure operations, favorable for hydrate formation [17]. The research entered its next phase, when the hydrates were concerned as a hindrance to the natural gas industry [15, 21].

The next period in the clathrate hydrates phenomena investigation began in 1967, when a group of Soviet geologists discovered the first major deposit in the permafrost. Although the estimated reserves of the Messoyakha field were extremely small (30×10^9 m^3) compared to other gas fields of Western Siberia, this field catalyzed development in investigation of natural gas hydrates. During the decade beginning in 1969, more than 5×10^9 m^3 of natural gas were produced from hydrates in the field. An existence of gas hydrates reservoirs as well as possible industrial development thereof was proven [3, 18].

This discovery provoked great interest to gas hydrates as an unconventional energy source. During the next decades large gas hydrate accumulations have been located in the terrestrial permafrost regions as well as in several hundred meters below marine sediments [16]. Potential resources in hydrates are estimated to be $1{,}5 \times 10^{16}$ m^3 of natural gas, which is higher than all the surveyed world resources of hydrocarbon gases by two order of magnitude [23].

A further motivation for performing time-dependent hydrate studies is the increasing interest in assessment and production of energy from natural gas hydrates in permafrost and oceanic deposits [3]. The last step in hydrate research was provoked in the mid 90-s by India and Japan, countries with large energy needs but substantially confined resources. These two countries began self-funded hydrate programs in preparation for production of methane from marine reservoirs [15].

Energy concentrated in natural gas-hydrates may possibly provide humankind with cheap, ecological energy for the majority of the 21st Century. However, so far there is not any efficient technique to develop the marine unconventional hydrates, and an engineering breakthrough is required for energy recovery to be economically feasible. Studies of natural gas hydrates must be coordinated on a world scale, which could speed up technology development [27].

2.3 Gas hydrate structures

It is important to get basic understanding of gas hydrate phenomena on microscopic level, in particular to review different structures of these compounds, which impact microscopic hydrate accumulations such as hydrate plugs and natural gas hydrate reservoirs [28].

Depending upon the arrangement of water molecules in the crystal and the conditions of formation, three common structures (sometimes often referred to as types) of clathrate hydrates can form: two cubic structures sI and sII and one hexagonal sH [17]. Figure 2 provides an overview and comparison of these hydrate structures.

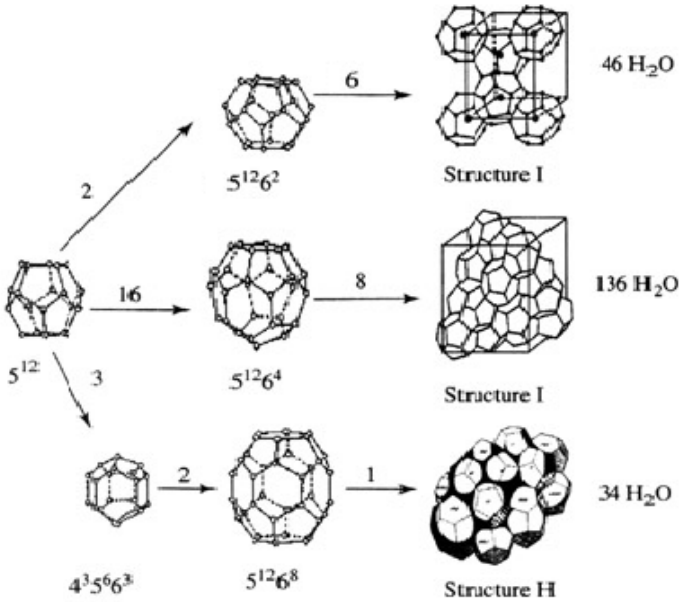

Figure 2 – Clathrate hydrate sI, sII and sH structures and their constituents [28]

2.3.1 Structure I (sI), body centered cubic

The least complex and the most common hydrate structure is the type I, which unit cell consists of 46 water molecules. It forms two types of cages – small and large. Each unit cell contains two small cages and six large ones. The small cage has the shape of a pentagonal dodecahedron, 12-sided cavity (5^{12}) with equal angles and edge lengths. The large cage presents tetracaidecahedron ($5^{12}6^2$), which in addition to 12 pentagonal faces has 2 hexagonal ones [17, 28].

The guest molecules of the type I structure are typically represented by methane, ethane, carbon dioxide and hydrogen sulphide. One interesting feature is that CH_4, CO_2, and H_2S can occupy both the small and large cavities, but C_2H_6 molecule occupies only the large cages [17].

2.3.2 Structure II (sII), diamond lattice

Although type II hydrates are also constructed from two types of cages, their structure is significantly more complicated. Sixteen small cages are represented by a pentagonal dodecahedron (5^{12}), like in sI structure. Six big cages have a shape of hexakaidecahedrons that are denoted $5^{12}6^4$ because in addition to 12 pentagonal faces they contain 4 hexagonal faces. The number of water molecules in Type II hydrate is 136 [3, 17] .

Among the common type II formers in natural gas are nitrogen, propane, and isobutane. Nitrogen can occupy both the large and small cages, but propane and isobutane occupy only the large cages [17, 28].

2.3.3 Structure H (sH), hexagonal

Type H hydrates are much less common then type I or II. The unit cell of type H consists of 34 water molecules, forming three types of cages – two small of different type and one large. In this case, the unit cell comprises three small cavities of type 5^{12}, two small ones of type $4^3 5^6 6^3$ called an irregular dodecahedron and one large of type $5^{12}6^8$, known as irregular icosahedron [3].

Type H hydrate includes both a small molecule such as methane and a larger, type H forming molecule. The important condition is that formation of this type of clathrate to be stable requires cooperation of large and small molecules. Type H formers are represented by such complex species as 2-methylbutane, 2,2-dimetilbutane, methylcyclopentane, which are not commonly found in natural gases [17].

Besides sI, sII and sH hydrate structures, a few other clathrates phases have been discovered, for instance Jeffrey's structures III to VII. Some other structures can exist only at very high (in order of GPa) pressure conditions [29, 30].

As for synthetic natural gas mixtures SNG-2 and SNG-7, used in the present work, both of them form type-II hydrate structures [31].

2.4 Hydrates prevention and mitigation

In petroleum technology, gas hydrates are undesirable compounds that may form during production or transportation of natural gas [16]. Typically normal functioning of oil and gas industrial systems is violated by hydrate plug formation.

In general, there are three well-known techniques to prevent hydrate formation:

1) maintaining thermobaric conditions outside the hydrate stability zone (lowering the pressure and increasing the temperature),

2) gas dehydration, i.e. decreasing a water content in gas,

3) utilization of gas hydrate inhibitors.

The first group of methods is represented mainly by burial, insulation and heating of the equipment [3]. One effective method of providing heat is the use of electrical-resistance heating cables. Confining the pressure below hydrate formation region is not suitable method for industrial processes since it may decrease the energy density to a point that is not economical. Therefore the pressure reduction is often used when gas hydrates are already formed and there is necessity of hydrate plug removal. This potentially dangerous procedure must be performed in accordance with all the safety rules. Particularly, depressurizing of the plug has to be done from both sides [3].

Gas dehydration implies removal of free and dissolved water from the system by means of separators, glycol dehydrators, molecular sieves or other methods.

The third group of methods includes various inhibition techniques. Traditional hydrate inhibitors such as alcohols (mainly methanol), glycols, and ionic salts have been in use for many years [16]. They basically inhibit the hydrate formation by depressing the freezing point — a thermodynamic effect [17]. The major drawback of thermodynamic inhibition is in large amounts of methanol or glycol required (often more than 20 wt% of the aqueous phase) [32, 33]. In addition, alcohols can cause safety problems since they are highly flammable liquids.

It was of great importance to develop low-dosage inhibitors, and in the 1990s, two types of chemical inhibitor technologies (antiagglomerants and kinetic inhibitors) were introduced [3]. The interesting feature of antiagglomerants (AA) is that they do not prevent hydrates from forming, but do prevent them from agglomeration, i.e. sticking together. Thereby, his type of inhibition technique leaves hydrates dispersed in oil and prevents them from accumulation, thus removing the reason of blockages in pipelines [33]. The second technology with kinetic hydrate inhibitors (KHI) was designed to influence the induction time of nucleation thereby preventing crystal growth for a period exceeding the free-water residence time in a pipeline [3, 32, 34].

With a required concentration of less than 1 wt% kinetic inhibitors and anti-agglomerants have been extensively used in recent years. These two technologies showed successful results in several field applications in deep-water systems under various conditions [7, 35].

Recently a new potential chemical-free technique, known as cold (or stabilized) flow, has received much attention in both the industrial and academic communities. The key principle of the technology is to convert all free water droplets entirely to hydrate as rapidly as possible. Two different concepts of the technique were patented by ExxonMobil and SINTEF Petroleum Research [28].

At the same time, with the extensive utilization of low-dosage hydrate inhibitors and developing new techniques, relatively new area in the oil and gas industry known as flow assurance shifted from hydrate avoidance to hydrate risk management that has been proven to be more economical. It involves the use of transient methods to delay hydrate formation or prevent particles from agglomeration, thus preventing hydrate blockages. However, selection of the hydrate control scheme is always a question of balancing capital against operating cost [3, 28].

2.5 Gas hydrate nucleation

Gas hydrate formation is a crystallization process in which two distinct regions can be distinguished, namely nucleation and a growth stage [1].

According to Sloan and Koh [3], hydrate nucleation is a process, during which small clusters of water and gas (hydrate nuclei) grow and disperse in an attempt to achieve critical size for continued grow.

The nucleation period is characterized by experimental difficulties. The two principle reasons for difficulties are: (1) nucleation is a microscopic phenomenon which involves only small numbers of molecules so that it is difficult to observe, and (2) it is characterized by high degree of metastability, i.e. the ability of a nonequilibrium state to persist for a long period of time [1, 36].

The phenomenon of metastability can be explained by means of a plot in Figure 3, where AB is an equilibrium curve and CD is a so-called thermodynamic spinodal curve that defines metastable limit. When a system is at point P, nucleation is impossible since the solution is superheated by the amount PR. To the left of the line CD nucleation occurs spontaneously because driving force is very high. However, between these lines (point Q), there is possibility to form as well as not to form nuclei in area that is called metastable region [1].

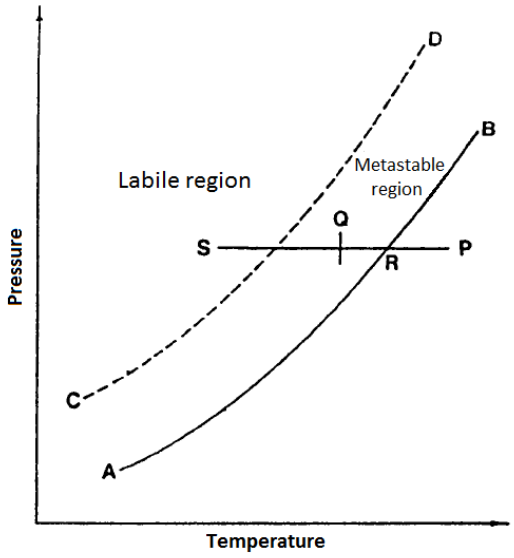

Figure 3 – Hydrate formation as function of subcooling, AB – equilibrium line, CD – spinodal line (modified from [1])

At the end of the nucleation period the cages are long-lived but unstable. The structures can either dissipate or grow to hydrate unit cells or agglomerations of unit cells, thus forming metastable nuclei. Since these metastable unit cells are of subcritical size, they may grow or dissipate too. The metastable nuclei are in quasi-equilibrium with the cages until the nuclei reach the critical radius, thus becoming stable for monotonic growth [1].

2.5.1 Main parameters of hydrate nucleation process

Like other crystallization processes hydrate formation is not governed by thermodynamic laws but is rather stochastic [3-6]. Actually, many researchers observed significant variations in time intervals elapsed for detection of hydrates from repeated experiments in a bulk phase. Accordingly, in order to compare and combine experiments of hydrate formation key parameters defining nucleation have to be discussed, namely driving force, induction time, and so-called "memory effect" [6].

2.5.1.1 Driving force

The rate at which hydrate nuclei are formed and then grow up is related to thermodynamic driving force [32]. The driving force is a key component of a hydrate nucleation correlation. Numerous driving forces can be found in literature on hydrate nucleation phenomena. The most common ones are subcooling and supersaturation [13, 37, 38].

1) Supersaturation

A solution is supersaturated when the solute concentration exceeds its solubility limit. It is possible for the solution to maintain its supersaturation over a concentration range for a certain period without formation of crystals [39, 40].

According to Mullin [41], supersaturation ratio S for a given temperature is defined as the ratio of the supersaturated solution concentration to the equilibrium concentration at the solubility curve (line AB in Figure 3).

Physically, the supersaturation is the gain in free energy per molecule associated with the passage of the phase from the minimum with higher Gibbs free energy to the minimum with lower Gibbs free energy. In other words, it is the difference between the chemical potentials of the old (solution) and the new (nuclei) phases at isothermal conditions [42].

Effect of the supersaturation on the rate of nucleation for solutions and melts is described by Figure 4, where a plot, shown by the solid curve, indicates the exponential increase in the rate once some critical level of supersaturation is exceeded. In reality one have the rapid increase up to a particular level, followed by slowdown and subsequent rapid decrease of the nucleation rate with the same speed as for growth (dashed line). This could be caused by the sharp increase in viscosity with supercoiling which restricted molecular movement and lowered the rate of formation of ordered crystal structures [41]. However, this behavior often is not the case for gas hydrate nucleation, since the solubility of gas in water is relatively low [20].

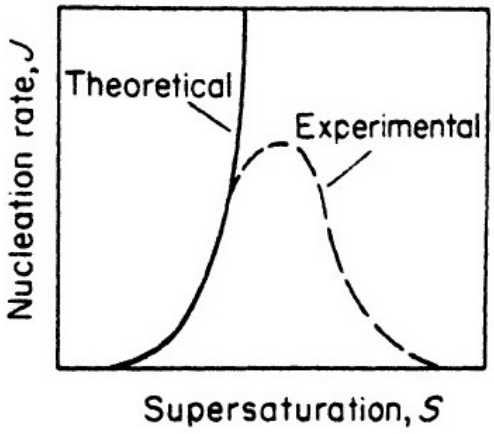

Figure 4 – Effect of supersaturation on the nucleation rate [41]

2) Subcooling (sometimes also called supercooling) ΔT can be defined as the difference between the hydrate equilibrium temperature T_{eq} and the lower system temperature at a given pressure. The driving force is increasing with increasing of subcooling.

Figure 5, where C_S and C_{Sh} are the methane solubility in water curves without and with hydrate present, gives the relation between subcooling and the supersaturation. The points A through F correspond to different temperatures during the continuous cooling process [3]. So, if hydrate formation occurs at equilibrium temperature T_{eq}, the length of the line segment CC between the curves C_S and C_{Sh} would represent the degree of supersaturation that will grow as the temperature decreases. In turn, the subcooling as temperature difference between T_{eq} and lower temperature is also increasing when moving from right to left along the temperature axis. In connection with Figure 4, nucleation rate is also function of the subcooling.

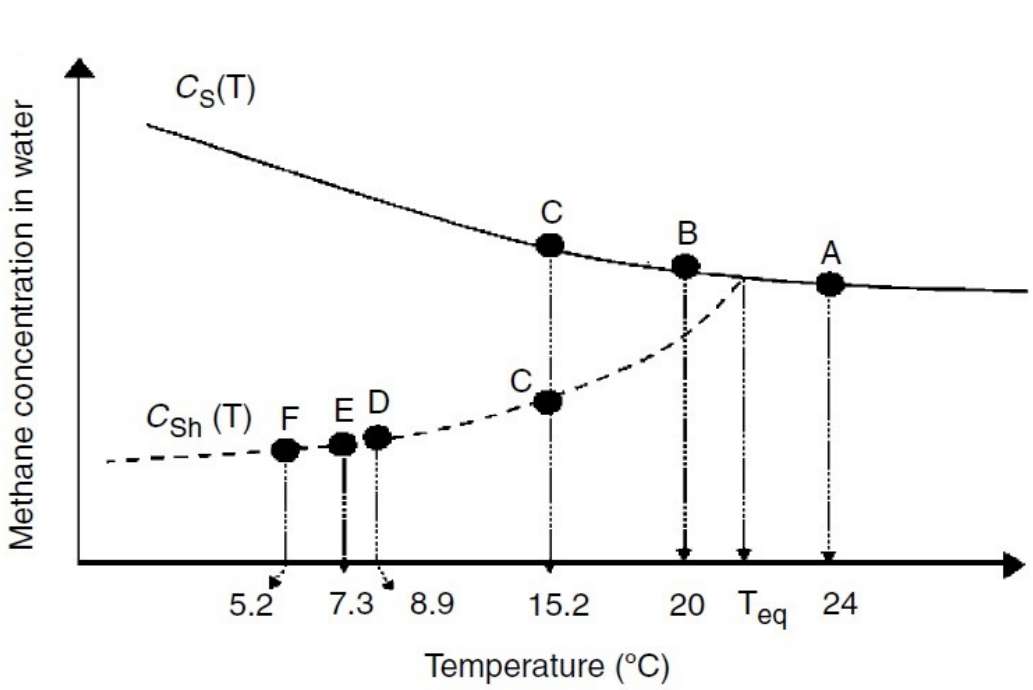

Figure 5 – Effect of temperature on the equilibrium methane concentration in water (C_S – without hydrate, C_{Sh} – with hydrate). Points A through F correspond to different temperatures during the continuous cooling process [3]

If we look at Figure 3, the degree of subcooling will increase when entering the metastable region and moving to the left of the equilibrium line AB along PS. For instance, the degree of subcooling at Point Q equal RQ.

The subcooling is widely used as driving force for hydrate formation. The main drawback of subcooling is that it does not encompass the effect of pressure [38].

A comprehensive driving force for hydrate formation is a function of pressure, temperature and gas composition. However, the results of Arjmandi et al. investigation [38] show that for pure gas–water systems subcooling provides good approximation of the driving force over a wide pressure range at a given temperature. For multicomponent systems (e.g., natural gases) driving force is significantly underestimated by subcooling for pressures less than 20 MPa. For higher pressure conditions the driving force is matched well and can be represented solely by subcooling [3].

In the present MSc study the driving force is assumed to be represented by the degree of subcooling, since the experiments were conducted at isobaric conditions.

2.5.1.2 Induction time of nucleation process

According to Kashchiev and Firoozabadi [32], induction time is measure of the ability of a supersaturated system to remain in the state of metastable equilibrium.

Looking back to Figure 3, induction time is a specific feature for hydrate nucleation process for the system located in the region between the lines AB and CD (metastable region). However, with approaching to the line AB the induction time in metastable region is not an object of correlation or prediction since the driving force is too low [3].

Induction time can be defined as a period of time elapsed from the moment of supersaturation creation in the solution to the first appearance of the secondary (solid) phase. As supersaturation (and subcooling) increases, the induction time is reduced [41].

The induction time is most likely to be dominated by nucleation period, but also includes growth up to a point at which hydrates are first detected [3]. Thus, the induction time comprises this growth time, i.e., a lag between the time of appearance of a nucleus and the time of detection thereof [4]. This time offset depends on the size at which nuclei are detectable (by means of macroscopic instruments) and the growth rate [5].

In diversity of the scientific studies that investigate the gas hydrates formation phenomena and include their own definitions and notations, there is no general agreement on terminology. Particularly, there is big inconsistency concerning such parameters as induction time and lag time [11]. In order to avoid confusion and misunderstanding, in the present MSc study the main parameters are defined as follows:

The induction time is a time interval that begins from the moment of establishment of initial conditions and lasts to the moment of hydrate detection by means of equipment. Thereby in the present study this time is counting from the moment when the stirrer is turned on until the hydrates are first detected.

The lag time τ_0 is a time elapsed between the occurrence of a critical nucleus and detection of thereof by means of equipment. Therefore the probability of detection gas hydrates before this time equals zero. This parameter was estimated for a particular series of experiments by statistical methods.

2.5.1.3 Memory effect phenomenon

Besides the induction time and the driving forces, the history of aqueous phase involved in the hydrate formation seems to play a significant role in the kinetics of hydrate formation [6].

Several studies proved that water molecules can keep a local spatial organization during a certain time after hydrate dissociation or ice melting [7, 43-45]. Consequently, hydrates form more easily from gas and water obtained by melting hydrate, than from fresh water with no previous hydrate history. However, this effect will be destroyed if the hydrate system is heated sufficiently above the hydrate formation temperature at a given pressure [3].

This strange phenomenon can be explained by two opposing hypothesis that are so far an object of speculations among several researchers [3, 46, 47]:

1) Hydrate structure (which is not visible to the naked eye) remains in solution after hydrate dissociation. Such molecular arrangement would increase the probability for the hydrate nuclei to achieve the critical size and then will fasten their growth [3, 6, 32],

2) Dissolved gas remains in solution after the hydrates were decomposed.

Some researchers apply this so-called “memory effect” in order to explain the apparent reduction in induction time for hydrates formed repetitively from supercooled solutions [47]. Duchateau et al. [8] proposed a new experimental procedure for laboratory evaluation of kinetic hydrate inhibitors that utilizes the memory effect to reduce degree of stochastic behavior and produce repeatable results.

However, although evidences of the memory effect are plentiful, there is no general consensus about this phenomenon. Furthermore, there have been only a limited number of direct molecular investigations to verify this effect [3]. In contrast, the experiments of some other researchers put into doubt existence of the memory effect [46-48].

2.5.2 Effect of agitation on hydrate nucleation

Nucleation may occur spontaneously or it may be induced artificially. The condition of supersaturation or subcooling alone is not sufficient cause for a system to begin to crystallize.

Nucleation can often be induced by agitation, mechanical shock, friction and extreme pressures within solutions. In some cases stirred water will allow only 0.5 °C of supercoiling before spontaneous nucleation occurs, whereas undistributed water allows over 5 °C [41].

Artificial stirring in a laboratory vessel will have a particularly strong effect on the surface structure of water. The water surface will be more ruffled, and bubbles of different sizes may be forced longer into the water and even get distributed in the water bulk phase, depending on the stirring characteristics (stirring rate, stirrer and vessel geometry). There can be significant amount of gas trapped inside these bubbles, which leads to increase in concentration and contact area of the interface between the active phases [39].

In summary, the effects of stirring may significantly alter the nucleation rate [4]. Normally increased stirring leads to reduction in induction times. At the same time intense stirring may induce destructive shear forces that can increase the induction times [39].

2.6 Types of nucleation

According to Mullin [41], the nucleation may be divided on primary and secondary. In turn, primary is divided on homogeneous and heterogeneous. Thereby the simple scheme may be proposed (Figure 6):

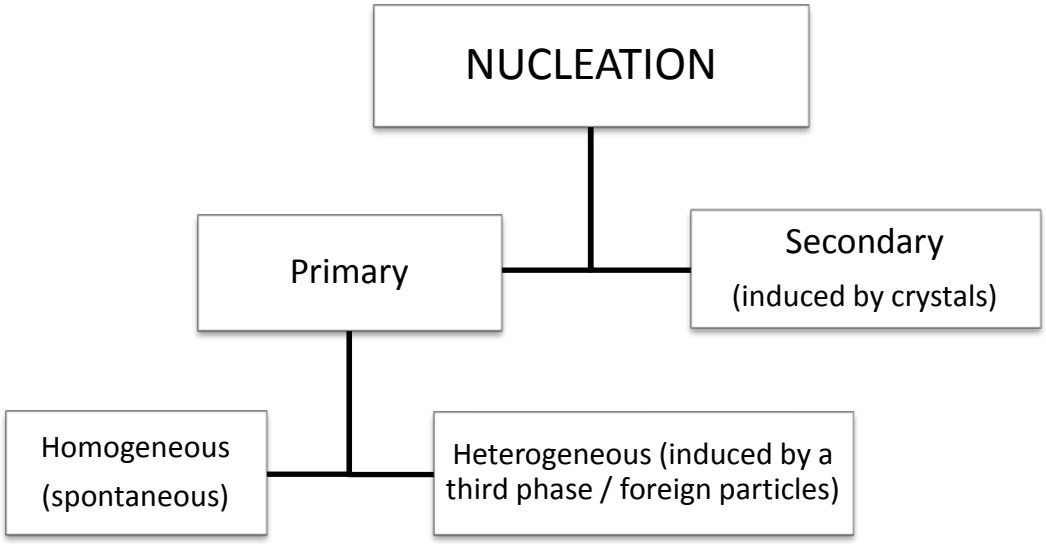

Figure 6 – Classification of nucleation (modified from [41])

Term “primary” imply nucleation in systems that do not contain crystalline substance. Accordingly, when formation of crystals takes place at the presence of formerly created crystals, the process is called secondary nucleation. A solution nucleates more readily when crystals of the solute are already present or intentionally added.

Homogenous nucleation (HON) can be defined as formation process of nuclei in the bulk of the solution free of impurities (only two phases involves, see Figure 7a). It cannot be stated with any degree of certainty exactly how a stable nucleus is formed within a homogenous fluid. Since the formation of a cluster requires many more molecules than could collide simultaneously, it is assumed that a sequence of bimolecular additions of an autocatalytic nature is more probable. The critical cluster size (also called critical nucleus) is the cluster size that must be reached before nuclei can grow spontaneously [3, 41].

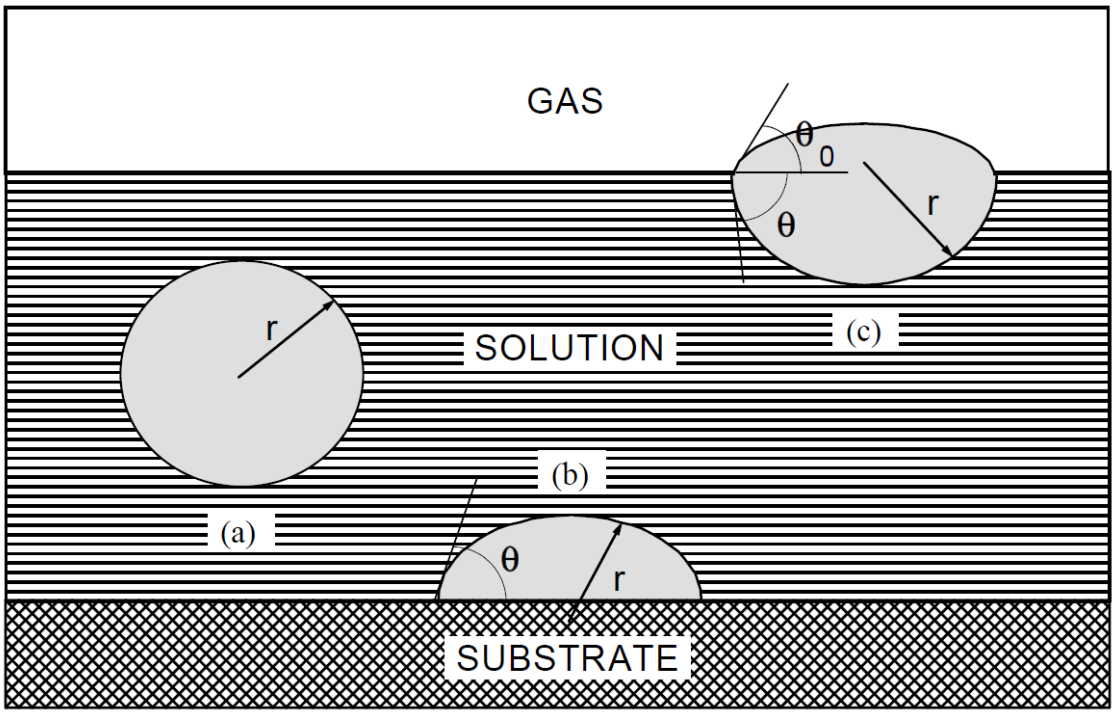

Figure 7 – Schematic of (a) spherical cluster of radius ***r*** in HON; (b) cap-shaped cluster in HEN on a substrate; (c) lens-shaped cluster in HEN at the gas/solution interface [8]

It is generally accepted that true homogenous nucleation is rarely encountered phenomenon in the real world since is practically impossible to achieve a solution completely free of foreign bodies For example, aqueous solutions as normally prepared in the laboratory may contain more than 10^6 solid particles per cm^3 of sizes less than 1 micrometer, although careful

filtration can reduce the numbers to less than 10^3 per cm^3 to obtain a solution more or less insensitive to spontaneous nucleation [3, 41].

The heterogeneous nucleation (HEN) takes place when more than two phases are in contact at site of nucleation (e.g. the forming nuclei – the solution – gas/a substrate/a foreign surface). HEN is much more widespread in both nature and technology than is HON, because, as has been said, the old phase practically always includes in itself foreign particles, or is in contact with other phases or container walls which limit it spatially. Thereby HEN does not require the energy to construct the surface for the clusters to grow on, in contrast with HEN: the surfaces of the other phases or of the various particles serve as places on which the formation of the clusters takes place [12, 42].

The vapor-liquid interface (or rather a film located at this interface) is often the most favorable site of hydrate formation, not only because the interface lowers the Gibbs free energy of nucleation, but also because the interface is the location of the required very high concentrations of host and guest molecules [3, 17].

2.7 Method of calculation of hydrate nucleation rate and lag time from probability distribution of induction times

The method to estimate the parameters of nucleation process was first proposed by Toschev et al. [9]. They stated that the formation of nuclei for one particular set of conditions can be treated as a sequence of random events along the time axis.

When the appearance of nuclei is independent, the probability of forming m nuclei in a time interval can be described by the Poisson formula [4]:

$$P_m = \frac{N^m}{m!}\exp(-N), \tag{1}$$

where N is the average number of nuclei that form during this time interval.

Consequently, the probability of the event “0 nuclei were formed” is then:

$$P_0 = \frac{N^0}{0!}\exp(-N) = \exp(-N). \tag{2}$$

By subtracting this probability from unity (total probability of all the events), we obtain the probability of having $m \geq 1$ nuclei:

$$P_{\geq 1} = 1 - P_0 = 1 - \exp(-N). \quad (3)$$

In turn, the average number of nuclei N can be determined as

$$N = JV\Delta t, \quad (4)$$

where J – nucleation rate, i.e. a number of nuclei which appear in a unit solution volume per unit time [38], $m^{-3}s^{-1}$ (or $m^{-3}min^{-1}$);

V – volume where the hydrate formation takes place, m^3;

Δt – time interval, s (or min).

In the present MSc study it is assumed that the volume where hydrate formation takes place is a very thin layer just below the interface between water and a gas mixture. The volume is function of contact area / interfacial area A between gas and water and the volume V could be approximated by A in Equation (4). It is also assumed that this volume is the same for every experiment for a particular cell since the dimensions of the cell do not change and the stirring rate is kept constant throughout the experiments. Therefore, in order to compare different experiments for a given cell, one can neglect the volume factor in Equation (4). Thus the estimated nucleation rates for the laboratory experiments will be measured in units of [min^{-1}]. If stirring was not kept constant, the contact area could be estimated as a function of stirring rate by measurements on water surface levels in open cell assuming smooth conical surface.

The formed nuclei have to grow to appropriate sizes before they can be detected by means of macroscopic instruments. This causes an offset or a lag time τ_0. Thus the time interval where we are looking for probability of nuclei occurrence can be introduced as a difference between induction time t and lag time:

$$\Delta t = t - \tau_0. \quad (5)$$

Finally, the probability to detect crystals at time t which were nucleated at an earlier time can be thus determined as:

$$P(t) = 1 - \exp\big(-J(t - \tau_0)\big). \quad (6)$$

On the other hand, for M isolated experiments, the cumulative probability $P(t)$ to measure an induction time between time zero and t can be found as [4, 5]:

$$P(t) = \frac{n(t)}{M}, \tag{7}$$

where $n(t)$ – number of experiments in which nuclei were detected within the time t,

M – total number of experiments.

Having the distribution of cumulative probabilities for induction times calculated by the Equation (7), one can then utilize curve-fitting algorithms in KaleidaGraph® in order to obtain hydrate nucleation rate J and a delay τ_0 in detection of a critical nucleus.

2.8 Method of calculation of activation energy and critical nuclei radius

Formation of hydrate nuclei may be described by means of classical nucleation theory, which was developed from the 1920s [49, 50]. According to the theory, the energy barrier that the system needs to overcome for nucleation to occur is considered as the excess in free energy due to the needed increase of energy by creation of a new interface and the decrease of energy from formation a more stable phase. In terms of Gibbs free energy ΔG such a phenomenon can be described by equation:

$$\Delta G = \Delta G_s + \Delta G_v = 4\pi r^2 \sigma + \frac{4}{3}\pi r^3 \Delta g_v, \tag{8}$$

where ΔG_s – Gibbs surface excess free energy, J;

ΔG_v – Gibbs volume excess free energy, J;

r – radius of nuclei, assuming spherical cluster, m;

σ – interfacial tension between the phases, N/m^2;

Δg_v – Gibbs free energy change per unit volume of hydrate bulk phase, J/m^3.

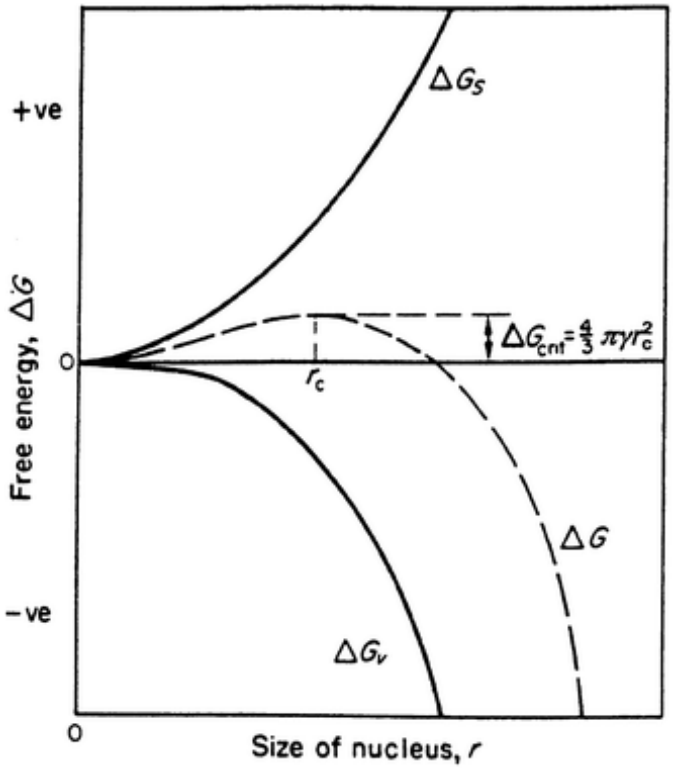

Figure 8 – Volume excess free energy $\boldsymbol{\Delta G_v}$ and surface excess free energy $\boldsymbol{\Delta G_s}$ which sum results in the energy barrier $\boldsymbol{\Delta G_{crit}}$ to form a stable nucleus of radius $\boldsymbol{r_c}$ [41]

With reference to Figure 8, the inflection point represents activation energy of formation a critical size nucleus with a radius r_c. In this point the derivative of Equation (8) with respect to r equals zero:

$$\frac{\partial(\Delta G)}{\partial r} = 8\pi r\sigma + 4\pi r^2 \Delta g_v = 0, \tag{9}$$

Solving it for Δg_v gives

$$\Delta g_v = -\frac{2\sigma}{r_c}. \tag{10}$$

Inserting this in Equation (9) and setting $r = r_c$ gives

$$\Delta G_c = E_a = \frac{4}{3}\pi\sigma r_c^2. \tag{11}$$

A cluster of size r_c has equal probabilities to grow and decay, and, hence, such clusters are called critical and they represent nuclei of a new phase [51]. Before achieving the critical size, clusters of molecules form the bulk metastable liquid, and these clusters may either grow or shrink as a result of density or composition fluctuations. When a certain energy level is attained

in the system and the cluster achieves a critical size, the growth of hydrate becomes self-sustaining [3].

The relation between activation energy and rate of nucleation can be found in Arrhenius formula, the widely used equation for description of temperature influence on the rate of chemical and other processes. For hydrate formation process it can be defined as [52]:

$$J = J_0 \exp\left(-\frac{E_a}{k_B T}\right), \tag{12}$$

where J_0 – pre-exponential factor (constant),

k_B – Boltzmann's constant, J/K;

T – absolute temperature of the system, K.

By taking logarithms of both sides and rearranging the equation, it can be rewritten as

$$\ln(J) = \ln(J_0) - \frac{1}{T}\left(\frac{E_a}{k_B}\right). \tag{13}$$

One then can easily find similarities with the equation of a straight line:

$$y = bx + c, \tag{14}$$

where $y = \ln(J)$ and $x = 1/T$ (see Figure 9).

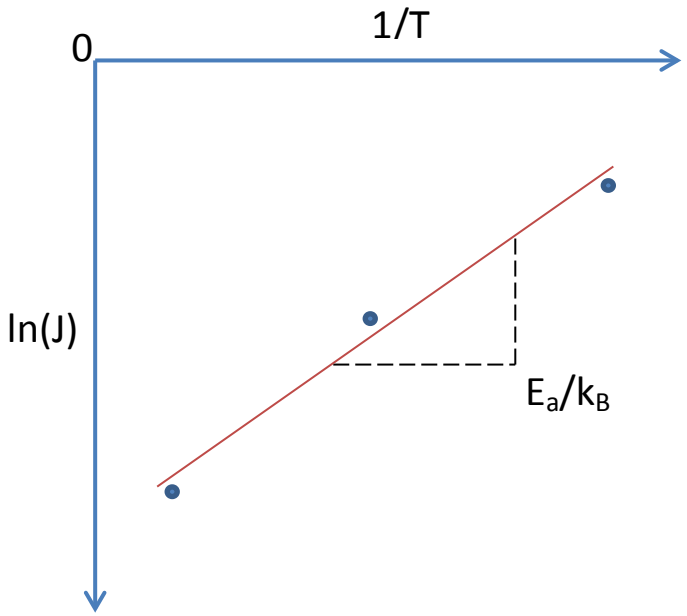

Figure 9 – Linear regression analysis for determination of activation energy of hydrate nucleation using Arrhenius equation

Thus plotting $ln(J)$ versus $1/T$ for values of J found experimentally at different temperatures should ideally (according to Arrhenius theory) give a straight line. Consequently, a slope of this line will represent E_a/k_B.

Thereby the activation energy E_a can be determined from the equation of a straight line that can be generated by means of linear regression in KaleidaGraph®. Multiplication of a slope with the Boltzmann constant k_B will give a desired value of the energy barrier.

In order to find a critical radius of nuclei, one then can use Equation (11), which after rearrangement gives:

$$r_c = \sqrt{\frac{3E_a}{4\pi\sigma}} \tag{15}$$

The critical radius represents a minimum size of a stable nucleus. Clusters smaller than r_c will dissolve or evaporate because only in this way the particle can achieve a reduction in its free energy. In turn, clusters with radius larger than r_c will continue to grow [41].

In absence of exact values for the specific surface tension on the interface between hydrate clusters and aqueous phase we assume it to be equal 0.0276 J/m^2 that corresponds to that of the interface between ice and liquid water.

2.9 Maximum Likelihood Estimator (MLE) method

The chapter is written in accordance with [53].

In addition to best-fit estimation of the nucleation parameters in KaleidaGraph®, Maximum Likelihood Estimation (MLE) technique was used in the present work.

MLE method is often utilized in order to find characteristics of a two-parameter exponential distribution, which has many real world applications. This type of distribution can be used to model the data such as the service times of agents in the system (queing theory), the time it takes before a piece of machinery breaks down, the time until a radioactive particle decays, the distance between mutations of a DNA strand, and the extreme values of annual snowfall or rainfall [53].

The principle of the method lies in selection of a set of values that maximizes the likelihood function. This method helps to find particular parameters for a confined number of observations, i.e. when one does not have a full set of data.

The likelihood function for the two-parameter exponential distribution is:

$$L(\theta, \eta \mid x_1, x_2, \dots, x_n) = \frac{1}{\theta^n}^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\eta)}, x_{1:n} \geq \eta, \quad (16)$$

where θ – continuous scale parameter, determines the statistical dispersion of the probability distribution;

η – continuous location parameter, determines the shift of the distribution;

$x_{1:n}$ – minimum observed value;

n – number of observations (experiments).

Under some regularity conditions, MLE method has good properties such as consistency and efficiency. But the regular MLE method is too conservative because it always chooses minimum of a sample to estimate a location parameter [53]. Therefore it was customized by introducing a penalty multiplier $(x_{1:n} - \eta)$ with a condition $x_{1:n} \geq \eta$. Such modification forces the likelihood function to deviate from monotonic behavior with respect to location parameter. The penalized maximum likelihood function is then:

$$L(\theta, \eta \mid x_1, x_2, \dots, x_n) = (x_{1:n} - \eta) \cdot \frac{1}{\theta^n}^{-\frac{1}{\theta}\sum_{i=1}^{n}(x_i-\eta)}, x_{1:n} \geq \eta. \quad (17)$$

By taking logarithm of the function, differentiation it with respect to θ and η, and setting the derivatives equal to 0, one can obtain penalized maximum likelihood estimators for θ and η:

$$\theta = \frac{n(\bar{x}-x_{1:n})}{n-1}, \quad (18)$$

$$\eta = \frac{nx_{1:n}-\bar{x}}{n-1}, \quad (19)$$

where $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ – sample mean.

The Cumulative Distribution Function, CDF for a two-parameter exponential distribution function looks like

$$F(x) = 1 - \exp\left(-\frac{x-\eta}{\theta}\right), \quad (20)$$

One then can easily find similarities with Equation (6) that describes the probabilities of nuclei detection. Comparing these two equations we have $J = \frac{1}{\theta}, \tau_0 = \eta$.

Thereby such characteristics of hydrate nucleation phenomenon as induction time and lag time can be calculated by means of MLE method.

2.10 Confidence interval

Confidence interval is a type of interval estimate of a population parameter and is utilized to indicate reliability of an estimate. The interval covers an unknown parameter of interest with a set limit of reliability.

When referring to a confidence interval, $(1 - \alpha)$ is known as a confidence level. This level describes the uncertainty associated with a sampling method. By saying "a 95% confidence interval of parameter" before the data are observed we can assert that with probability 0.95 the interval that will be obtained will contain mean value of the parameter, whereas after the data are obtained we can only assert that the resultant interval indeed contains mean with 95% confidence. This means that, if all possible means of size n were taken from the population and a 95% confidence interval were calculated from each sample, 95% of those intervals would contain the mean value of the parameter [54, 55].

The $100(1-\alpha)\%$ confidence interval for the scale parameter θ of an exponential distribution is given by [54, 56]:

$$\frac{v\bar{x}}{\chi^2_{1-\alpha/2,v}} < \frac{1}{\theta} < \frac{v\bar{x}}{\chi^2_{\alpha/2,v}}, \tag{21}$$

where $\chi^2_{p,v}$ is the 100(1 – p) percentile of a chi squared distribution,

v – degrees of freedom, $v = 2n - 2$ for a two-parameter exponential distribution;

$\bar{x}$ – sample mean,

n – number of elements in the sample.

It is understandable that small standard deviation of the mean will result in a smaller confidence interval, meaning that mean value is estimated more precisely when standard deviation is small. In turn, standard deviation becomes small as number of elements in the

sample becomes large. Thereby, in general, a parameter estimate from a large sample is more precise than an estimate of the same parameter from a small sample [55].

In summary we can conclude that a narrower confidence interval normally associated with a larger sample size, a smaller standard error, a smaller confidence coefficient $(1 - \alpha)$ [55].

In the present study 95% confidence intervals for a nucleation rate have been calculated by means of R Studio® software. A listing of the program for constructing the confidence interval for MLE method is given in Appendix A. The confidence interval for J estimated by KaleidaGraph® best-fit technique was obtained by inserting this value into the program instead of a formula for calculation of nucleation rate by MLE method.

2.11 Permutation test

Permutation test is the oldest of all nonparametric procedures that are still widely applied today. The use of this method dates back to at least Fisher in 1935 [57, 58].

Nonparametric tests (or distribution-free tests) do not make any assumptions about the distribution of the data. These tests rank the data from low to high and analyze the rank. In general, a nonparametric test is used if the data represents a rank or score, or if the measurements are drawn from a population that is not normally distributed [59].

Permutation test is widely applied to compare two datasets and determine whether they are parts of one distribution. The basic idea embodied in permutation methods is to generate a reference distribution by recalculating a statistic for many rearrangements of the data [57].

As with any testing of statistical hypotheses we need the following elements: data, null hypothesis, test statistic and the sampling distribution of the test statistic under the null hypothesis. The null hypothesis includes the assumptions under which probability distribution for the data holds. It normally implies that the two classes have identical probability distributions. The sampling distribution of the test statistic is a set of statistic values with equal probabilities corresponding to all the possible replacements of labels of the initial randomization scheme [58, 60, 61].

The goal of permutation test is to determine whether the observed difference between the sample means is large enough to reject the null hypothesis at a defined acceptable significance level. The significance level α defines the maximum acceptance probability of false positive, i.e., it declares that the classes are different if the null hypothesis is true [60].

As the result, the permutation test calculates p-value that is the highest level of significance α at which the null hypothesis can still be rejected [60, 61]. This value determines if there is a statistically significant difference between the median of the sample and the hypothetical mean. If the value is below a certain level (usually 0.05), one can conclude that there is a significant difference between these means [59].

The number of sampling iterations one should perform is always a trade-off between precision and computer time since the number of permutations with n objects is n! (n-factorial). For small datasets, one can compute all possible permutations in a systematic way and obtain the complete permutation distribution of the statistic. However, for large datasets, only a part of all possible permutations can be computed since there are too many. The more permutations the better, but it may be unacceptable to wait for the permutations results when studying a large dataset. Therefore it is important to select the number large enough to guarantee accurate estimation. In many cases the number of 10000 permutations is set for the large datasets as one that gives a reliable result [58, 60].

A permutation test in the present study was run in order to compare two parts (groups) of an experimental series. A significance level was set to be 5%. Listing of the program code for execution of a permutation test in R Studio® is given in Appendix A.

2.12 Treatment of experiments with non-occurrence of hydrates

During the laboratory experiments on the large cell the situation occurred when hydrates did not form during a long time period. It was decided to stop these experiments after exceeding approximately 1000 minutes due to the limit of time. Probability to form hydrates after that time still exists, therefore the observation time value may be included in the dataset of induction times. In that case we have so-called incomplete dataset that cannon be treated according to the scheme described previously.

In the literature different techniques of treatment of the incomplete datasets can be found, going from naive ones, like ignoring records with unknown entries, to more sophisticated methods which take into account the fact that those data with missing values might be relevant to our purpose [62].

If the “unsuccessful” experiments have to be included in analysis, the Kaplan-Meier estimator can be used in order to get the probability distribution of such datasets with so-called censored observations.

The Kaplan–Meier estimator developed by Kaplan and Meier [63] in 1958 for censored lifetime data analysis, is the nonparametric maximum likelihood estimate of the so-called survival function [64]. Thus it measures probability that the failure in the system will not occur beyond a specified time.

For non-censored observations, the Kaplan-Meier estimator is identical to the regular MLE. The difference occurs when there is a censored observation – in that case the Kaplan-Meier estimator takes the "weight" normally assigned to that observation and distributes it evenly among all observed values to the right of the observation. This is reasonable because we know that the true value of the censored observation must be somewhere to the right of the censored value, but we don't have any more information about what the exact value should be [64].

If we have observations within a time interval t, then the Kaplan-Meier estimator gives a cumulative probability of the form

$$F_{KM}(t) = 1 - \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}, \tag{22}$$

where n_i – number of observations up to time t_i minus the number of censored cases;

d_i – number of events occurred at time t_i.

Thus, for our case the cumulative probability $P(t)$ to measure induction time before time t can be determined as:

$$P(t) = 1 - \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}, \tag{23}$$

where d_i will represent the number of hydrate detections at time t_i, i.e., $d_i = 1$ for successful experiments, whereas $d_i = 0$ for experiments with non-occurrence of hydrates.

For analyzing by MLE method the experimental series with incomplete dataset were treated according to censored scheme for incomplete data, particularly Type-II hybrid censoring. This scheme guarantees that at least r events are observed when the experiment terminates at a random time $t^{*}=\max(t_{r:n},t)$. The scale and location parameters are [53]:

$$\theta = \frac{nrx_{1:n}-\sum_{i=1}^{n} x_{i:n}-(n-r)x_{r:n}}{n(r-1)}, \tag{24}$$

and

$$\eta = \frac{-nx_{1:n}+\sum_{i=1}^{n} x_{i:n}+(n-r)x_{r:n}}{r-1}. \tag{25}$$

where n – total number of experiments;

r – number of failures (in our case it is number of hydrate detections);

$x_{1:n}$ – time minimum, s (or min);

$x_{r:n}$ – time spent until termination of observation (for experiments with non-occurrence of hydrates).

As for the case of penalized MLE method for complete datasets, the nucleation rate and the lag time are $J = 1/\theta$ and $\tau_0 = \eta$.

3 EXPERIMENTAL SECTION

3.1 Experimental setup

A schematic outline of the experimental setup used to obtain necessary data in the present MSc study is presented in Figure 10.

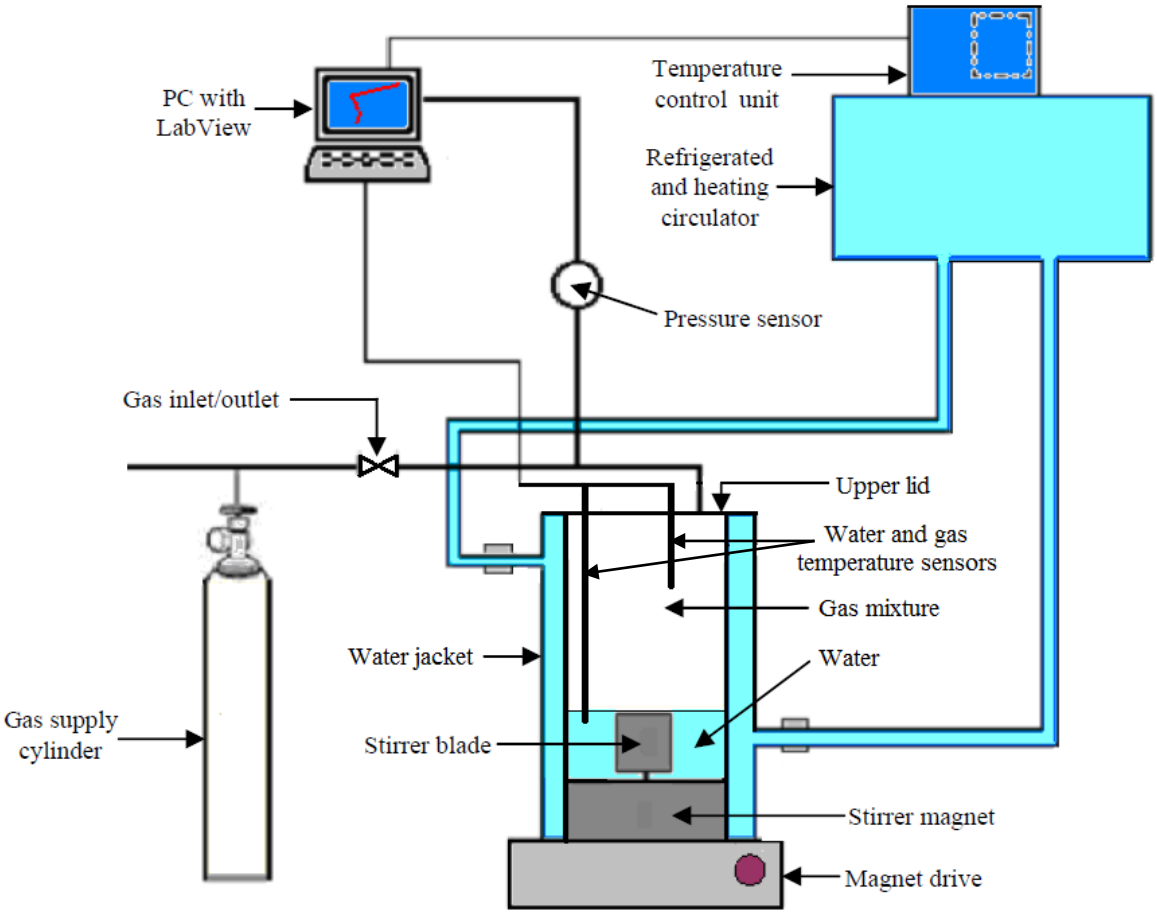

Figure 10 – Outline of the experimental installation used in the present work (modified from [11])

A main part of the experimental installation is a high-pressure autoclave cell. Experiments in the present MSc work were conducted on two cylindrical cells with similar construction but different dimensions. The main characteristics of the cells and some parameters of the experiments on them are presented in Table 1.

Table 1 – Main characteristics of the setup and experiments by cells

Characteristic	Large cell	Cell #3
Inner diameter, mm	90	60
Inner volume, ml	318.1	141.4
Volume of water used, ml	112.5	50.0
Type of a gas mixture	SNG-7	SNG-2
Stirring rate, rpm	450	730
Initial temperature, °C	23	23
Model of a refrigerated - heating circulator	Julabo F33 HL	Julabo F34 HL

The main components (Figure 11) of each autoclave cell are the metallic (titanium) cylindrical body, the stirring blade, the magnet, bottom and upper lids. Two rubber O-rings were used to provide sealing between the lids and the cell body. The temperature sensors for water and gas temperature measurements Pt 100 of class 1/10 (DIN) provide measurement accuracy within ±0.03 °C at 0 °C. The pressure inside the cell was measured using Rosemount 3051 TA gauge with accuracy within ±0.025 bar up to 100 bar.

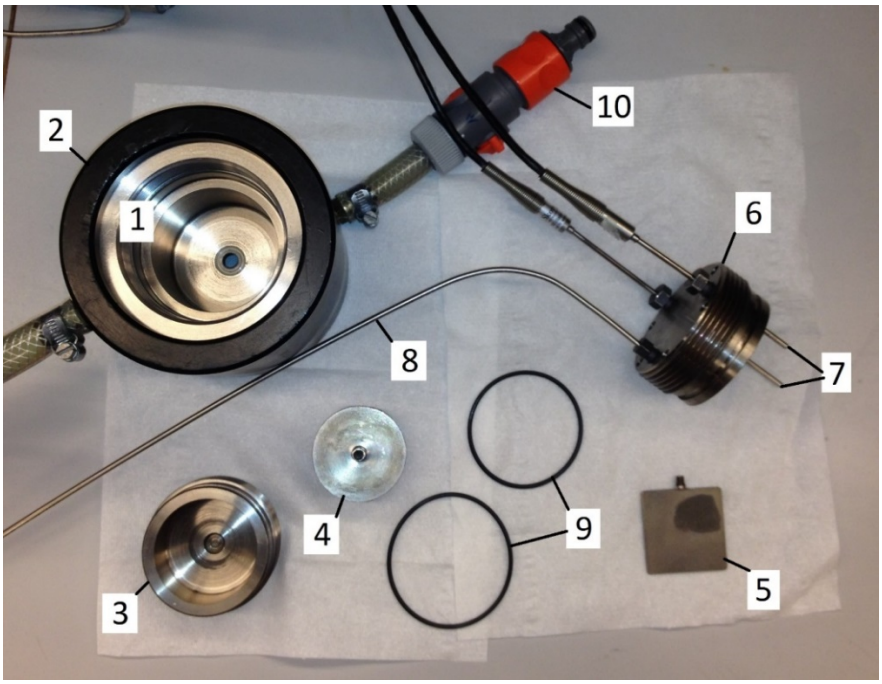

Figure 11 – Cell components: 1 – cell body, 2 – plastic jacket, 3 – bottom lid, 4 – magnet with bearing, 5 – stirring blade, 6 – upper lid, 7 – temperature sensors, 8 – gas supply tubing, 9 – upper and bottom O-rings, 10 – water hose with quick connect.

Both cells are packed into an outer plastic jacket that during an experiment was connected with the refrigerated and heating circulator via the inlet and outlet flexible pipes. Thus the temperature regulation inside the cell was provided by means of heat exchange with distilled water as a cooling-heating agent.

Two different programmable refrigerated and heating circulators were used: Julabo of model F33-HL for the large cell and F34-HL (see Figure 12) for cell #3. Both of them have a working temperature range of -30…150 °C and the temperature stability of ±0.01 °C Integrated programmers are able to contain up to 60 regulation steps per regulation loop, and the loop can be repeated up to 6 times. The main difference between the circulators is an operating liquid volume of 20 liters for F34 and 16 liters for F33-HL.

Figure 12 – Refrigerated - heating circulator Julabo F34-HL. The electronic timer is attached to the controlling block.

Water-gas mixture was agitated in the cell by means of the magnetic stirrer with an option of controlling rotation speed. Besides manually, the stirrer on/off function has been performed by the electronic timer.

Temperatures of bath, water and gas phase, as well as pressure in the cell were constantly monitored with a time step of 3 s by means of LabVIEW® program installed on a personal computer.

3.2 SNG-2 and SNG-7 gas mixtures

Laboratory experiments in the present MSc study were conducted using distilled water and two different gas mixtures:

- 2-components Synthetic Natural Gas SNG-2,
- 7-components Synthetic Natural Gas SNG-7.

The compositions of both mixtures are presented in Table 2.

Table 2 - Composition of SNG-2 and SNG-7 gas mixtures used in the present study

Specific component	Composition [mol%]	
	SNG-2	SNG-7
Methane	92.5	80.40
Ethane	-	10.30
Propane	7.5	5.00
i-butane	-	1.65
n-butane	-	0.72
Carbon dioxide	-	1.82
Nitrogen	-	0.11
Total	100	100

The pressure-temperature equilibrium curves for SNG-2 and SNG-7 mixtures, calculated in CSMGem® are shown in Figure 13. Above this curves the mixture is in metastable region. The temperature was initially set to be 23 °C, which was calculated to be outside the hydrate formation region for pressures more than 100 bars for both SNG-2 and SNG-7. During the induction period, the temperature and pressure were kept inside the metastable region. The equilibrium temperatures at 90 bars are 21.85 °C and 20.97 °C for SNG-2 and SNG-7 respectively.

From the personal communication with Thor Martin Svartås [20], the academic supervisor of this master's thesis, it was known that CSMGem® program gives accurate results on SNG-2. There were several measurements run in the UiS lab for equilibrium temperatures at

30, 45, 60, 90, 120 and 175 bar for this mixture which show that agreement between predicted and measured data is better than ± 0.2 °C.

However, CSMGem® prediction of equilibrium temperatures for SNG-7 is very poor and experimental data indicates discrepancy up to + 1.7 °C at 65 bar. In order to verify correct equilibrium properties of SNG-7 experimental measurements are currently running.

Thus the plot for SNG-7 in Figure 13 gives incorrect data, which will further affect the quantification of the degree of subcooling. However, incorrect estimation of subcooling does not influence the experimental results.

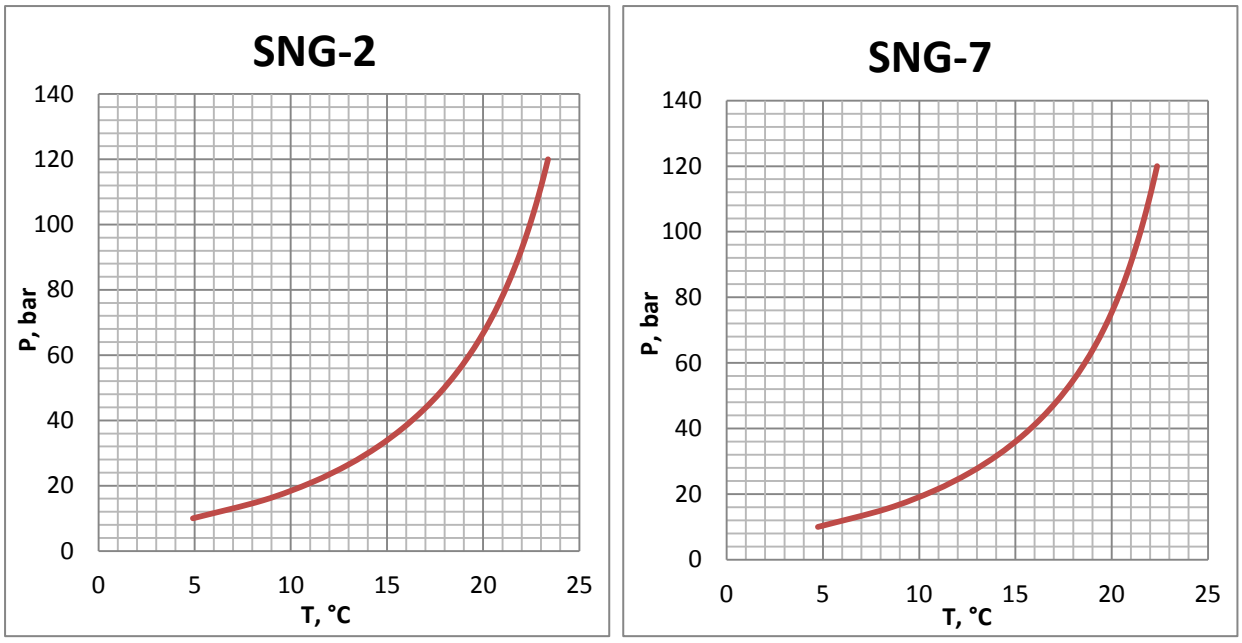

Figure 13 – P-T equilibrium curves for water and SNG-2, water and SNG-7 calculated in CSMGem®

3.3 Experimental procedure

Every experimental procedure started from preparation of the cell. In order to eliminate the memory effect that could be caused by potential residuals of hydrate structures from previous experiment, the cell and all of the components were washed with brush under a stream of tap water. After that every detail was rinsed with distilled water again and dried by pressurized air.

Upper and lower O-rings were lubricated by high vacuum silicone grease to protect them from drying and wearing while screwing the lids. Also small amounts of copper paste were

applied on the threads, as required, to prevent their mutual adhesion. The O-rings then were mounted on the lids. The stirring blade was placed inside the cell through the connection with the magnet.

Only for cell #3 a 0.5 mm thick nylon gasket was placed to the recess for magnet holder bearing in the bottom lid in order to lift the stirring blade and prevent scraping of the cell inner surface. Then the bottom lid was filled with approximately 15 ml of distilled water and screwed in the cell. The water yield through the blade bearing indicated absence of air in the lower chamber.

After that the cell was filled with required amount of pure distilled water depending on the cell size. The upper lid was mounted, water hosepipes from the bath were attached to the plastic jacket through the fast connects and the cell was ready for gas filling.

The next step was lowering air concentration inside the cell by purging it with gas. The pressure inside the cell was adjusted by the system of valves on the gas bottle. After increasing the pressure in the cell to approximately 40 bars, the stirrer was turned on in order to relieve air extraction from water. After pressure stabilizing, the stirrer was stopped before the cell was depressurized to avoid ejection of water droplets in the gas supply tubing. Having reached ambient pressure, the stirrer was restarted to remove residues of supersaturated gas in the water phase. The procedure was repeated once more to decrease the air concentration in a gas mixture for 1600 times (40×40).

The initial temperature was set to be 23 °C. As the cell temperature is lowered the pressure decreases, principally due to gas contraction as well as increased gas solubility upon cooling at constant volume. Therefore, in order to reach the final experimental pressure of 90 bar after cooling the system down to the desired temperature of the experiment, the initial pressure in the cell was increased above 90 bar. The cell charging pressure for every experimental series was calculated utilizing a formula:

$$P = P_{exp} + GradP \cdot \Delta T', \tag{26}$$

where P_{exp} – experimental pressure, P_0 = 90 bar,

$GradP$ – pressure gradient, specific for the gas mixture and the cell,

$\Delta T'$ – difference between the initial temperature and the temperature of the experiment.

After achieving the required pressure, the connection between the gas supply tubing and the upper lid was checked for possible gas leaks.

Allowing the pressure inside the sell to stabilize, the appropriate program in the programmable module of the refrigerated-heating circulator was started simultaneously with a LabVIEW® program that generated pressure-temperature log. The timer was turned on 15 second after that.

According to each program, after a relaxation time of 10 minutes without stirring, the container was cooled at a rate of 6°C/h from 23 °C down to the desired experimental temperature. Having reached the desired bath temperature, the system was allowed to rest for 10 minutes to reach thermal equilibrium in the cell. After that agitation was started with keeping the temperature constant throughout the experiment.

The experiment was stopped after hydrate formation was detected. In order to dissociate hydrates, the content of the cell was warmed up slowly to achieve the initial temperature of 23 ºC which causes the absorbed gas to release. The pressure achieved approximately the initial value indicated that there are no hydrates in the cell. After that the pressure was reduced to ambient by discharging the gas from the cell. The installation was ready to be prepared for another experiment.

Detection of characteristic events such as start of agitation and hydrate occurrence was performed in KaleidaGraph® by analyzing the logs generated in LabVIEW®.

Since formation of gas hydrates is an exothermic process [58], i.e. heat is released during nucleation and consequent growth, a moment of time just after nuclei formed and grew enough to be detected can be determined by temperature increase. But in the present work another method to find this moment was used. Small but distinct pressure spikes on the log in the beginning of hydrate formation can be utilized to increase accuracy of data readings. It was assumed that the most correct time is measured through these pressure pulses because they are giving the quickest response in comparison with the temperature reaction on the process changes. As can be seen in a plot of pressure vs. time (Figure 14), there was a slight rise in pressure both at start of stirring (time of beginning of the experiment) and before the first detection of hydrates by pressure drop in the cell.

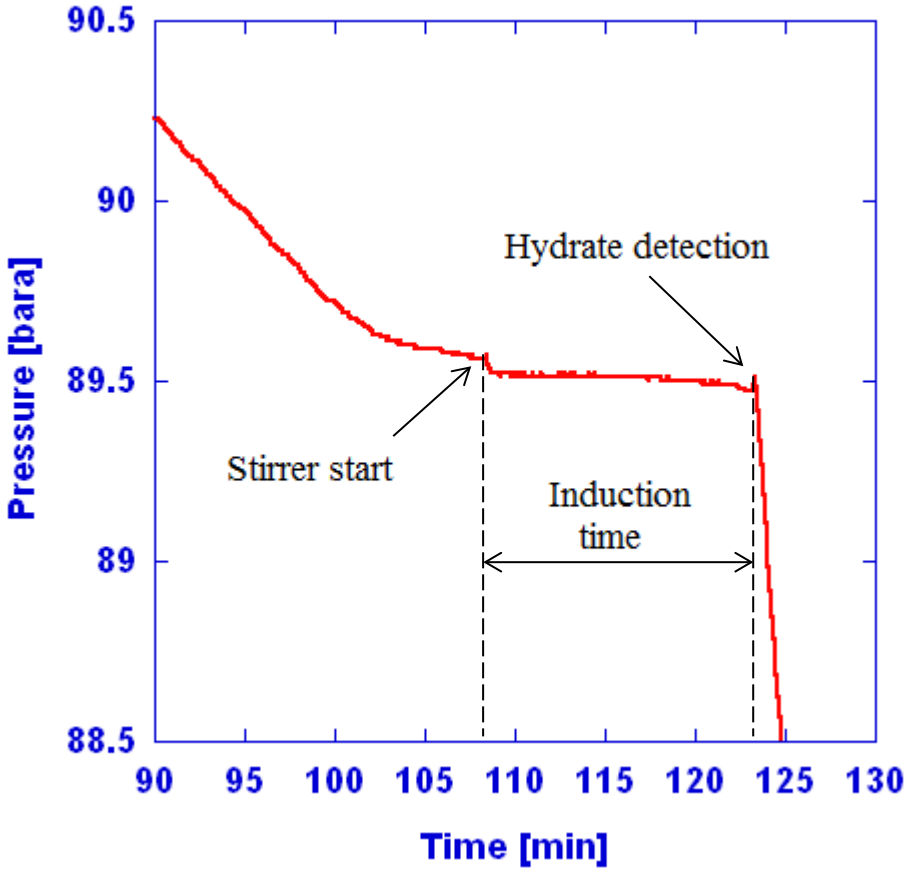

Figure 14 – Pressure versus time plot obtained from LabVIEW® log, where the moments of stirrer start and hydrate detection are identified and further used for the induction time calculation

However, in case of absence of pressure readings, in the experiments where the pressure sensor stopped to respond, gas temperature readings were used since the similar pulses could be detected on their logs. But in these cases the events of the interest were assumed to occur 4 time steps (12 s) before the pulse. When trying to detect the beginning of stirring, the readings of rotational speed of the stirrer were used in addition to the temperature response. The stirrer was assumed to be switched on also 4 time steps before changing the rate from zero to a positive value in the log.

3.4 Experimental analysis description

The induction time of every experiment was obtained in KaleidaGraph® program by analyzing the data from the logs generated in LabVIEW®.

Then the induction periods for each experimental series were ranked in ascending order, and the cumulative probabilities of detecting hydrates within those periods were assigned in accordance with Equation (7) for complete datasets and Equation (23) for incomplete datasets.

Utilizing the curve fitting tool in KaleidaGraph®, estimations of nucleation rate J and lag-time τ_0 were obtained. This method will be referred to as KaleidaGraph® best-fit.

The same parameters were calculated in Microsoft® Excel for MLE method by means of Equations (18) and (19) for complete induction time datasets (J was calculated as the opposite of θ). For experimental series contained the time intervals with non-occurrence of hydrates, the parameters were obtained according to Equations (24) and (25).

The confidence intervals for J were calculated in R Studio® according to the program presented in Appendix A.

Having the values of the nucleation rates for 3 different temperatures, the activation energy was calculated according to the theory of finding the activation energy through an Arrhenius plot based on Equation (13). Consequently, the radius of a critical nucleus was calculated using Equation (15).

For the data analysis specific for this MSc thesis approach was used, according to which we assume that the case including all the available experiments for particular series gives the most accurate result and we can consider it as a reference for the further analysis. Having particular number of datapoints (induction times), we start to reduce the number of experiments by groups of 6 (or 2 for small datasets) calculating the parameters of interest at every such step. The lower limit is considered equal 6 experiments, because it is a minimum logical amount of observations sufficient for parameters estimation of a phenomenon with stochastic nature [20].

The minimum number of parallels required to determine particular parameter is dependent on the acceptable accuracy of estimation. Test criteria were set to maximum deviation of 20 and 30 % from estimated value obtained with the maximum number of parallels included. In other words, we will consider the estimate of a parameter to be acceptable if the deviation in the value from the most accurate result (reference level) is in the range 0 – 20% or 0 – 30% respectively for two limits. Analyzing a trend of parameter variation, we are looking for an interval where the trend first intersects the accuracy limit in order to find an exact number of experiments with deviation equal or less than the set limits.

4 RESULTS AND DISCUSSION

Depending on temperature, there were 3 series of experiments conducted on each of the cells. Independently of the cell, every experiment was carried out at an experimental pressure of 90 bar and a cooling rate of 6 °C/h down to the experimental temperature.

The initial part of cell#3 experiments was performed by Aina Undersrud Bratland the same semester as part of her BSc thesis work. Table 3 shows the number of parallels at each experimental condition and their order of performance when data contains experiments from both the present MSc and Aina Bratland's BSc study.

Table 3 – Number of experiment by experimental series and performers

Temperature, °C	Number of experiments		
	Carried out by Aina	Part of the present MSc study	Total
Cell #3			
11.75	**20** (# 1-20)	**10** (# 21-30)	**30**
13.00	**25** (# 1-24, 46)	**35** (# 25-45, 47-60)	**60**
14.25	**12** (# 1-12)	**18** (# 13-30)	**30**
Large cell			
15.50	-	**12**	**12**
16.75	-	**20**	**20**
17.75	-	**12**	**12**

4.1 Nucleation rate and lag time calculation and analysis

The results and analysis will be presented by cells and experimental series.

4.1.1 SNG-2 cell #3 data analysis

As was mentioned previously, the first part of experiments in cell#3 was carried out by Aina Bratland. First Aina performed 6 experiments for every series on SNG-2 in cell#3. After the pause caused by gathering the data for another cooling rate, she conducted several more experiments to give 12, 24 and 20 experiments in total for series of 14.25 °C, 13.00 °C and 11.75 °C respectively.

When Aina finished her part of work, the experiments of the present MSc commenced continuing on the series at 13.00 °C in order to produce a dataset containing 60 parallels in total for the further analysis. Then the series at 11.75 °C and 14.25 °C were continued to obtain a total number of 30 parallels at each of these temperature conditions.

The arrangement of the time periods of conducting the experiments by series is presented in chronological order in Figure 15. This information could be important if the measurements were influenced by some unknown condition due to e.g. wear and tear of equipment or unknown effects introduced through needed maintenance during the experimental period.

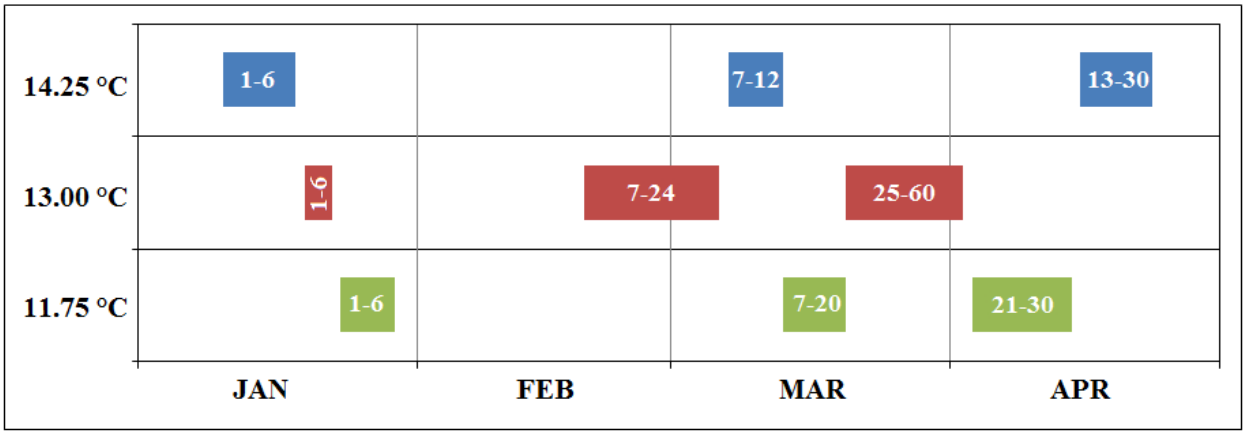

Figure 15 – Chronology of experiments on SNG-2 in cell #3

4.1.1.1 Series with 60 experiments on SNG-2 at 13 °C and 90 bars

There were 60 experiments performed at the temperature of 13 °C in order to obtain relatively large number of parallels for the further analysis.

The driving force represented by subcooling at 13.0 °C and 90 bars is:

$$\Delta T = 21.85 - 13.00 = 8.85 \text{ °C}.$$

All the measured induction times are shown in chronological order in Figure 16. The induction times measured covered a range from 0.35 minutes (i.e. 21 seconds) to 322.45 minutes with the majority being less than 1 hour and only 3 measurements above (experiment #2, #7 and #18).

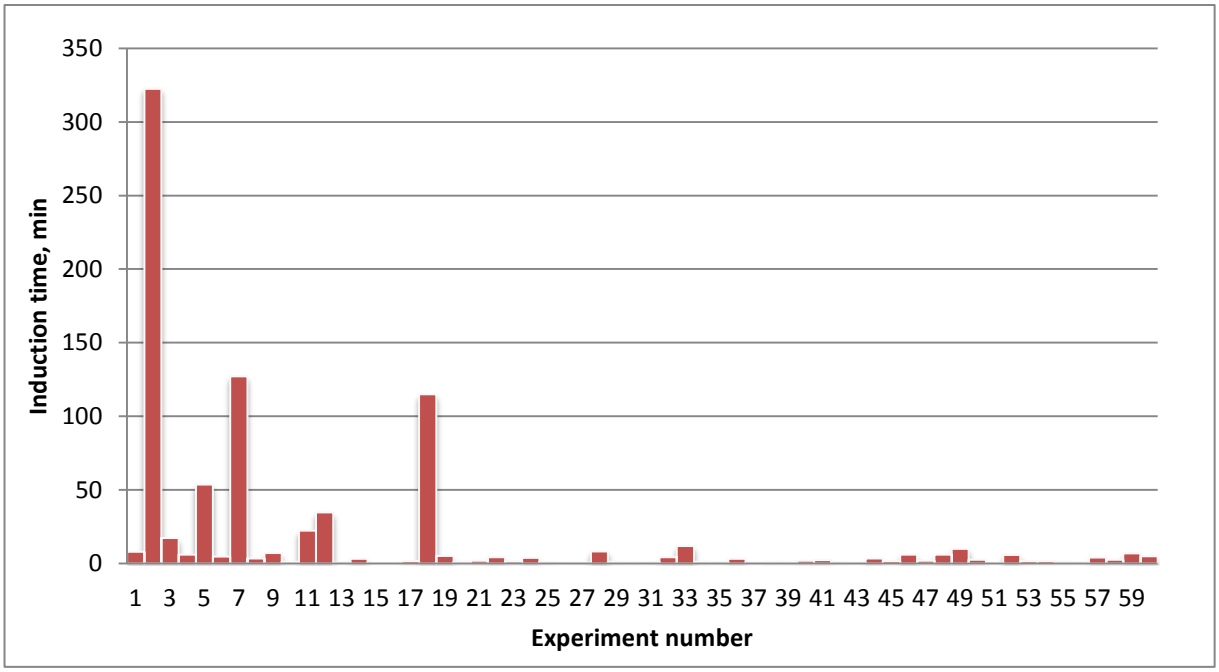

Figure 16 – The induction time in 60 experiments for SNG-2 in water at 13.0 °C

At first sight strong inconsistency in the series in Figure 16 may be noticed. Although gas hydrate nucleation is highly stochastic phenomenon, the great variation in induction times could be caused by other reasons since all the long induction times are in the beginning.

Looking back to Figure 15, all the experiments could be sub-sectioned into 3 groups by time periods they have been conducted within. One may see that all the long induction times are in the first and in the second groups.

However, there is no reason to state it was the personal influence since:

1) The first experiment conducted in the present MSc work was performed with Aina participating, checking that all the parameters and the sequence of operations were in accordance with her previous experimental procedure;

2) Aina performed one of the experiments (#46) after the first part of the present experiments, were carried out to check whether the long induction times could be due to some effect influenced by the person conducting the experiment or not. This check gave low induction time and no indication that the long induction times experienced were due to personal effect.

The real reason of such unexpected results cannot be stated with 100% reliability. However, the assumption that the cause was due to some unknown effect in equipment / system

looks the most reasonable. Since not any part of the equipment was modified through maintenance or replaced (except the O-rings), the reason can be in wear and abrasion of the rotating parts of the setup. Microscopic particles could be introduced in the water-gas mixture through the bearing of the magnet holder and promote the nucleation as the centers of crystallization. Furthermore, the wear and tear of the bearing could also cause loosing of the stirrer blade which in turn results in vibration. Mechanical shock was declared as one of the impacts that promote nucleation [41].

It was decided to run permutation test in order to compare different parts of the distribution of induction times and find probability of their belonging to the same distribution region. All the experiments were divided by groups of 6, and the groups were numerated following the chronological order. The results of the permutation test having 5% significance level are presented in Table 4. *p*-values less than 0.05 are colored red.

Table 4 – Permutation test for 60 induction times divided by 10 groups contained 6 experiments at a temperature of 13 °C, red color indicates p-value is less than 0.05

Group number	1	2	3	4	5	6	7	8	9
2	0.3601								
3	0.2321	0.2286							
4	**0.0023**	**0.0199**	0.1933				**< 0.05**		
5	**0.0029**	**0.0183**	0.2429	0.4788					
6	**0.0167**	**0.0388**	0.3240	0.3042	0.2686				
7	**0.0000**	**0.0068**	0.1183	0.0850	0.4792	0.1488			
8	**0.0129**	**0.0458**	0.4582	0.1811	0.2292	0.4756	0.3840		
9	**0.0207**	0.0512	0.4482	0.1802	0.1969	0.4136	**0.0475**	0.3693	
10	**0.0060**	**0.0408**	0.3790	0.2303	0.2542	0.4224	0.0568	0.4074	0.3243

There may be some conclusions made from the permutation test:

The first two groups are significantly different from the others since the induction times they contain are long compared to the other groups. According to the permutation test groups 1 and 7 are not interrelated at all since group 7 contains the smallest values out of all groups and group 1 – the largest. Groups 7 and 9 also produce p-value less than 0.05 because the group 9 contains induction times of the greatest magnitudes among the groups 3-10.

Permutation test was also conducted for groups related to the time periods the experiments were carried out (let them be numerated I, II and III). The p-values are 0.1095, 0.0037 and 0 for the test comparing the groups I + II, II + III, I + III respectively, which shows that there is significant difference between the probability distribution of group III and two others. Nevertheless, all the experiments were treated as one sample.

The results of estimation of nucleation rate J and lag time τ_0 by both KaleidaGraph® best-fit and MLE methods are presented in Table 5. A 95% confidence interval for the nucleation rate and absolute deviation of J from the reference level (i.e. the case based on 60 experiments) are also included in Table 5 to demonstrate the effect of the number of parallels on the accuracy. In this table and for tables that appear later in the text the minimum acceptable number of experiments as well as corresponding values will be colored red for 20% accuracy level and colored green for the level of 30%. If the minimum number is the same for two levels, only red color will be used.

Table 5 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for the experiments at 13.0°C on SNG-2 in water

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev., %	Conf. interval	τ_0, min	J, min^{-1}	Dev., %	Conf. interval	τ_0, min
60	0.2650	0	[0.2018 , 0.3368]	0.096	0.0706	0	[0.0538 , 0.0897]	0.114
54	0.2624	0.989	[0.1966 , 0.3376]	0.039	0.0649	8.084	[0.0486 , 0.0835]	0.065
48	0.2585	2.471	[0.1899 , 0.3375]	0.015	0.0590	16.43	[0.0434 , 0.0770]	-0.003
46					0.0570	19.25	[0.0416 , 0.0749]	-0.031
45					0.0562	20.47	[0.0408 , 0.0739]	-0.046
42	0.251	5.316	[0.1801 , 0.3334]	-0.063	0.0527	25.44	[0.0378 , 0.0700]	-0.102
40	0.2306	12.99	[0.1640 , 0.3084]	-0.155	0.0503	28.82	[0.0357 , 0.0672]	-0.147
39	0.2187	17.49	[0.1548 , 0.2935]	-0.230	0.0491	30.50	[0.0347, 0.0659]	-0.172
38	0.2092	21.06	[0.1473 , 0.2818]	-0.261	0.0478	32.26	[0.0337 , 0.0615]	-0.200
36	0.1926	27.32	[0.1342 , 0.2615]	-0.278	0.0453	35.82	[0.0316 , 0.0615]	-0.263
30	0.1726	34.89	[0.1156 , 0.2408]	-0.442	0.0386	45.33	[0.0259 , 0.0539]	-0.513
24	0.1379	47.98	[0.0874 , 0.1997]	-0.399	0.0311	56.02	[0.0197 , 0.0450]	-0.991
18	0.0647	75.60	[0.0377 , 0.0988]	-2.743	0.0235	66.75	[0.0137 , 0.0359]	-2.016
12	0.0437	83.52	[0.0218 , 0.0730]	-2.552	0.0185	73.77	[0.0093 , 0.0310]	-3.238
6	0.0637	75.97	[0.0207 , 0.1304]	0.075	0.0130	81.54	[0.0042 , 0.0267]	-7.984

One can notice a tendency of shifting the confidence intervals towards smaller values with the decreasing number of experiments. This behavior is explained by decreasing the mean value of the distribution, which is the estimated nucleation rate. The other general trend is in shrinking of the interval range with minimizing the number of experiments. This is quite an unusual behavior since a larger sample size normally leads to a better estimate of the population parameter. In other words, a confidence interval is supposed to shrink with increasing number of experiments [65]. However, since all the longest induction times detected are in the beginning and just small values are in the end (see Figure 16), eliminating the later values results in increasing of the significance of former ones, whose high magnitudes will influence the estimated nucleation rates more and more.

Figure 17 shows nucleation rate J as function of the number of parallels included in the estimation as obtained through KaleidaGraph® and MLE analysis. In this figure thin lines indicate confidence interval limits while thick lines are fitted to the estimated values.

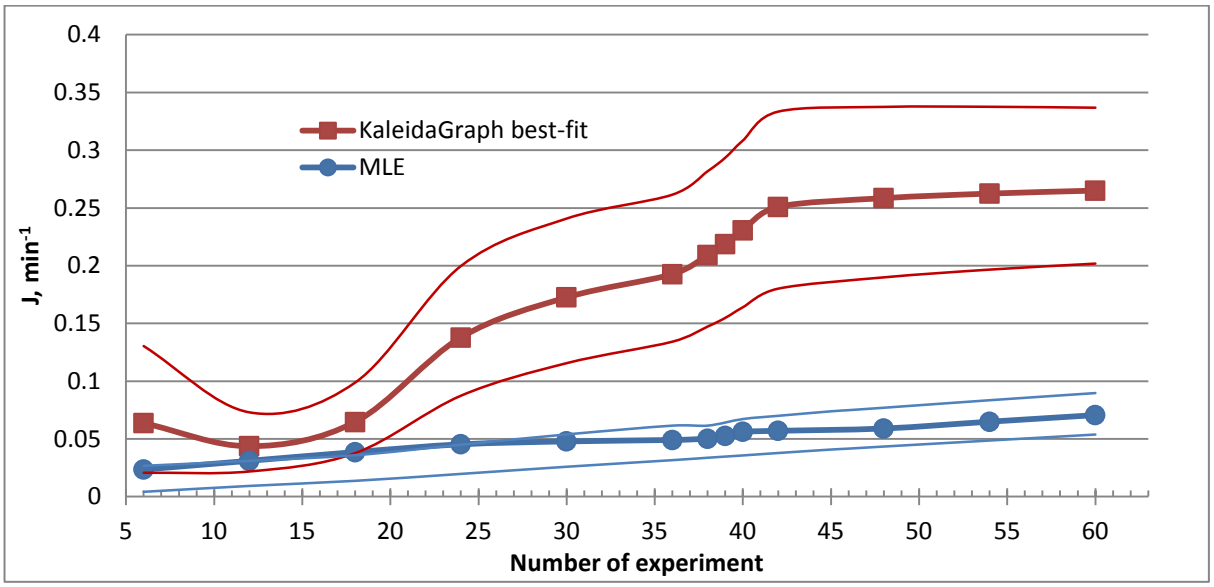

Figure 17 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 13°C

As indicated in Figure 17 KaleidaGraph® trend obtains an apparent steady state behavior after achieving the number of parallels equal 40. Thereby it can be assumed that one will not gain a big increment in accuracy by doing more than this number of experiments. However, the behavior of the trend possibly can be explained by uneven distribution of induction times within the experimental series. Since we are interested in 20% and 30% accuracy levels, more precise

analysis has been done with the results of 39 and 36 experiments respectively as the least sufficient number (see Table 5) based on KaleidaGraph® best-fit analysis.

The trend for MLE estimation is not stabilized at any level, but is rather monotonically decreasing with lowering the number of parallels. At the same time, MLE method results in less discrepancy than the best-fit method. Having 20% acceptable deviation from the reference level, a number of 46 experiments was assumed as sufficient, whereas for a tolerance of 30% a minimum acceptable number of parallels equal 40 (see Table 5).

For this particular case KaleidaGraph® always gives higher estimates for nucleation rate than MLE does, but the difference is normally decreasing with reduction of the number of parallels (except for the case of 6 experiments).

Figure 18 where fitting curves for both KaleidaGraph® and MLE methods are presented can be used to compare the techniques. The red curve was generated by means of a curve-fitting algorithm integrated in KaleidaGraph® software by matching the experimental data with Equation (6) where J and $\tau_0$were used as fitting parameters. By inserting values of J and τ_0 estimated by MLE method into Equation (6), the similar curve for this method (colored blue) was generated.

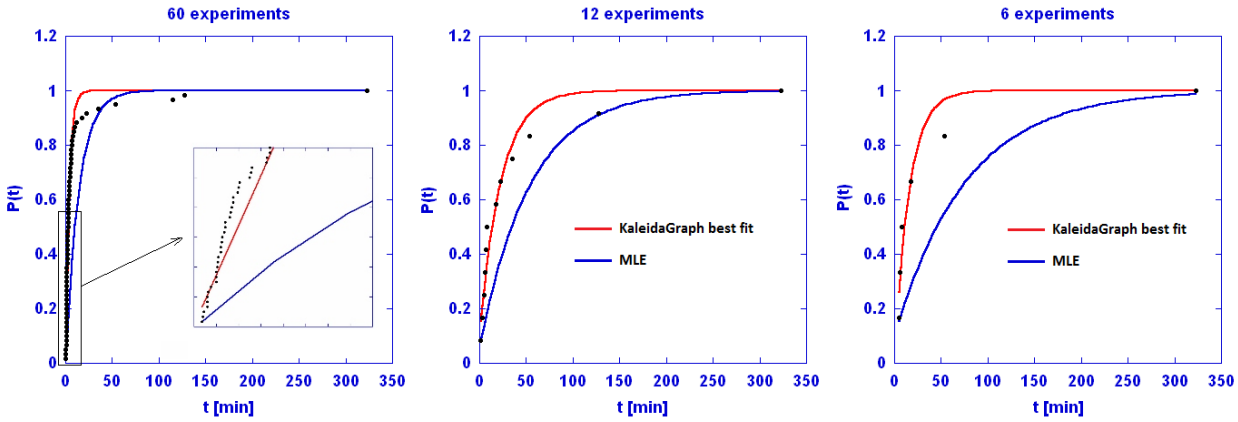

Figure 18 – Comparison of the fitting curves for KaleidaGraph® method and MLE for 60, 12 and 6 experiments at 13 °C

As can be seen from Figure 18, a curve for KaleidaGraph® matches well the whole distribution, excluding few datapoints in the upper part. In contrast, the fit curve for MLE does not describe a considerable part of the distribution. In particular, all the points it always goes through are the first and the last ones. This behavior can be explained by the fact that the lowest value of the distribution is a variable in the equations of the scale and location parameters

(Equations 18 and 19 respectively) estimated by MLE. The highest value in turn strongly affects the mean value, which is also included in these equations. Thereby we can suppose that MLE method is “linked” to extreme values of the population.

4.1.1.2 Series with 30 on SNG-2 at 11.75 °C and 90 bars

There were 30 parallels produced in contrast to 60 at 13.00 °C due to limited time of the present MSc work. Subcooling at 11.75 °C and 90 bar was calculated to be $\Delta T = 10.1$°C, which is the highest among the three series, therefore the short induction times were expected. The measured induction times are shown in chronological order in Figure 19.

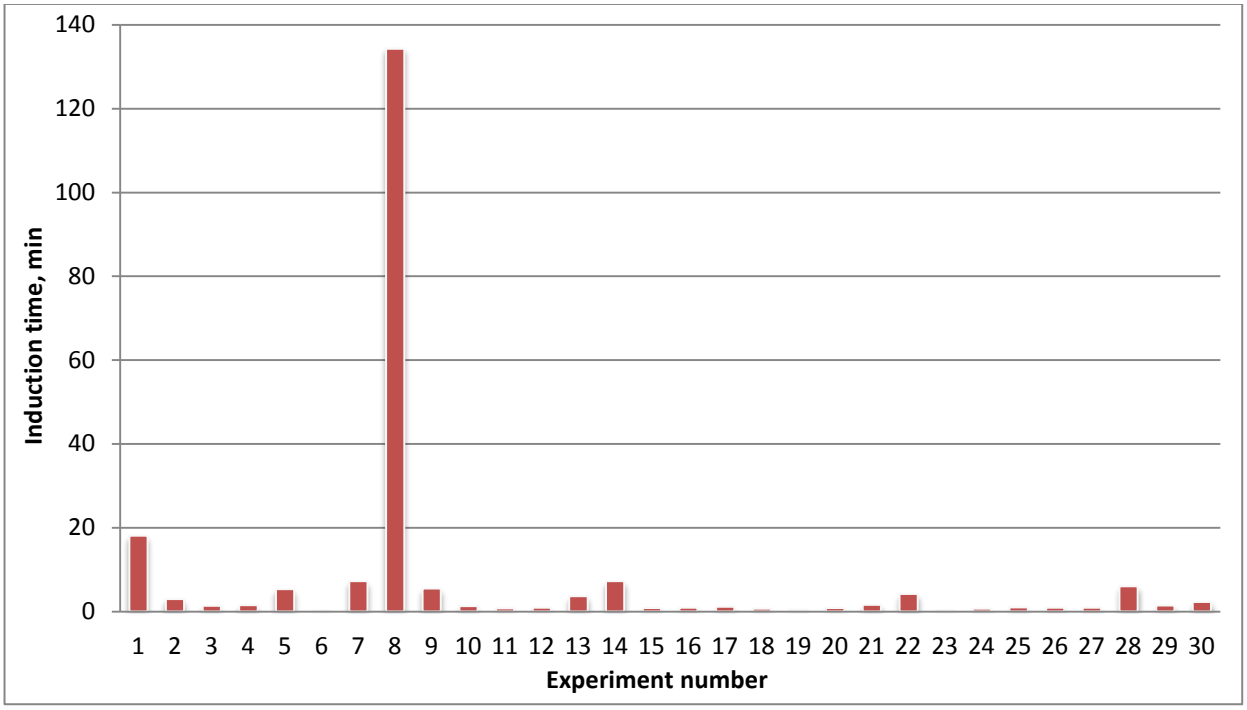

Figure 19 – The induction time in 30 measurements for SNG-2 in water at 11.75 °C

As can be seen from the figure, the overall picture is violated by one relatively high induction time in the experiment #8. But in general the data inconsistency is not as large as it is for the experimental series at 13 °C: the stochastic behavior of the hydrate nucleation is lowered by high driving force.

Experiment #8 gave very long induction time as compared to the others. Thus this point will result in reduced nucleation rate when included. This effect is seen for both KaleidaGraph®

and MLE analysis of Table 6 comparing calculations based on the first 6 parallels and the first 12 containing experiment #8. One long induction time among 29 shorter cannot be assumed as un-normal.

A permutation test was run in order to compare the groups by time periods they have been performed (numbered as I, II and III). The obtained p-values are 0.3733, 0.1247 and 0.2535 for the test between the groups I + II, II + III, I + III respectively, which shows that there is no significant difference between the distribution ranges of the groups and they can be treated as one distribution. This supports the assumption that parallel #8 belong to the same distribution range as the other parallels and can't be said to be "un-normal".

The results of calculations are presented in Table 6.

Table 6 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for experiments at 11.75 °C on SNG-2 in water

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev., %	Conf. interval	τ_0, min	J, min^{-1}	Dev., %	Conf. interval	τ_0, min
30	0.4853	0	[0.3250 , 0.6771]	0.086	0.1394	0	[0.0934 , 0.1946]	-0.039
24	0.4190	13.66	[0.2656 , 0.6068]	-0.061	0.1169	16.18	[0.0741 , 0.1693]	-0.157
23	0.3931	18.99	[0.2464 , 0.5736]	-0.042	0.1118	19.80	[0.0703 , 0.1636]	-0.189
22	0.3628	25.24	[0.2246 , 0.5336]	-0.074	0.1070	23.25	[0.0663 , 0.1574]	-0.225
20	0.3685	24.06	[0.2219 , 0.5517]	0.134	0.0996	28.60	[0.0599 , 0.1491]	-0.302
19	0.3343	31.10	[0.2184 , 0.5572]	-0.195	0.0946	32.12	[0.0561 , 0.1431]	-0.356
18	0.3053	37.09	[0.1778 , 0.4666]	-0.204	0.0895	35.83	[0.0521 , 0.1368]	-0.421
12	0.2457	49.37	[0.1224 , 0.4101]	-0.380	0.0622	55.38	[0.0311 , 0.1040]	-1.139
6	0.3214	33.76	[0.1044 , 0.6584]	-0.320	0.1773	27.17	[0.0576 , 0.3636]	-0.740

As can be seen from Table 6, the required numbers of parallels to obtain accuracies within 20 and 30% deviation are 23 and 20 respectively for both KaleidaGraph® best-fit and MLE methods.

Figure 20 shows nucleation rate J as function of the number of parallels included in the estimation as obtained through best-fit and MLE analysis.

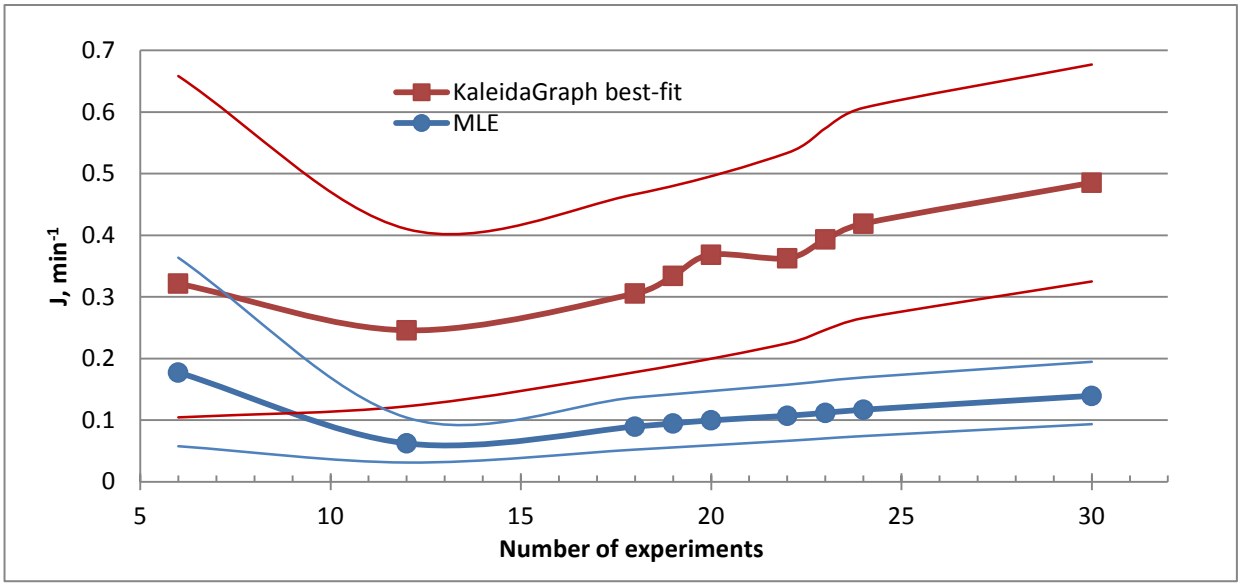

Figure 20 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 11.75 °C

As well as for the case with 60 parallels at 13.0 °C, the plots of nucleation rate as function of amount of experiments (see Figure 20) follow nearly the same trend for both KaleidaGraph® and MLE methods. MLE consequently produces results with lower nucleation rates, although discrepancy between estimates by two different methods is decreasing with reducing a number of parallels. Both plots follow a nearly linear trend while reducing a number of experiments down to 12. The estimates of nucleation rate by both best-fit and MLE methods for the case of 6 parallels become higher due to elimination of the group containing the value with the highest magnitude (experiment #8).

Apart from the case with 60 parallels at 13.0 °C, there is no stabilization in trend behavior for KaleidaGraph® best-fit method.

The confidence interval width is pretty stable for both methods down to the number of 12 experiments, whereas for the case of 6 parallels their boarders diverge. This behavior can be explained by both influences of mean value and the size of sample, which act in the opposite directions.

Comparison of fitting curves for KaleidaGraph® and MLE methods for cases of 30 and 6 parallels are presented in Figure 21.

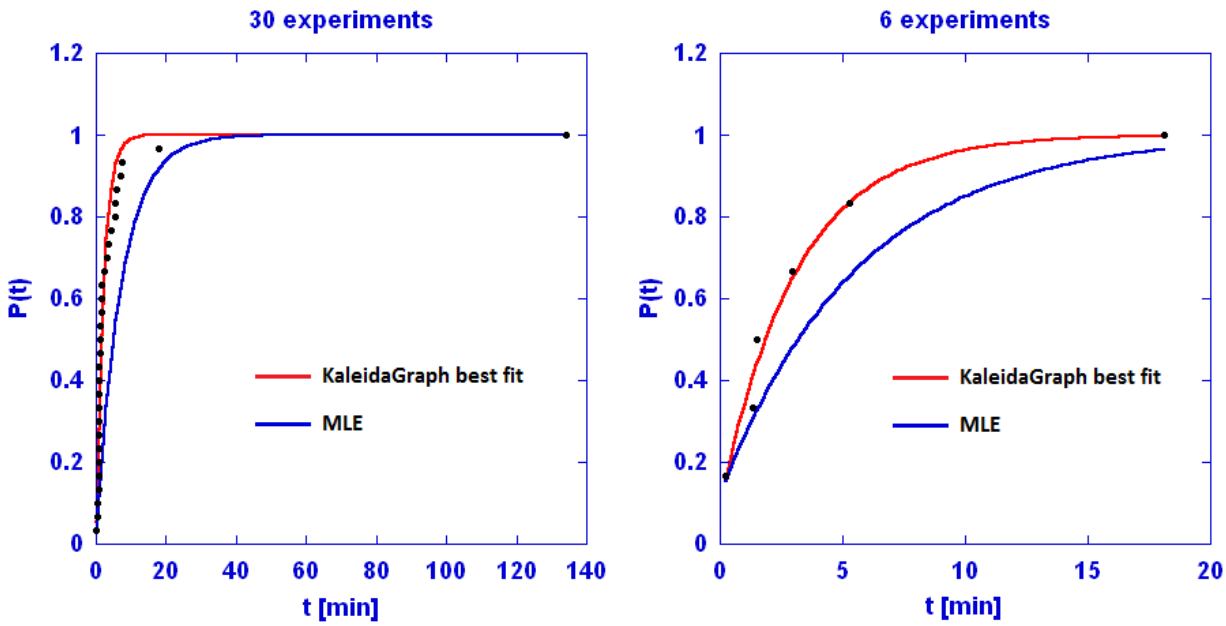

Figure 21 – Comparison of the fitting curves for KaleidaGraph® and MLE methods for 30 and 6 experiments at 11.75 °C

From Figure 21 one can see that the fitting curves show behavior similar to that for the experimental series at 13.0 °C: the curve for KaleidaGraph® fits almost all the experimental datapoints, whereas MLE curve matches just few.

4.1.1.3 Series with 30 experiments on SNG-2 at 13 °C and 90 bars

Finally the experimental series with the lowest driving force is analysed. The main results are presented in Figure 22, Figure **23** and Figure **24** and Table 7. The subcooling during the experiments at 14.25 °C and 90 bars is $\Delta T = 21.85 - 14.25 = 7.6$ °C.

The calculated induction times are shown in chronological order in Figure 22.

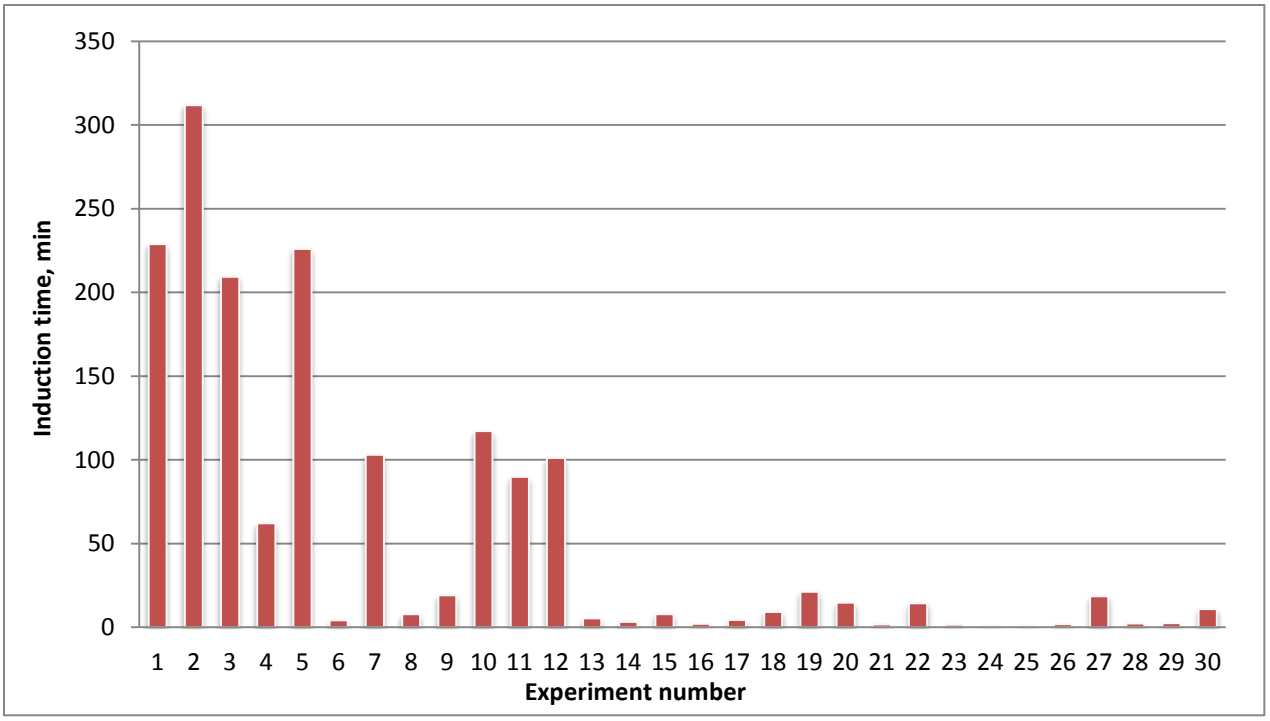

Figure 22 – The induction time in 30 measurements for SNG-2 in water at 14.25°C

If compare the experiments by groups according to the sequence they have been conducted, one can see the distinct differences between the distributions of induction times within the different groups. The general trend is decreasing induction times from the first group to the last one. In particular, all the longest times are in the first group, the average ones are in the second, and the third group includes just low induction times.

With reference to Figure 15, one can see that the first group of 6 experiments was run in January, the second one, also containing 6 experiments was produced in March after the experiments 6-24 at 13 °C were finished, and the series was completed in April to obtain total number of 30 parallels. Thus the assumption of the equipment influence on the results is still valid. On the other hand, the low driving force is reflected in more stochastic behavior of the process of nucleation. Therefore the induction time variation is assumed as a result of influence by both these factors.

The permutation test for these three groups gave p-values of 0.0281, 0 and 0 for comparison of the groups I + II, II + III and I + III respectively, which indicates that there is no interrelation between the groups. Nonetheless, the analysis for the whole dataset has been done.

Table 7 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for experiments at 14.25°C on SNG-2 in water

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev., %	Conf. interval	τ_0, min	J, min^{-1}	Dev., %	Conf. interval	τ_0, min
30	0.0545	0	[0.0365 , 0.0761]	-1.920	0.0184	0	[0.0123 , 0.0256]	-0.914
28	0.0459	15.85	[0.0302 , 0.0647]	-2.881	0.0172	6.199	[0.0114 , 0.0243]	-1.172
27	0.0396	27.27	[0.0259 , 0.0563]	-3.644	0.0166	9.589	[0.0109 , 0.0236]	-1.330
26	0.0254	53.36	[0.0165 , 0.0363]	0.9263	0.0162	12.08	[0.0105 , 0.0231]	-1.481
24	0.0164	69.95	[0.0104 , 0.0237]	-15.64	0.0149	19.06	[0.0094 , 0.0215]	-1.902
23	0.0147	72.94	[0.0092 , 0.0215]	-17.04	0.0144	21.82	[0.0090 , 0.0210]	-1.477
21	0.0123	77.46	[0.0075 , 0.0182]	-20.33	0.0132	28.08	[0.0081 , 0.0196]	-1.803
20	0.0115	78.87	[0.0069 , 0.0172]	-20.55	0.0126	31.40	[0.0076 , 0.0189]	-1.867
18	0.0097	82.24	[0.0056 , 0.0148]	-25.37	0.0115	37.30	[0.0067 , 0.0176]	-2.722
12	0.0073	86.61	[0.0036 , 0.0122]	-10.22	0.0077	58.15	[0.0038 , 0.0129]	-6.638
6	0.0051	90.61	[0.0017 , 0.0105]	-23.39	0.0049	73.31	[0.0016 , 0.0101]	-29.79

As can be seen from Figure 7, 28 parallels for best-fit and 24 parallels for MLE are required to obtain accuracies within 20 %. The minimum acceptable number is slightly lower for the accuracy limit of 30%: 27 and 21 parallels for best-fit and MLE methods respectively.

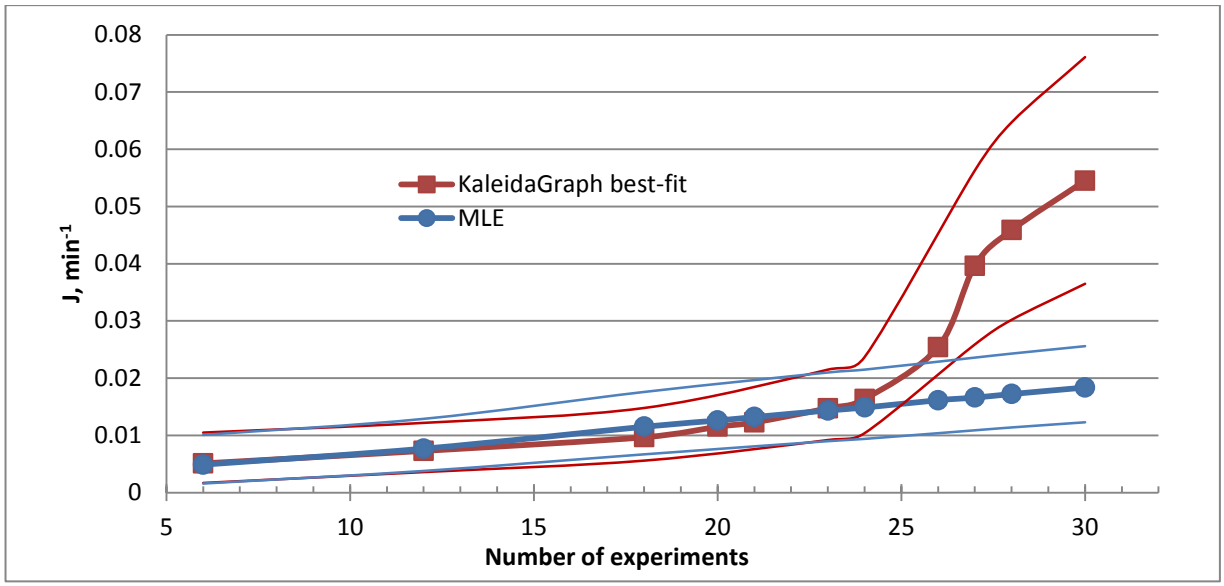

Figure 23 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 14.25°C

Figure 23 shows that the values of estimates by best-fit method are decreasing rapidly when lowering the number of experiments, but the behavior is stabilized by the number of parallels equal 24. The estimates by MLE method decrease constantly following a nearly linear trend. Thus MLE results in less discrepancies in comparison to best-fit method.

As for the previous experimental series, the difference between the estimates of nucleation rate by two methods is decreasing with lowering the number of parallels. It is interesting to notice that after reducing the number of experiments below 24, both best-fit and MLE methods produce very similar results. By eliminating the datapoints with low values the distribution of the experiments is becoming more and more uniform, which has a positive influence on the MLE estimation. This tendency is well illustrated by Figure 24, where the divergence between the curves is gradually decreasing with reducing the number of parallels and for the case of 6 experiments difference between the curves is barely noticeable.

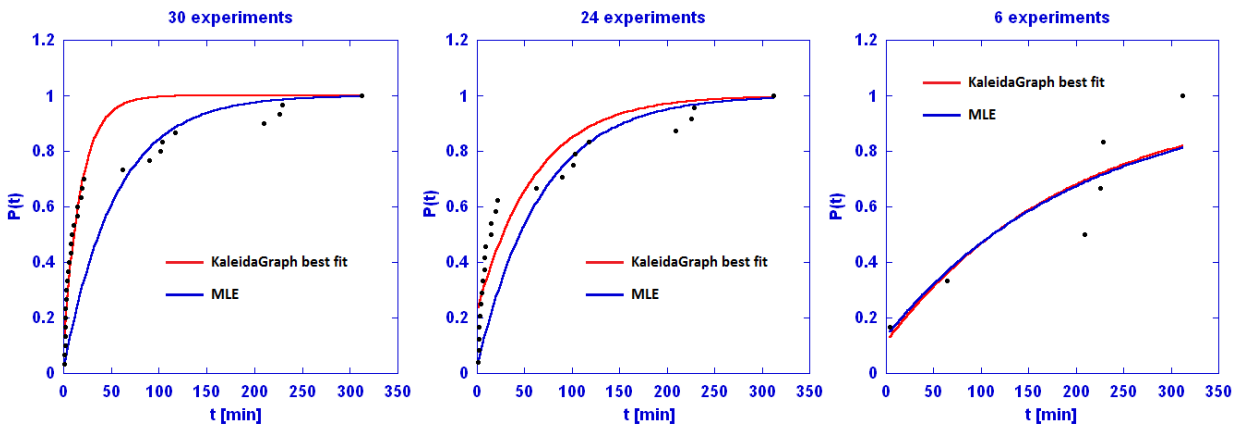

Figure 24 – Comparison of the fitting curves for KaleidaGraph® and MLE methods for 30, 24 and 6 experiments at 14.25 °C

4.1.2 SNG-7 large cell data analysis

Since a time span of the present thesis work was confined by just few months, there was no opportunity to perform the amount of experiments on the large cell similar to that for cell #3. Therefore, having a relatively small number of parallels in each series, the main purpose of the large cell experiments was to examine the required number of parallels to estimate activation energy and a radius of a critical nucleus within an acceptable accuracy. Analysis of nucleation rate and a lag time was also included for every series.

4.1.2.1 Series with 12 experiments on SNG-7 at 15.5 °C and 90 bars

It is the series with the highest driving force among the three for the large cell. The results are presented through Figures Figure 25, Figure 26 and Table 8. The subcooling at 15.5 C and 90 bar was calculated to be:

$$\Delta T = 20.97 - 15.50 = 5.47\ ^{\circ}\mathrm{C}.$$

All the measured induction times are given in chronological order in Figure 25. At first sight it can be declared that the dispersion of values is relatively low compare to that for cell #3 experimental series.

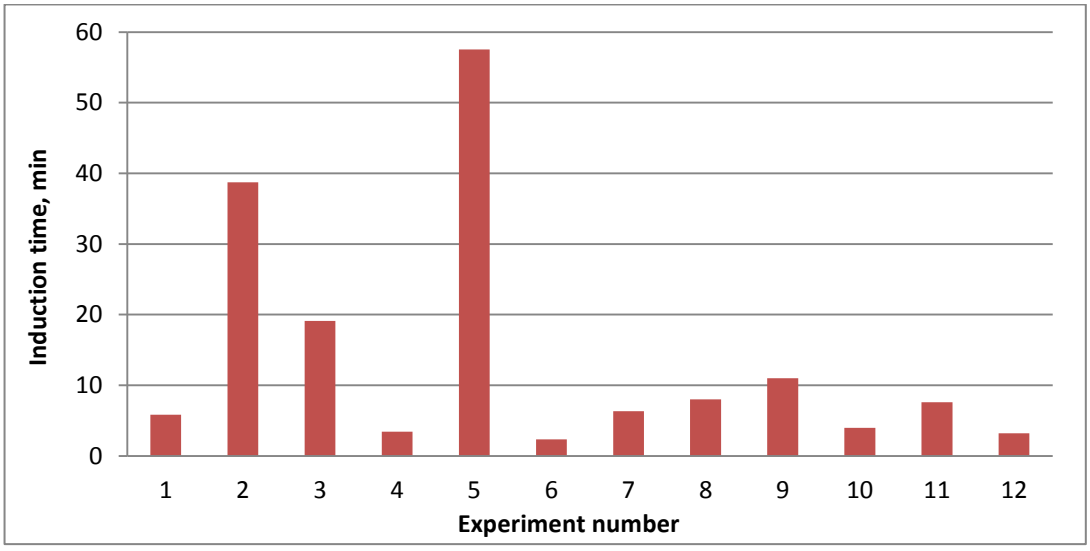

Figure 25 – The induction time in 12 measurements for SNG-7 in water at 15.50 °C

Table 8 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for experiments at 15.50 °C for SNG-7 in water

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev, %	Conf. interval	τ_0, min	J, min^{-1}	Dev, %	Conf. interval	τ_0, min
12	0.1453	0	[0.0725 , 0.2429]	1.616	0.0792	0	[0.0396 , 0.1325]	1.298
10	0.1200	17.39	[0.0549 , 0.2103]	1.261	0.0678	14.42	[0.0310 , 0.1187]	0.875
9	0.0986	32.12	[0.0426 , 0.1778]	0.917	0.0610	23.00	[0.0263 , 0.1100]	0.528
8	0.0978	32.73	[0.0393 , 0.1839]	0.405	0.0571	27.87	[0.0230 , 0.1066]	0.163
6	0.0517	64.39	[0.0168 , 0.1060]	-3.670	0.0443	44.09	[0.0144 , 0.0907]	-1.410

As can be seen from Table 8, accuracies within 20% are obtained by 10 experiments for both methods of estimation. The required number of parallels to be inside 30% accuracy level is 8 for MLE and still 10 for best-fit technique.

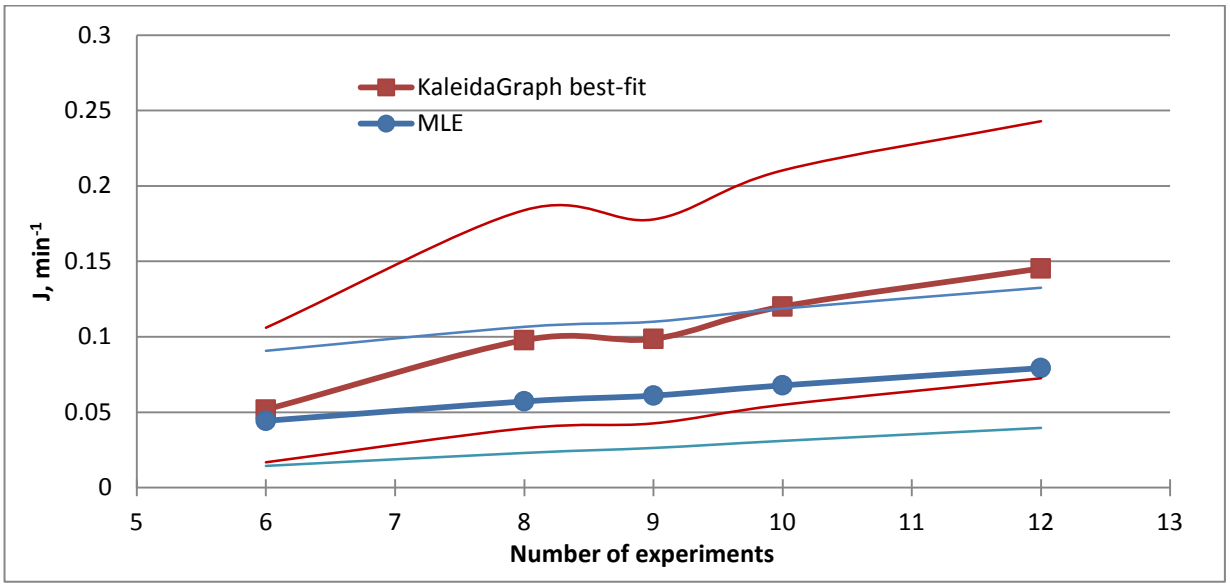

Figure 26 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 15.50°C

Figure 26 shows that all the main tendencies from the cell #3 data analysis are also valid for the present series, namely:

1) Nucleation rate is decreasing as function of a decreasing number of parallels;

2) KaleidaGraph® best-fit estimates are higher than that for the MLE method;

3) The difference between estimates is decreasing with reducing a number of parallels;

4) The trend is nearly linear for MLE method and results in less discrepancy that that for best-fit technique.

4.1.2.2 Series with 20 experiments on SNG-7 at 16.75 °C and 90 bars

The subcooling during the experiments at 16.75 °C and 90 bars is $\Delta T = 20.97 - 16.75 = 4.22$ °C.

First is should be said that the technical problem arouse while carrying experiments on the large cell at 16.75 °C. Relatively small to quite large leaks were detected through the upper O-ring sealing. The problem was attempted to be fixed by replacement the standard O-ring dimension (94.93 x 2.62 mm) by another dimension (86 x 3 mm) hoping that the stretch of a smaller ring with larger thickness would improve pressure sealing and eliminate the cause of leak. The replacement helps to reduce the leak, but not to eliminate it.

The average pressure drop was 1.0 – 1.5 bar during an experiment. Analyzing the plot P vs t in KaleidaGraph® it was clear that the pressure decreased during the induction period too though the pressure drop was not so great during this period. This problem possibly could lead to errors and inaccuracies in the parameters estimation.

In order to estimate whether the pressure drop affected the nucleation process permutation test was performed. The analysis gave the p-value of 0.1179 when comparing two populations of data points of ordinal experiments and of experiments with a leak. Since this value is higher than 0.05, it was concluded that there is no significant difference between these two datasets, with leaks and without leaks, therefore it was assumed they could be treated as one distribution.

In total there were conducted 20 experiments. In the experiment # 2 hydrates did not form during the long time and the experiment was terminated after 1055 minutes. The probability of getting hydrates later that time is still exist, therefore the value of observation time was included in the dataset which treatment was performed in accordance with a scheme for incomplete dataset, that has been described in the paragraph 2.12.

Chronology of the measured induction times (for the experiment #2 – observation time) in the experiments is given in Figure 27.

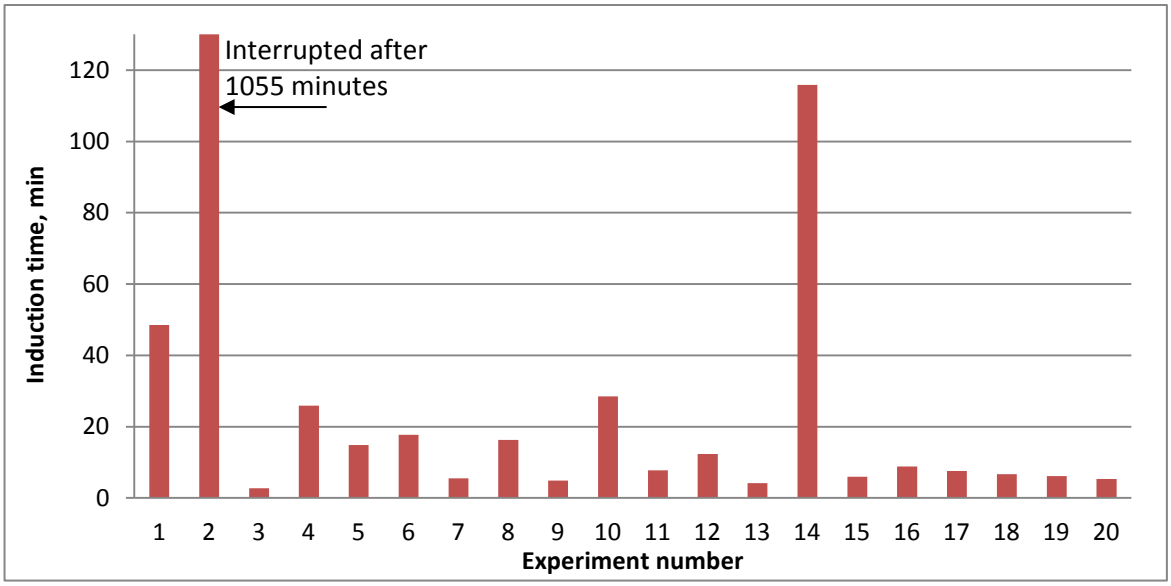

Figure 27 – The induction time in 20 measurements for SNG-7 in water at 16.75 °C

As can be seen from the figure, despite the longest observation time in the experiment #2, the overall distribution of induction times is more or less consistent, but one outlier (experiment #14) is still present.

The results of calculations are presented in Table 9.

Table 9 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for experiments at 16.75°C on SNG-7 in water

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev., %	Conf. interval	τ_0, min	J, min^{-1}	Dev., %	Conf. interval	τ_0, min
20	0.0902	0	[0.0543 , 0.1351]	2.048	0.0075	0	[0.0045 , 0.0112]	-3.970
18	0.0766	15.05	[0.0446 , 0.1171]	1.773	0.0067	10.89	[0.0039 , 0.0102]	-5.617
17	0.0700	22.36	[0.0400 , 0.1082]	1.526	0.0063	16.32	[0.0036 , 0.0097]	-6.678
16	0.0636	29.44	[0.0356 , 0.0996]	1.160	0.0059	21.74	[0.0033 , 0.0092]	-7.954
15	0.05728	36.48	[0.0313 , 0.0910]	0.603	0.0055	27.14	[0.0030 , 0.0087]	-9.507
14	0.0520	42.34	[0.0277 , 0.0838]	0.425	0.005	32.65	[0.0027 , 0.0081]	-11.45
12	0.0578	35.94	[0.0289 , 0.0966]	1.243	0.0044	41.03	[0.0022 , 0.0074]	-16.15
10	0.0467	48.26	[0.0214 , 0.0818]	-0.055	0.0036	52.52	[0.0016 , 0.0062]	-25.40
8	0.0392	56.52	[0.0158 , 0.0731]	-0.615	0.0027	63.94	[0.0011 , 0.0050]	-43.55
6	0.0345	61.70	[0.0112 , 0.0707]	-1.551	0.0036	51.33	[0.0012 , 0.0075]	-24.71

From Table 9 one can see that the minimum acceptable number of parallels for 20% tolerance is on the level of 18 for KaleidaGraph® estimation and 17 for MLE method. 30% accuracy level results in 16 and 15 experiment for the best-fit and MLE techniques respectively.

Estimation of induction time values by MLE method is strongly affected by the datapoint with the highest magnitude, since this point is in the beginning of the distribution. Thereby the MLE estimates are several times lower than these for the best-fit method.

Figure 28 shows nucleation rate J as function of the number of parallels included in the estimation as obtained through KaleidaGraph and MLE analysis.

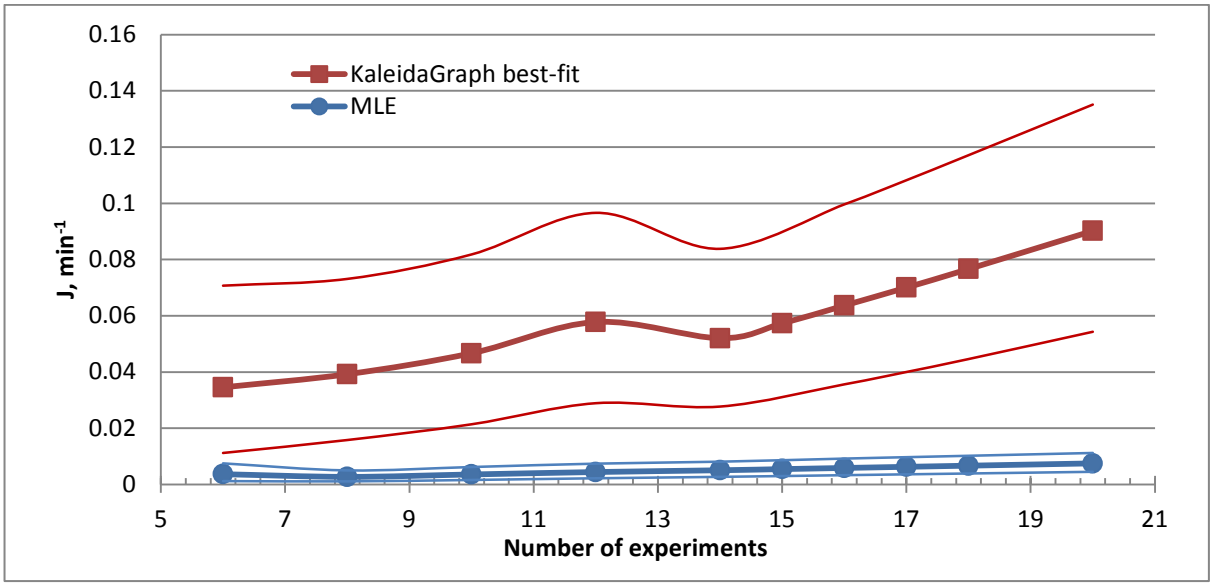

Figure 28 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 16.75°C

Figure 28 shows that in general all the tendencies in the trends behavior from the previous experimental series analysis are applicable for the present case.

From the analysis of experimental series at 16.75 ºC we can see that including the datapoint with the longest observation time to induction times dataset strongly affects MLE estimates of the nucleation rate since all the other values are relatively low (several times lower than the longest observation time). It was decided to do the analysis of the series ignoring this datapoint in order to identify the influence of the extreme values on both KaleidaGraph® and MLE methods. Thus we have a complete dataset containing 19 datapoints.

The results of the analysis of the main parameters are given in Table 10. As can be noticed, the values of nucleation rate estimates by KaleidaGraph® increase just slightly, whereas

the values that MLE method estimates raised several times, the difference between estimated values became comparable to that for other experimental series.

Table 10 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for experiments at 16.75°C on SNG-7 in water; an experiment with non-occurrence of hydrates is ignored

# of exp.	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev., %	Conf. interval	τ_0, min	J, min^{-1}	Dev., %	Conf. interval	τ_0, min
19	0.1047	0	[0.0620 , 0.1583]	2.336	0.0612	0	[0.0363 , 0.0926]	1.840
18	0.0968	7.53	[0.0564 , 0.1479]	2.243	0.0583	4.70	[0.0340 , 0.0892]	1.748
17	0.0893	14.69	[0.0518 , 0.1381]	2.087	0.0556	9.24	[0.0318 , 0.0859]	1.641
16	0.0819	21.72	[0.0459 , 0.1283]	1.862	0.0528	13.73	[0.0296 , 0.0827]	1.516
15	0.0749	28.39	[0.0410 , 0.1190]	1.534	0.0502	18.07	[0.0274 , 0.0796]	1.371
14	0.0681	34.9	[0.0363 , 0.1099]	1.041	0.0476	22.22	[0.0254 , 0.0768]	1.200
12	0.0749	28.44	[0.0374 , 0.1252]	1.251	0.0702	14.71	[0.0350 , 0.1174]	1.513
10	0.0659	37.02	[0.0301 , 0.1155]	0.980	0.0618	1.01	[0.0283 , 0.1083]	1.083
8	0.0672	35.83	[0.0270 , 0.1253]	0.310	0.061	0.346	[0.0245 , 0.1138]	0.651
6	0.0592	43.47	[0.0192 , 0.1212]	-0.443	0.0505	17.49	[0.0164 , 0.1035]	-0.600

As a result we have 17 parallels for KaleidaGraph® best-fit method and 15 – for MLE as a minimum acceptable number for accuracy limit of 20% and 15 and 6 respectively for accuracy of 30%.

Figure 29 presents nucleation rate J as function of the number of parallels included in the estimation as obtained through KaleidaGraph and MLE analysis.

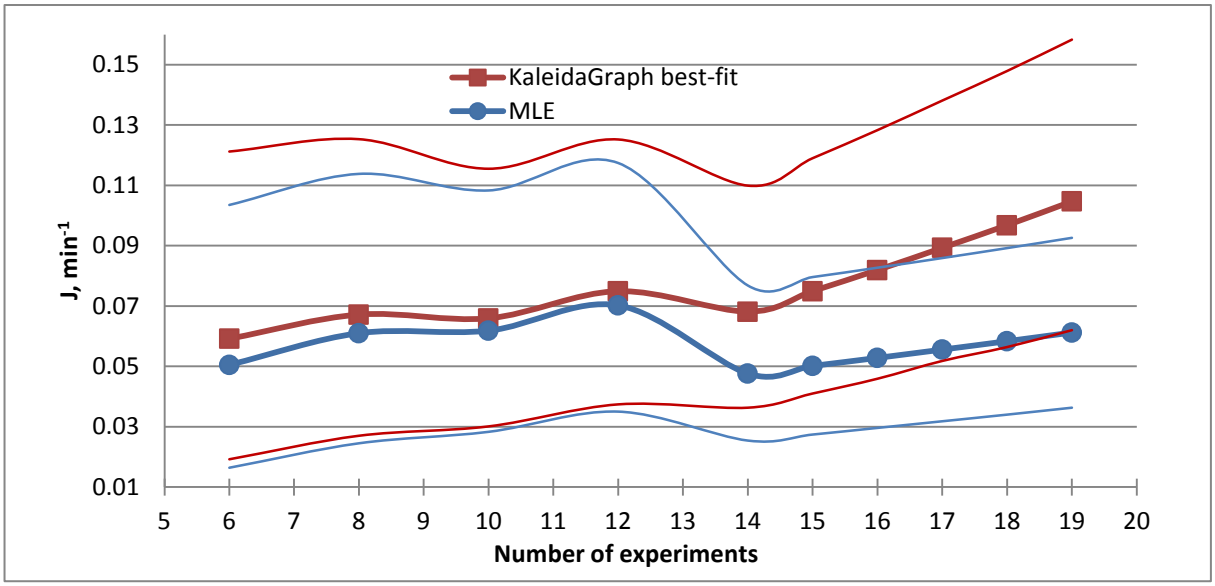

Figure 29 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 16.75°C

The trend for best-fit method did not change considerably compare to the trend from the analysis with including the long observation time. At the same time, the trend for MLE estimates changed drastically and does not follow the straight line. Furthermore, it produces larger discrepancies when reducing the number of parallels than KaleidaGraph® method.

Figure 30 where the fitting curves for both KaleidaGraph® and MLE methods are presented for the cases of incomplete (20 experiments) and complete (19 experiments) datasets can be used to illustrate effect of excluding the observation time in experiment with non-occurrence of hydrates on fitting curves.

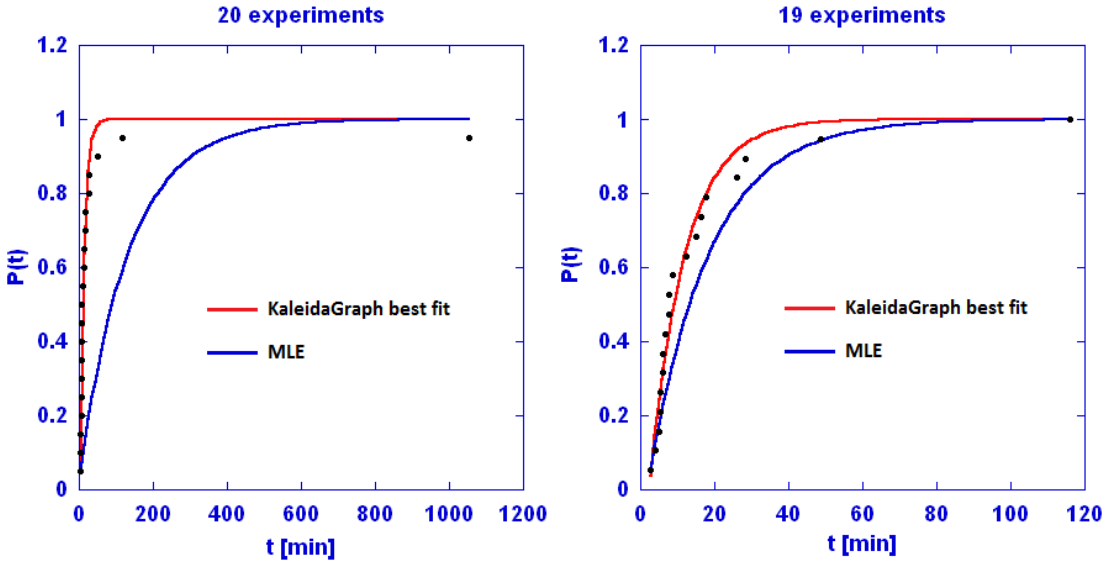

Figure 30 – Comparison of the fitting curves for KaleidaGraph® and MLE methods for complete (20experiments) and incomplete (19 experiments) datasets for experiments at 16.75 °C

Estimation of nucleation rate by MLE is assumed to produce more reliable results in case when we do not include the value of the observation time in the "unsuccessful" experiment #2. This is well illustrated by Figure 30, where the distance between the fitting curves produced by different methods is lower for the case with 19 experiments (when excluding the "unsuccessful" experiment). On the other hand discrepancies for KaleidaGraph® estimation of nucleation rate for complete (19 parallels) and incomplete dataset (20 parallels) are in order of several percent. Therefore for further calculation of activation energy and critical cluster radius we will use estimates of nucleation rate by best-fit method for incomplete dataset (20 experiments) and estimates of nucleation rate by MLE for complete dataset(19 experiments).

4.1.2.3 Series with 12 experiments at 17.75°C

The main results are presented through Figures Figure **31**, Figure **32** and Table 11. The driving force represented by subcooling at 17.75 °C and 90 bars is $\Delta T = 20.97 - 17.75 = 3.22$ °C.

As for the experimental series at 16.75 °C, the same situation occurred: in the experiment #4 hydrates did not form during the observation period that was interrupted after 1377 minutes. This observation time was included in the dataset since it does not affect the overall distribution of relatively low induction times. The chronology of the calculated induction times (for the experiment 4 – observation time) for the experiments are given on Figure 31.

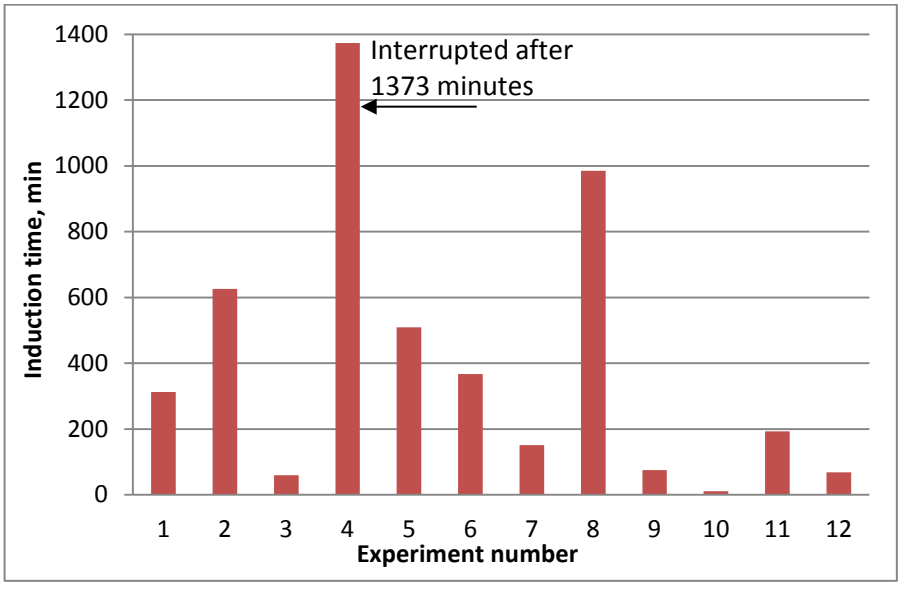

Figure 31 – The induction time in 12 measurements for SNG-7 in water at 17.75 °C

A span of the induction times is wide, but the main part of the distribution is in the middle part, i.e. the distribution is close to normal.

Table 11 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for 12 experiments at 17.75°C on SNG-7 in water

# of exp	KaleidaGraph® best-fit				MLE			
	J, min^{-1}	Dev, %	Conf. interval	τ_0, min	J, min^{-1}	Dev, %	Conf. interval	τ_0, min
12	0.00268	0	[0.00134 , 0.00448]	-36.90	0.00167	0	[0.00083 , 0.00279]	-39.37
10	0.00215	19.82	[0.00098 , 0.00376]	-57.35	0.00139	16.64	[0.00064 , 0.00243]	-61.27
9	0.00206	23.21	[0.00089 , 0.00371]	-23.31	0.00132	20.96	[0.00057 , 0.00238]	-24.39
8	0.00190	29.11	[0.00076 , 0.00354]	1.35	0.00114	32.06	[0.00046 , 0.00213]	-50.34
6	0.00189	29.59	[0.00061 , 0.00386]	-15.84	0.00094	43.9	[0.00031 , 0.00193]	-117.9

As can be seen from Table 11, the required number of parallels to obtain accuracies within 20% deviation is 10 for both KaleidaGraph® and MLE techniques. 30% deviation set the minimum number to be 6 for best-fit and 9 for MLE.

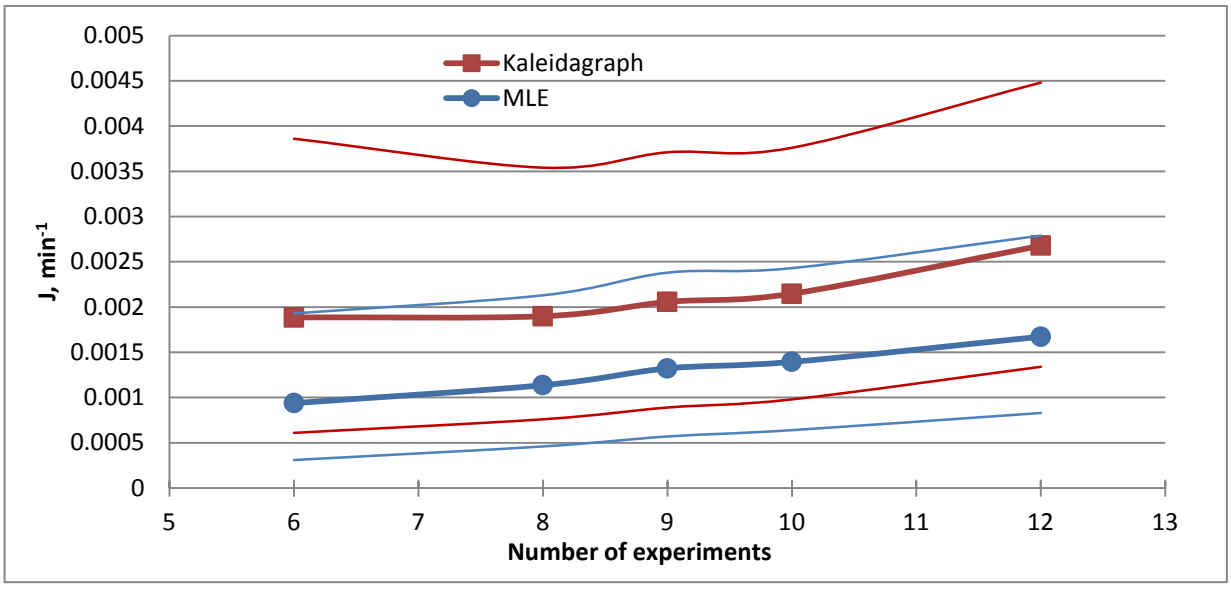

Figure 32 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of parallels for experiments at 17.75°C

From Figure 32 it may be noticed that the trends follow very similar path. However, the estimation of nucleation rate by MLE is again lower than by KaleidaGraph® best-fit. At the same time, apart from the previous cases, there is no tendency for convergence of the trends.

4.1.3 Analysis of data from other studies

The technique analyzing nucleation rates of the present thesis was used on relatively large datasets found in literature for comparison. Datasets from studies of Jiang and ter Horst (2011) and Kulkarni et al. (2013) on crystallization processes from supersaturated solutions, containing 80 and 144 datapoints respectively, were used for this analysis. Induction times presented in the papers by Jiang and ter Horst and Kulkarni et al. were read from graph and could thus contain minor inaccuracies. The inaccuracy of reading was assumed negligible for our purpose.

Although these works are not related to gas hydrates, the authors study a stochastic nucleation process following the same type exponential probability distribution. In both studies the multiple reactor setup Crystal-16 was utilized, which means that 16 experiments were run simultaneously. The first important feature of the experiments in these studies is that crystallization took place in in equal glass vials with internal magnetic stirring, which supposedly should decrease the impact of equipment on the results of experiments since the inner surface is extremely smooth compare, for instance, to our cells containing several metal (titanium) parts. The second feature is that they obtained the homogenous solubility of the solute in supersaturated solution and the small volume vials were cooled down to the experimental temperature relatively rapid to avoid incipient nucleation during the cooling stage. It should be noted that one cannot obtain similar conditions of pre-controlled saturation and rapid arrival at the experimental condition using our type of hydrate cell equipment.

4.1.3.1 Jiang and ter Horst study with 80 parallels

Jiang and ter Horst studied crystal nucleation rates in solutions of m-aminobenzoic acid and L-Histidine over the range of supersaturation ratios 1.83 – 2.15 and 1.55 – 1.79 for the former and later substance respectively. They perform experiments in 1.8 ml glass vials at a constant temperature of 25 ºC. A bottle-top dispenser was used to dispense 1 ml of clear solution into each vial. A cooling rate was chosen to be 5.0 Cº/min. A stirring speed was maintained at 900 rpm. After detection of crystallization the samples were reheated with a rate of 1 Cº/min and maintained above the equilibrium temperature to dissolve the crystals and to get a clear solution. Then the subsequent measurements of induction time started. By repetition of this cycle for 5 times they obtained a total of 80 induction time measurements for each supersaturation ratio [4].

The induction times in 80 measurements for m-aminobenzoic acid in water/ethanol mixture (50 wt%) at the supersaturation ratio S = 1.96 are given in Figure B1 in Appendix B. Analysing the figure it can be concluded that the behavior of crystallization process is less stochastic compare to the experiments on high pressure cells in the UiS lab.

The results of calculations of nucleation rate and lag time are listed in Table 12.

Table 12 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for data from Jiang and ter Horst study

# of exp	KaleidaGraph® best-fit				MLE			
	$J, m^{-3}s^{-1}$	Dev., %	Conf. interval	τ_0, s	$J, m^{-3}s^{-1}$	Dev., %	Conf. interval	τ_0, s
80	626.79	0	[496.24 , 772.36]	1164.5	587.55	0	[465.17 , 724.00]	1005.7
64	630.80	0.64	[484.72 , 795.82]	1195.2	581.38	1.05	[446.75 , 733.47]	1000.1
48	680.33	8.54	[449.99 , 888.16]	1274.8	599.96	2.112	[440.83 , 783.23]	1074.3
32	760.32	21.30	[516.60 , 1050.4]	1288.3	685.92	16.74	[466.05 , 947.60]	1144.4
16	807.74	28.87	[452.09 , 1264.9]	1228.9	622.56	5.959	[348.44 , 974.91]	1089.6

Analyzing Table 12 we can say that for 20% and 30% tolerance MLE method set 1 cool-hold-heat cycle for a set of 16 vials, i.e. 16 experiments, as sufficient to obtain acceptable results. The same result (16 experiments) is obtained by estimation in KaleidaGraph® concerning 30% accuracy level, but for the level of 20% acceptable accuracy would be obtained only by 3 those cycles (48 experiments).

Figure 33 shows nucleation rate J as function of the number of parallels included in the estimation as obtained through KaleidaGraph and MLE analysis.

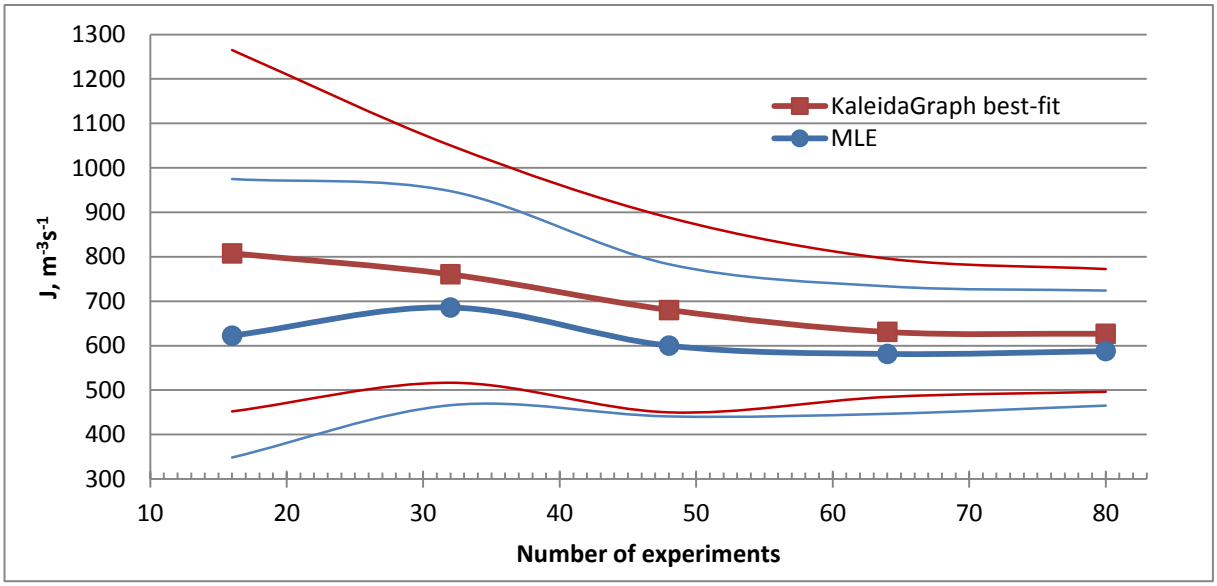

Figure 33 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of experiments for data from Jiang and ter Horst study

As can be seen from Figure 33, the trendlines for KaleidaGraph® best-fit and MLE methods behave similar while reducing numbers of parallels down to 32, but for 16 experiments they diverge. Unlike for the experiments in the UiS lab, the trend for MLE is neither linear, nor monotonic, but it can be noticed some stabilization in the behavior if the number of parallels is more than 48. The difference between the estimates by different methods keeps approximately constant when reducing the number of experiments down to 32; for 16 experiments the trends diverge. The same tendency is traced by analyzing the fitting curves generated by two different methods (see Figure 34). Also from this figure one can see that MLE for this particular case better describes the distribution than it was for the experiments on gas hydrates in high-pressure cells.

The confidence intervals show the expectable behavior by divergence while the number of datapoints in the sample is decreasing.

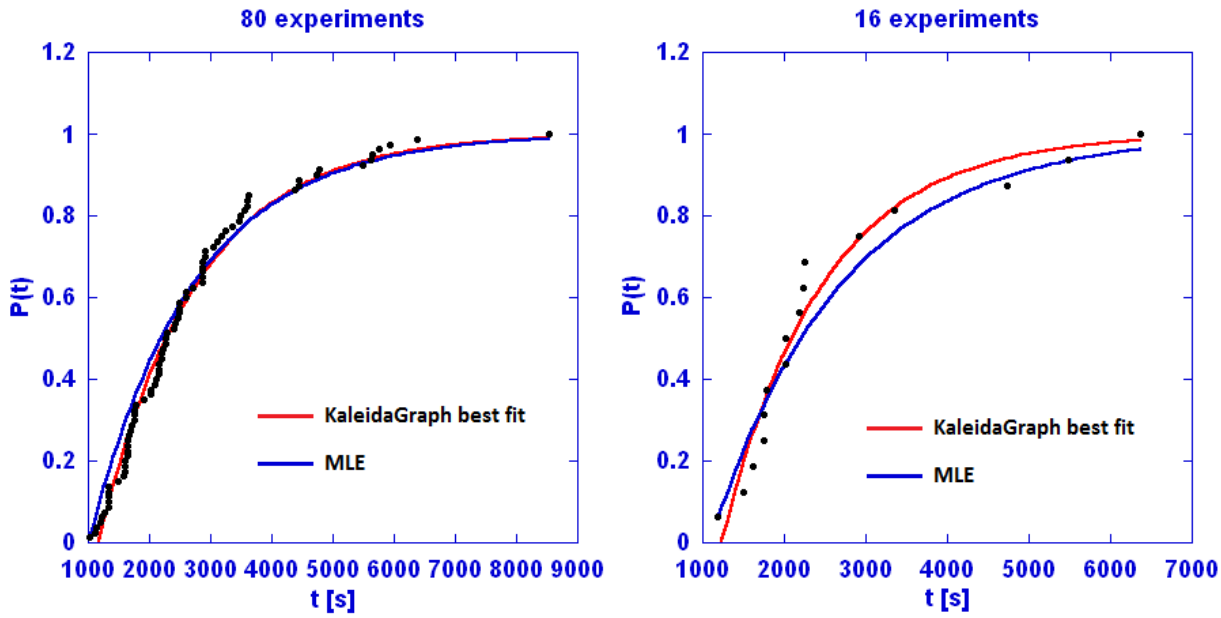

Figure 34 – Comparison of the fitting curves for KaleidaGraph® and MLE methods for 80 and 16 experiments for data from Jiang and ter Horst study

It should be mentioned that Jiang and ter Horst declare the nucleation rate to be 630±20 $m^{-3}s^{-1}$ for the supersaturation ratio of 1.96. Comparing this value with those for the nucleation rate obtained by best-fit and MLE methods for the case of 80 experiments (which equal 626.79 $m^{-3}s^{-1}$ and 587.55 $m^{-3}s^{-1}$ respectively), we can state that the technique they used for estimation in their study is close to KaleidaGraph® best-fit.

4.2.3.2 Kulkarni et al. study with 144 parallels

Kulkarni et al. studied crystal nucleation rate conducting the experiments on the solution of isonicotinamide in ethanol. As a reacting chamber they used 1.5 ml glass vials that have been placed in the Crystal16 reactor. Thereby 16 experiments were running simultaneously. They set a temperature of experiments to be 25 °C, used a cooling rate of 5 °C /min and a constant stirring speed of 700 rpm. After crystallization occurred, the samples were reheated with a rate 0.3 °C /min and maintained above the saturation temperature in order to dissolve the crystals. This cool–hold–heat cycle was repeated 9 times to get 144 measurements for each supersaturation ratio [5].

The induction times in 144 measurements for the case with a supersaturation ratio S = 1.40 are given in Figure B2 in Appendix B. As seen from this figure, there are very long

induction times, as well as very small, but in general the distribution of values is more or less consistent.

The results are presented through Table 13 and Figure 35,Figure **36**.

Table 13 – Results of calculations of nucleation rate $\boldsymbol{J}$ and lag time $\boldsymbol{\tau_0}$ for the data from Kulkarni et al. study

# of exp	KaleidaGraph® best-fit				MLE			
	$J, m^{-3}s^{-1}$	Dev, %	Conf. interval	τ_0, s	$J, m^{-3}s^{-1}$	Dev, %	Conf. interval	τ_0, s
144	461.14	0	[338.65 , 539.73]	211.32	401.41	0	[338.32 , 469.82]	182.70
128	529.79	14.89	[441.66 , 625.82]	239.69	432.98	7.86	[360.96 , 511.46]	181.96
112	494.68	7.27	[406.94 , 590.85]	236.89	406.57	1.28	[334.46 , 485.61]	178.04
96	503.75	9.24	[407.56 , 609.97]	262.86	407.85	1.60	[329.98 , 493.85]	174.46
80	567.36	23.03	[449.18 , 699.13]	251.79	430.21	7.17	[340.60 , 530.13]	170.94
64	477.75	3.60	[367.13 , 602.73]	236.62	371.85	7.37	[285.74 , 469.13]	157.98
48	485.61	5.30	[356.81 , 633.95]	233.00	374.13	6.78	[274.90 , 488.42]	144.32
32	530.2	14.98	[360.25 , 732.48]	166.97	450.29	12.20	[305.94 , 622.07]	130.60
16	413.24	10.39	[231.29 , 647.12]	56.95	326.49	18.66	[182.73 , 511.28]	8.57

As we can see from Table 13, if we consider MLE as a method to estimate nucleation rate, in this case it is enough to do just 1 cool-hold-heat cycle, i.e. 16 experiments to get acceptable result for both 20% and 30% accuracy levels. In contrast, for estimation in KaleidaGraph® the acceptable accuracy would be obtained only by 6 cycle runs (96 experiments) for the 20% accuracy limit, however only one cycle is enough to produce acceptable accuracy if we are looking for 30% level.

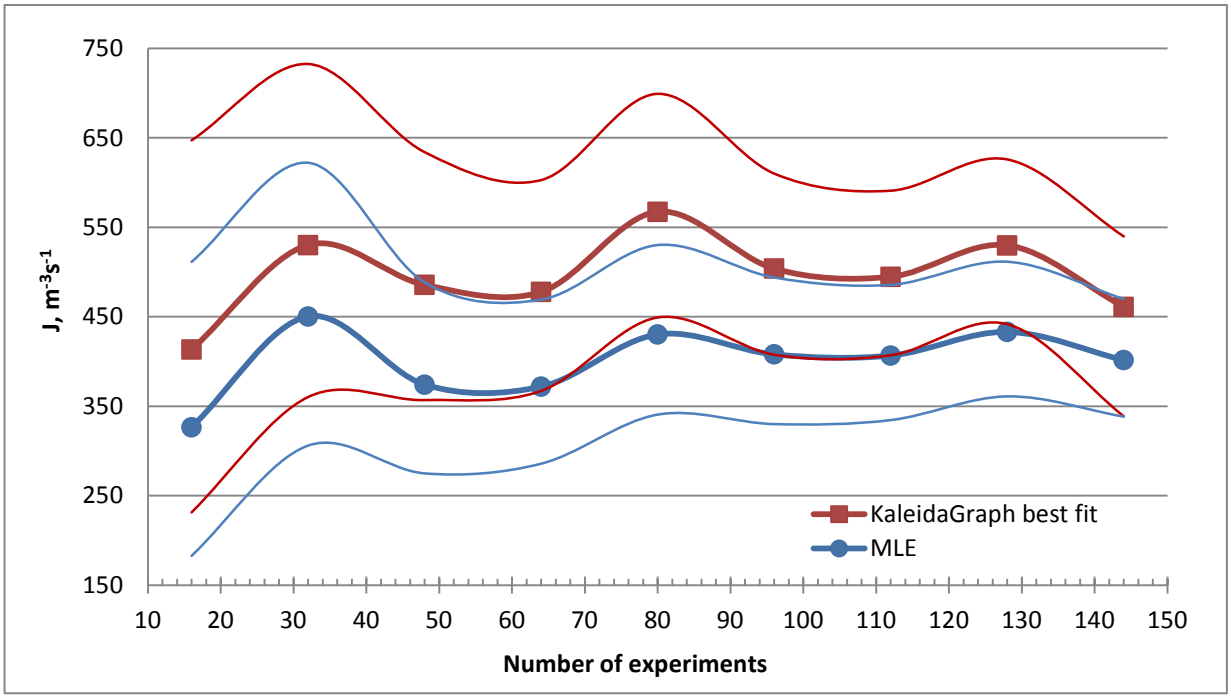

Figure 35 – Comparison of estimation of nucleation rate for KaleidaGraph® and MLE methods by different number of experiments for data from Kulkarni et al. study

Analyzing trendlines for best-fit KaleidaGraph® and MLE methods in Figure 35, one can state that they show similar behavior. As well as for the data produced by Jiang and ter Horst, the MLE method does not give a straight trendline. It is of great importance to notice that fluctuations for the MLE trend are decreasing in magnitude with increasing of the amount of parallels, so there is stabilization in the behavior. Also it should be said that MLE method consequently predicts lower nucleation rates than the best-fit and the differences between the estimates by two methods for the various numbers of experiments are approximately the same.

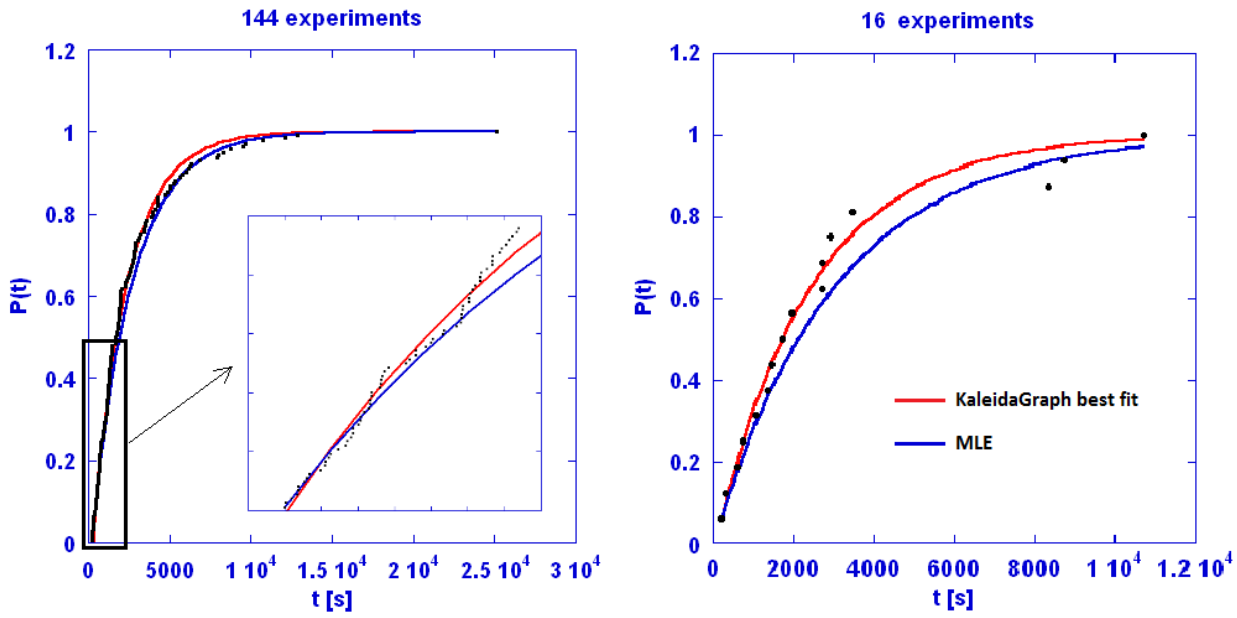

Figure 36 – Comparison of the fitting curves for the KaleidaGraph® and the MLE methods for 144 and 16 experiments data from Kulkarni et al. study

As well as for Jiang and ter Horst study the same comments can be made about the behavior of fitting lines for the methods (see Figure 36): the MLE curve differs not so much from one generated by KaleidaGraph®, although does not matches a considerable part of induction times cumulative distribution.

Kulkarni et al. declared the nucleation rate to be 450±2 $m^{-3}s^{-1}$ for a supersaturation ratio of 1.40 and 144 experiments. We obtained estimates of 461.14 $m^{-3}s^{-1}$ by best-fit method and 401.41 $m^{-3}s^{-1}$ by MLE technique. Thereby, as for Jiang and ter Horst study, the method they used should be similar to KaleidagGraph® best-fit.

4.2 Calculation and analysis of activation energy and critical nuclei radius

The calculated nucleation rates further have been used in order to calculate the activation energy of hydrate nucleation and the critical radius of a nucleus.

It was assumed that the case with all of the available experiments in each series gives the most reliable result.

The calculation procedure is described on the example of that "the most accurate" case for the induction times estimated by KaleidaGraph® best-fit method for cell#3. Thus we have 60 parallels at 13.00 °C and 30 parallels in both 11.75 °C and 14.25 °C series. All the required data are presented in Table 14.

Table 14 – Data for calculation of activation energy for the case of 30 parallels at 11.75 °C and 14.25 °C and 60 parallels at 13.0 °C; the induction times calculated in KaleidaGraph®

T, °C	T_a, K	1/T, K^{-1}	J, min^{-1}	ln(J)
11.75	284.90	0.0035100	0.48525	-0.7231
13.00	286.15	0.0034947	0.26504	-1.3279
14.25	287.40	0.0034795	0.05451	-2.9094

Plotting 3 points with abscissa $1/T$ and ordinate $ln(J)$ and applying linear regression method in KaleidaGraph® we obtain an equation $y = -251.73 + 71559x$. Consequently, using the value of the slope we get $E_a = 71559 \cdot 1.38076 \cdot 10^{-23} = 0.9881 \cdot 10^{-18}$ J.

The critical nuclei radius is then calculated according to Formula (15):

$$r_c = \sqrt{\frac{3 \cdot 0.9881 \cdot 10^{-18}}{4 \cdot \pi \cdot 0.0276}} = 29.23 \text{ Å}$$

In comparison, Englezos et al. obtained values of the critical radius of methane hydrate to be in the range of 30 – 170 Å. Nerheim et al. proposed approximately 100 Å. Larson and Garside estimated size of critical nuclei at around 32 Å using classical nucleation theory. Whereas computer simulations predict critical sizes to be around 14.5 Å [3, 12]. Therefore the calculated value of critical radius is assumed to be reasonable.

4.3.1 Calculations for cell #3

The calculation results for activation energy and critical radius for nucleation rates estimated by means of KaleidaGraph® best-fit method are presented in Table 15.

Table 15 – Results of calculations of activation energy and critical radius for SNG-2 in cell #3 by different number of experiments for nucleation rates estimated by KaleidaGraph® method

Number of experiments	E_a/k_B	E_a, 10^{-18} J	Deviation ,%	r_c, Å	Deviation,%
30-60-30	71559	0.9881	0	29.23	0
30	71600	0.9886	-0.0573	29.24	-0.0286
29	74044	1.0224	-3.4727	29.74	-1.7215
28	75026	1.0359	-4.845	29.93	-2.3938
27	85541	1.1811	-19.539	31.96	-9.3339
26	97420	1.3451	-36.139	34.11	-16.679
25	104540	1.4434	-46.089	35.33	-20.867
24	106130	1.4654	-48.311	35.6	-21.783
18	113020	1.5605	-57.94	36.74	-25.674
12	115110	1.5894	-60.86	37.08	-26.831
6	135560	1.8718	-89.438	40.24	-37.637

Reducing the number of parallels results in increase in both activation energy and critical radius. Reduction of the number of parallels has not so strong effect on the critical radius as on the activation energy. As can be seen from Table 15, the required number of parallels to obtain accuracies within 20% deviation is 27 for activation energy and 26 for critical radius. For deviation of 30 % it requires 12 experiments for critical radius and still 27 experiments for activation energy.

Table 16 presents the calculation results for activation energy and critical radius for nucleation rates estimated by means of MLE method.

Table 16 – Results of calculations of activation energy and critical radius for SNG-2 in cell #3 by different number of experiments for nucleation rates estimated by MLE method

Number of experiments	E_a/k_B	E_a, 10^{-18} J	Deviation,%	r_c, Å	Deviation,%
30-60-30	66349	0.91612	0	28.15	0
30	66407	0.91692	-0.0874	28.16	0.0437
24	67554	0.93276	-1.8162	28.4	0.9040
18	67166	0.92740	-1.2314	28.32	0.6138
12	68487	0.94564	-3.2224	28.6	1.5984
10	64246	0.88708	3.1696	27.7	1.5976
8	61913	0.85487	6.6859	27.19	3.4008
7	113310	1.56454	-70.779	36.79	30.682
6	117520	1.62267	-77.124	37.46	33.088

The estimates produced by MLE method result in less discrepancy under reduction of the amount of parallels, compare to best-fit. Therefore the minimum acceptable number of experiments for estimation of the activation energy and the critical radius is relatively small and equal 8 experiments for both 20% and 30% accuracy levels.

4.3.2 Calculations for the large cell

The reference level of accuracy for estimation of nucleation rate by best-fit method is represented by combination of 20 experiments for the series conducted at the temperature of 16.75 and by 12 experiments for both series at 15.5 °C and 17.75 °C. The calculation results are presented in Table 17.

Table 17 – Results of calculations of activation energy and critical radius for SNG-7 in the large cell by different number of experiments for nucleation rates estimated by KaleidaGraph® method

Number of experiments	E_a/k_B	E_a, 10^{-18} J	Deviation,%	r_c, Å	Deviation,%
12-20-12	144060	1.98912	0	41.48	0
12	145330	2.00666	-0.8816	41.66	-0.4398
10	146470	2.02240	-1.6729	41.82	-0.8330
8	143030	1.97490	0.71498	41.33	0.3581
6	119490	1.64987	17.0554	37.78	8.9261

The analysis of the data produced by KaleidaGraph® shows that 6 experiments is a sufficient amount that gives acceptable accuracy for both limits of 20% and 30%. The deviation in the estimate value is 17% for the activation energy and is less than 9% for the critical nucleus radius.

The reference level of accuracy for estimation of nucleation rate by MLE method is represented by combination of 19 experiments for the series conducted at the temperature of 16.75 and by 12 experiments in both series at 15.5 °C and 17.75 °C. The results of calculations of activation energy and critical radius are presented in Table 18

Table 18 – Results of calculations of activation energy and critical radius for SNG-7 in the large cell by different number of experiments for nucleation rates estimated by MLE method

Number of experiments	E_a/k_B	E_a, 10^{-18} J	Deviation,%	r_c, Å	Deviation,%
12-19-12	138570	1.91332	0	40.68	0
12	138180	1.90793	0.2814	40.62	0.14082
10	139060	1.92008	-0.3536	40.75	-0.1766
8	139810	1.93044	-0.8949	40.86	-0.4464
6	137360	1.89661	0.8732	40.5	0.43756

MLE estimates again show less discrepancy than these produced by KaleidaGraph®. The maximum deviation is obtained for 8 experiments. It is less than 1% for activation energy and less than 0.5 % for a nuclei critical radius. Thus the minimum acceptable number of parallels for both 20% and 30% accuracy levels is 6 for activation energy as well as for critical radius.

It was supposed that fit of Ln J as function of 1/T to a straight line in linear regression analysis could be used as criterion evaluating the minimum number of parallels required.

In order to determine how well datapoints fit a straight line coefficient of determination (R^2 values) should be used. The value of the coefficient close to 1 would indicate that $Ln\,J$ as function of $1/T$ approximately on a straight line.

The R^2 values have been found in KaleidaGraph® for both cell#3 and the large cell for different number of parallels in a series. The results of the analysis are shown in Table 19.

Table 19 – R^2 values from linear regression in KaleidaGraph®

Cell #3			Large cell		
Number of parallels	R^2 value for best-fit	R^2 value for MLE	Number of parallels	R^2 value for best-fit	R^2 value for MLE
30-60-30	0.96768	0.98198	12-20-12	0.88686	-
30	0.99943	0.98870	12-19-12	-	0.86279
24	0.98347	0.98525	12	0.93323	0.84600
18	0.99817	0.98525	10	0.93470	0.84239
12	0.99992	0.99605	8	0.91957	0.82297
6	0.99186	0.96772	6	0.88802	0.81453

From the table one can see that there is no any conclusive evidence of any relationship between the number of experiments and R^2 values. It can be explain by the fact that stochastic nature of the process can significantly impact estimated values of nucleation rate which in turn determine the slope of the plot $Ln\,J$ vs $1/T$.

It is interesting to compare results of the present study with results obtained by Eirik Høvring [12] in his MSc study. Eirik obtained value of critical radius to be 34.78 Å on SNG-7 in cell#3 for 6 experiments in each of 3 experimental series at 13.0 ºC, 14.25 ºC and 15.5 ºC. Other experimental conditions were as follows: pressure is 90 bar, stirring rate is 619 rpm and cooling rate is 6.75 ºC/h.

In the present MSc work for experiments on SNG-7 in the large cell the value of critical radius was estimated to be 37.78 Å for 6 experiments in a series. The experimental conditions were as follows: 90 bar, 450 rpm, 6 ºC/h.

According to Mullin [41] the critical size is function of subcooling and increases with increasing temperature and decreasing subcooling. This probably could explain the difference between the present study (37.78 Å) and Eirik's (34.78 Å) critical size measurements, since Eirik conducted experiments at lower temperatures. On the other hand, induction times, which values would affect estimation of critical radius, appear to be apparatus-dependent [3]. Larger cells have larger contact area of the interface between gas and aqueous phases, so nucleation can occur more readily. Some other parameters as agitation speed and cooling rate may also influence nucleation. Therefore it cannot be any strong conclusions made since many parameters are different between the studies. Moreover there is no certainty if 6 parallels in a series is enough to obtain accurate estimation of critical radius (e.g. see Table 15).

5 SUMMARY OF RESULTS

5.1 Critics on the approach used in the present work

The approach used in the present MSc work could be arguable. For instance, if we perform more experiments after we did the analysis according to this approach for fewer experiments, then the results of previous analysis would be unreliable since we set a new reference level that represents 100% accuracy. It can be assumed that the approach is reliable only for relatively large number of parallels, when the later experiments do not add new information that would improve a distribution. However, we will still have some discrepancy between two different reference levels, which could be assumed as negligible for large amount of experiments.

The method of creating cumulative probability after assigning equal probabilities for detected induction times can be considered as too simplified. Possibly the probabilities should be assigned as follows. First, to create a probability density function (PDF), a time scale should be divided by equal time intervals and probabilities for a given interval should be assigned according to a fraction of induction time values falling to the interval. Moreover, the last time interval should begin from longer time that the largest observed and it should not have a right border since we never know the longest possible induction time for a given series concerning highly stochastic nature of hydrate nucleation phenomenon. The probability of the last time interval, despite being low, cannot be assigned with 100% certainty, but rather should be assumed. Then creating CDF is simply attained by consequent summation of probabilities of the intervals and arranging them along the same time scale. Thus we will obtain more reasonable probability to detect hydrate formation within a given time interval.

5.2 Induction time

All the minimum acceptable numbers of experiments for estimation of a nucleation rate by both KaleidaGraph® and MLE methods, concerning 20% and 30% accuracy levels are presented in Table 20.

Table 20 – Minimum acceptable number of experiments for estimation of the nucleation rate by KaleidaGraph® best-fit and MLE methods for the experimental series studied in the thesis work

Study		Cell #3			Large cell			Jiang and ter Horst	Kulkarni et al.
Temp, ºC		11.75	13.00	14.25	15.5	16.75	17.75	25.00	25.00
Total number of parallels		**30**	**60**	**30**	**12**	**20**	**12**	**80**	**144**
Accuracy limit of 20%	Best-fit	23	39	28	10	18	10	48	96
	MLE	23	46	24	10	17	10	16	16
Accuracy limit of 30%	Best-fit	20	36	27	10	16	6	16	16
	MLE	20	40	21	8	6	9	16	16

From Table 20 we can see that the minimum number of parallels required to obtain accuracies within both set limits in general increase with increasing the total number of experiments. Thus it can be assumed that 60 parallels is not enough for our approach to make some conclusions on the minimum necessary amount of experiments. Therefore there may be need in further investigation by increasing the total number of parallels performed to see how much the results may change.

Stochastic nature of nucleation normally tends to decrease with increasing of driving force [3], so in general the minimum required number of experiments should decrease with increasing subcooling because distribution of induction times would be more uniform within an experimental series. However this hypothesis was not confirmed in present study.

Results of analysis of experimental data from Kulkarni et al. and Jiang and ter Horst studies with large amount of parallels show that minimum number of experiments required to produce acceptable accuracy could possibly be only several tens. However, these results cannot be related to gas hydrates nucleation in the present study since the experimental systems are very different (glass vials, uniform distribution of solute in solution, fast cooling to experimental temperature).

5.3 Activation energy and critical radius

All the minimum acceptable numbers of experiments for estimation of activation energy and critical nuclei radius by both KaleidaGraph® and MLE methods, concerning 20% and 30% accuracy levels are presented in Table 21.

Table 21 – Minimum acceptable number of experiments for estimation of activation energy and critical radius by KaleidaGraph® best-fit and MLE methods

Cell used for experiments		Cell #3		Large cell	
Method		Best-fit	MLE	Best-fit	MLE
Total number of parallels for the series at 15.5 ºC, 16.75 ºC and 17.75 ºC respectively		**30-60-30**		**12-20-12**	**12-19-12**
Accuracy limit of 20%	Estimate for E_a	27	8	6	6
	Estimate for r_c	26	8	6	6
Accuracy limit of 30%	Estimate for E_a	27	8	6	6
	Estimate for r_c	12	8	6	6

From the results for cell#3 it may be assumed that the influence on the critical radius by reduction of the number of parallels is not as strong as the effect of the same reduction on the activation energy.

If we look at the results of minimum acceptable number of experiments for the data produced on the large cell we can suppose that the high dispersion of induction times in experimental series not so highly affects the results of calculations of the activation energy and the critical radius of a hydrate cluster.

Acceptable accuracies for estimates of activation energy and critical radius were obtained even though the number of parallels was reduced to 6 for the large cell experiments, and the discrepancies from the case with all available parallels included were relatively low. In order to decide where this will be valid in other situations, the number of parallels was supposed to depend on "goodness" of fit of $Ln\, J$ as function of $1/T$ to a straight line that is described by R^2 value. However, there was not found clear relationship between the number of parallels and the coefficient of determination (R^2 value).

5.4 Lag time

The lag time variation has not been analyzed in the present MSc study since this parameter has very confined practical application. As was mentioned in the theoretical part, the lag time is a delay in detection of nucleation event caused by the reason that macroscopic instruments are unable to detect the first stable nucleus formation. Therefore the probability of hydrate detection within this time interval equal zero.

Negative lag times estimated by both best-fit and MLE methods (see Table 11) mean that there is a probability that nucleation process could commence during the cooling section prior to start of stirring (beginning of experiment). In general it looks reasonable since the system is in metastable region long time before the actual start of the experiment.

5.5 Best-fit and MLE techniques comparison

Best-fit method has been extensively used for estimation of nucleation rate in many studies and investigations carried out in the hydrate laboratory at Department of Petroleum Engineering, University of Stavanger.

It should be mentioned that Jiang and ter Horst and Kulkarni et al. in their study used the method similar to KaleidaGraph® that is confirmed by results of the present study.

If we assume that correct estimation of nucleation rate depends on matching the experimental data by fitting curve, then we can conclude that KaleidaGraph® best-fit method is more accurate than MLE. Indeed, in most of the cases discussed in the present work KaleidaGraph® matches well a main part of cumulative distribution of experimentally measured induction times, whereas a fitting curve for MLE matches mainly the first and the last points.

From Figures 23 and 24 it may be noticed that shift of MLE fitting curve to the left results in estimates of nucleation rate by this method that are closer to KaleidaGraph® estimates. Thus it can be concluded that MLE underestimate the real value of nucleation rate (if the previous assumption about fitting curves still holds).

MLE method may be considered as less time consuming, since the estimation of induction time is done by means of equation and not regression analysis as for best-fit. MLE

method in principle could be used e.g. for comparison of different inhibition techniques, when one is interested to find deviation in parameter, and not its absolute value. But further analysis is required to find applicability borders of this method.

Based on the results of the laboratory experiments MLE method in general shows less discrepancy than best-fit when we reduce a number of parallels. However, we cannot see this tendency from the results of analysis of data from Jiang and ter Horst and Kulkarni et al. studies. This controversial behavior could be caused by different distribution of induction times within a given series. It is required to run simulations with different rearrangements of induction times within one series to quantify the impact of induction time sequence on the method.

It was noted that in some cases KaleidaGraph® gives greater values of lag time than the minimum measured induction time. For instance, the minimum induction time in Jiand and ter Horst study for 80 parallels is 1027 s, and KaleidaGraph® estimation of lag time is 1164.5 s. But this is in conflict with definition of lag time. From this point of view MLE technique looks more reliable since the estimation of lag time by this method always gives smaller values than the minimum observed induction time in a series. This estimation is guaranteed by the equation this method based on.

5.6 Accuracy limit

The acceptable limit of accuracy (tolerance) highly affects final results for the approach used in the present work. In general this limit depends on the purpose of the study. In experimental work inaccuracies of the order 20 to 30% may be acceptable [20]. But for the industrial purposes these limits are considerably stricter.

In the present study acceptance of 30% tolerance reduces a minimum number of parallels considerably as compared to a limit of 20%. It is well illustrated on the example of the analysis of data from Jiang and ter Horst and Kulkarni et al. studies, where changing of the limit from 20% to 30% lowers the minimum number from 48 and 96 to 16 respectively for estimation by best-fit method.

As a conclusion, the number of experiments to perform should always be increased if lower limit of acceptable accuracy is set.

5.7 Equipment influence

The analysis of Jiang and ter Horst, Kulkarni et al. studies showed that in general the stochastic behavior of nucleation may be lowered if one uses an experimental vessel with extremely smooth inner surface. However, from the personal communication with Thor Martin Svartås [20], supervisor of the present work, it was known that i.e. sapphire cells also may show unexpected behavior that cannot be explained.

A size of an experimental vessel also influences the results. For instance, in the large cell it is possible to form hydrates at subcooling lower than that for smaller cell. This can be explained by fact that the volume involved in the nucleation process becomes larger with increase in diameter of a cell. It was possible to form hydrates at 18.25 ºC on the large cell, although the induction time was several hundred minutes [20]. The analogy can be drawn between high pressure cells and pipelines. The larger the diameter, the greater the probability to form hydrates in the pipeline at the lower subcooling degree.

In general it may be introduced a “cell constant”, that determines a part of activation energy due to contribution of the cell system itself [12]. The “cell constant” can be changed due to problems of mechanical character or due to precipitation of the chemicals on the walls. Therefore this concept can be used in explanation of unexpected results / unpredictable behavior of the cell in particular cases.

6 CONCLUSIONS AND FUTURE WORK

The main purpose of the present work was to determine whether it is acceptable to conduct just few experiments in order to estimate key parameters of gas hydrate nucleation.

Based on the results of statistical analysis of experimental data in the present study it can be concluded that the minimum number of experiments to obtain a nucleation rate estimate should possibly be not less than several tens to provide accuracy within 20-30%. Analysis of a series with more parallels is required in order to make some conclusions about the exact number.

For the conducted experiments acceptable accuracies were obtained for estimation of nuclei size even though the number of parallels was reduced to 6 for the large cell experiments. However, there was not found clear relationship between number of parallels and fit of $Ln\,J$ as function of $1/T$ to a straight line. Therefore other possible criteria evaluating the minimum parallels required should be found and verified

The set accuracy limit that determines maximum acceptable deviation in parameter estimate has great influence on the required number of parallels. In general increasing the number of parallels increases the chances on getting a representative estimate of a parameter of interest. In any case one should conduct more experiments in order to obtain more accurate results.

Best-fit KaleidaGraph® technique is possibly more accurate for estimation of nucleation rate than MLE based on the fitting curves comparison. At the same time MLE method is less labor intensive and possibly can be used to find relative deviation in parameter value, i.e. for comparison of different hydrate inhibitor treatments. The reliability of MLE method for different conditions should be further tested experimentally.

Analysis of different rearrangements of induction times within one series could clarify the impact of stochastic nature of nucleation on the approach used in the present study for both methods.

In order to reduce overall time of analysis it is beneficial to develop a software that would use LabVIEW® logs as input data, process them using best-fit or MLE algorithm and produce key nucleation parameters as output.

REFERENCES

1. Muller-Bongartz, B., T. Wildeman, and R. Sloan Jr. *A Hypothesis For Hydrate Nucleation Phonemena*. in *The Second International Offshore and Polar Engineering Conference*. 1992. International Society of Offshore and Polar Engineers.
2. Bishnoi, P.R. and V. Natarajan, *Formation and decomposition of gas hydrates.* Fluid Phase Equilibria, 1996. **117**(1): p. 168-177.
3. Sloan Jr, E.D. and C. Koh, *Clathrate hydrates of natural gases*. 2007: CRC press.
4. Jiang, S. and J.H. ter Horst, *Crystal nucleation rates from probability distributions of induction times.* Crystal Growth & Design, 2010. **11**(1): p. 256-261.
5. Kulkarni, S.A., et al., *Crystal nucleation kinetics from induction times and metastable zone Widths.* Crystal Growth & Design, 2013. **13**(6): p. 2435-2440.
6. Fandiño, O. and L. Ruffine, *Methane hydrate nucleation and growth from the bulk phase: Further insights into their mechanisms.* Fuel, 2014. **117**: p. 442-449.
7. Ohmura, R., et al., *Statistical study of clathrate-hydrate nucleation in a water/hydrochlorofluorocarbon system: Search for the nature of the "memory effect".* The Journal of Physical Chemistry B, 2003. **107**(22): p. 5289-5293.
8. Duchateau, C., et al., *Laboratory evaluation of kinetic hydrate inhibitors: a procedure for enhancing the repeatability of test results.* Energy & Fuels, 2009. **23**(2): p. 962-966.
9. Toschev, S., A. Milchev, and S. Stoyanov, *On some probabilistic aspects of the nucleation process.* Journal of Crystal Growth, 1972. **13**: p. 123-127.
10. Wilson, P., D. Lester, and A. Haymet, *Heterogeneous nucleation of clathrates from supercooled tetrahydrofuran (THF)/water mixtures, and the effect of an added catalyst.* Chemical engineering science, 2005. **60**(11): p. 2937-2941.
11. Palm, M., *The effect of some hydrocarbon liquids on methane hydrate nucleation.* 2012.
12. Høvring, E., *On the activation energy for the formation of a critical size water cluster in structure I and structure II gas hydrates.* 2012.
13. Kashchiev, D. and A. Firoozabadi, *Driving force for crystallization of gas hydrates.* Journal of crystal growth, 2002. **241**(1): p. 220-230.
14. Hammerschmidt, E., *Formation of gas hydrates in natural gas transmission lines.* Industrial & Engineering Chemistry, 1934. **26**(8): p. 851-855.
15. Taylor, C.E. and J.T. Kwan, *Advances in the study of gas hydrates*. 2004: Springer.
16. Fink, J., *Petroleum Engineer's Guide to Oil Field Chemicals and Fluids*. 2011: Gulf Professional Publishing.
17. Carroll, J., *Natural gas hydrates: a guide for engineers*. 2009: Gulf Professional Publishing.
18. Макогон, Ю., *Природные газовые гидраты: распространение, модели образования, ресурсы.* Российский химический журнал, 2003. **47**(3): p. 70-79.
19. Lasich, M., et al., *Phase equilibria of methane clathrate hydrates from Grand Canonical Monte Carlo simulations.* Fluid Phase Equilibria, 2014. **369**: p. 47-54.
20. Svartås, T.M., *Personal communication.* February - June 2014.
21. Carson, D.B. and D.L. Katz, *Natural gas hydrates.* Transactions of the AIME, 1942. **146**(01): p. 150-158.
22. Thomas, E., *Clathrates: little known components of the global carbon cycle.* Wesleyan University (November 2004), 2007.
23. Makogon, Y.F., *Hydrates of hydrocarbons*. 1997: Pennwell Books.
24. Buffett, B. and D. Archer, *Global inventory of methane clathrate: sensitivity to changes in the deep ocean.* Earth and Planetary Science Letters, 2004. **227**(3): p. 185-199.
25. Kvenvolden, K.A., *Gas hydrate and humans.* Annals of the New York Academy of Sciences, 2000. **912**(1): p. 17-22.
26. Atilhan, M., et al., *Natural Gas Hydrates*. 2012. 542.

27. Makogon, Y., S. Holditch, and T. Makogon, *Natural gas-hydrates—A potential energy source for the 21st Century.* Journal of Petroleum Science and Engineering, 2007. **56**(1): p. 14-31.
28. Sloan, E.D., C.A. Koh, and A. Sum, *Natural gas hydrates in flow assurance*. 2010: Gulf Professional Publishing.
29. Hirai, H., et al., *Structural changes in gas hydrates and existence of a filled ice structure of methane hydrate above 40GPa.* Journal of Physics and Chemistry of Solids, 2004. **65**(8): p. 1555-1559.
30. Hirai, H., et al., *Structural changes of argon hydrate under high pressure.* The Journal of Physical Chemistry B, 2002. **106**(43): p. 11089-11092.
31. Abay, H.K., T.M. Svartaas, and W. Ke, *Effect of Gas Composition on sII Hydrate Growth Kinetics.* Energy & Fuels, 2011. **25**(4): p. 1335-1341.
32. Kashchiev, D. and A. Firoozabadi, *Induction time in crystallization of gas hydrates.* Journal of crystal growth, 2003. **250**(3): p. 499-515.
33. Huo, Z., et al., *Hydrate plug prevention by anti-agglomeration.* Chemical Engineering Science, 2001. **56**(17): p. 4979-4991.
34. Clark, L.W. and J. Anderson. *Low Dosage Hydrate Inhibitors (LDHI): Further advances and developments in flow assurance technology and applications concerning oil and gas production systems*. in *IPTC 2007: International Petroleum Technology Conference*. 2007.
35. Frostman, L. and J. Przybylinski. *Successful applications of anti-agglomerant hydrate inhibitors*. in *SPE International Symposium on Oilfield Chemistry*. 2001. Society of Petroleum Engineers.
36. Mullin, J., *Crystallization*. 1993: Butterworth-Heinemann, Oxford.
37. Anklam, M.R. and A. Firoozabadi, *Driving force and composition for multicomponent gas hydrate nucleation from supersaturated aqueous solutions.* The Journal of chemical physics, 2004. **121**(23): p. 11867-11875.
38. Arjmandi, M., et al., *Is subcooling the right driving force for testing low-dosage hydrate inhibitors?* Chemical engineering science, 2005. **60**(5): p. 1313-1321.
39. Kvamme, B., *Kinetics of hydrate formation from nucleation theory.* International Journal of Offshore and Polar Engineering, 2002. **12**(4): p. 256-263.
40. Tung, H.-H., et al., *Crystallization of organic compounds: an industrial perspective*. 2009: John Wiley & Sons.
41. Mullin, J.W., *Crystallization*. 2001: Butterworth-Heinemann.
42. Kashchiev, D., *Nucleation*. 2000: Butterworth-Heinemann.
43. Wu, Q. and B. Zhang, *Memory effect on the pressure-temperature condition and induction time of gas hydrate nucleation.* Journal of Natural Gas Chemistry, 2010. **19**(4): p. 446-451.
44. LIU, Y., et al., *Memory effects of structure I and II gas hydrates.* Acta Physico-Chimica Sinica, 2011. **27**(6): p. 1305-1311.
45. Takeya, S., et al., *Freezing-memory effect of water on nucleation of CO2 hydrate crystals.* The Journal of Physical Chemistry B, 2000. **104**(17): p. 4164-4168.
46. Rodger, P., *Methane hydrate: melting and memory.* Annals of the New York Academy of Sciences, 2000. **912**(1): p. 474-482.
47. Wilson, P. and A. Haymet, *Hydrate formation and re-formation in nucleating THF/water mixtures show no evidence to support a "memory" effect.* Chemical Engineering Journal, 2010. **161**(1): p. 146-150.
48. Buchanan, P., et al., *Search for memory effects in methane hydrate: structure of water before hydrate formation and after hydrate decomposition.* The Journal of chemical physics, 2005. **123**(16): p. 164507.
49. Davies, S.R., et al., *Studies of hydrate nucleation with high pressure differential scanning calorimetry.* Chemical Engineering Science, 2009. **64**(2): p. 370-375.
50. Becker, R. and W. Döring, *Kinetische behandlung der keimbildung in übersättigten dämpfen.* Annalen der Physik, 1935. **416**(8): p. 719-752.
51. Vekilov, P.G., *Nucleation.* Crystal growth & design, 2010. **10**(12): p. 5007-5019.

52. Zatsepina, O.Y. and B. Buffett, *Nucleation of CO_2-hydrate in a porous medium.* Fluid Phase Equilibria, 2002. **200**(2): p. 263-275.
53. Zheng, M., *Penalized Maximum Likelihood Estimation of Two-Parameter Exponential Distributions*. 2013, University of Minnesota.
54. Ross, S.M., *Introduction to probability and statistics for engineers and scientists*. 2009: Academic Press.
55. Zar, J., *Biostatistical Analysis. 5th*. 2009, Prentice Hall, Upper Saddle River, NJ. USA.
56. Epstein, B., *Estimation of the parameters of two parameter exponential distributions from censored samples.* Technometrics, 1960. **2**(3): p. 403-406.
57. Ernst, M.D., *Permutation methods: A basis for exact inference.* Statistical Science, 2004. **19**(4): p. 676-685.
58. Legendre, P. and L. Legendre, *Numerical ecology. 2nd.* Elsevier Science, Amsterdam, NL, 1998.
59. *KaleidaGraph® version 4.0 tutorial*.
60. Golland, P. and B. Fischl. *Permutation tests for classification: towards statistical significance in image-based studies*. in *Information processing in medical imaging*. 2003. Springer.
61. Nichols, T.E. and A.P. Holmes, *Nonparametric permutation tests for functional neuroimaging: a primer with examples.* Human brain mapping, 2002. **15**(1): p. 1-25.
62. Acid, S., L.M.d. Campos, and J.F. Huete, *Estimating probability values from an incomplete dataset.* International Journal of Approximate Reasoning, 2001. **27**(2): p. 183-204.
63. Kaplan, E.L. and P. Meier, *Nonparametric estimation from incomplete observations.* Journal of the American statistical association, 1958. **53**(282): p. 457-481.
64. Kvam, P.H. and B. Vidakovic, *Nonparametric statistics with applications to science and engineering*. Vol. 653. 2007: John Wiley & Sons.

APPENDIX A

The listing of program code in R Studio® for calculation of confidence interval and running permutation test are presented below:

Function for calculation of estimate by MLE and confidence interval

```
estrate=function(data,a=0.05){
  n=length(data)
  mean=mean(data)
  min=min(data)
  J=(n-1)/(n*(mean-min))
  tau0=(n*min-mean)/(n-1)
  CI_lower=J*qchisq(a/2,df=2*n-2)/(2*n-2)
  CI_upper=J*qchisq(1-a/2,df=2*n-2)/(2*n-2)
  cat("tau0=",tau0,"\n")
  cat("J=",J,"\n")
  cat((1-a)*100,"% confidenceinterval for J: (",CI_lower,",", CI_upper,")","\n")
}
```

Function for execution of permutation test

```
estdiff=function(data1,data2){
  n1=length(data1)
  mean1=mean(data1)
  min1=min(data1)
  n2=length(data2)
  mean2=mean(data2)
  min2=min(data2)
  tau01=(n1*min1-mean1)/(n1-1)
  J1=(n1-1)/(n1*(mean1-min1))
  tau02=(n2*min2-mean2)/(n2-1)
  J2=(n2-1)/(n2*(mean2-min2))
  diff=J1-J2
  return(diff)
}
permutetest=function(data1,data2,P){
  n1=length(data1)
  n2=length(data2)
  data=c(data1,data2)
  tdist=vector(length=P)
  for(i in 1:P){
    pdata=sample(data)
    pdata1=pdata[1:n1]
    pdata2=pdata[(n1+1):(n1+n2)]
    tdist[i]=estdiff(pdata1,pdata2)
  }
  tobs=estdiff(data1,data2)
  pvalue=min(length(tdist[tdist<tobs])/P,length(tdist[tdist>tobs])/P)
  return(pvalue)
}
```

APPENDIX B

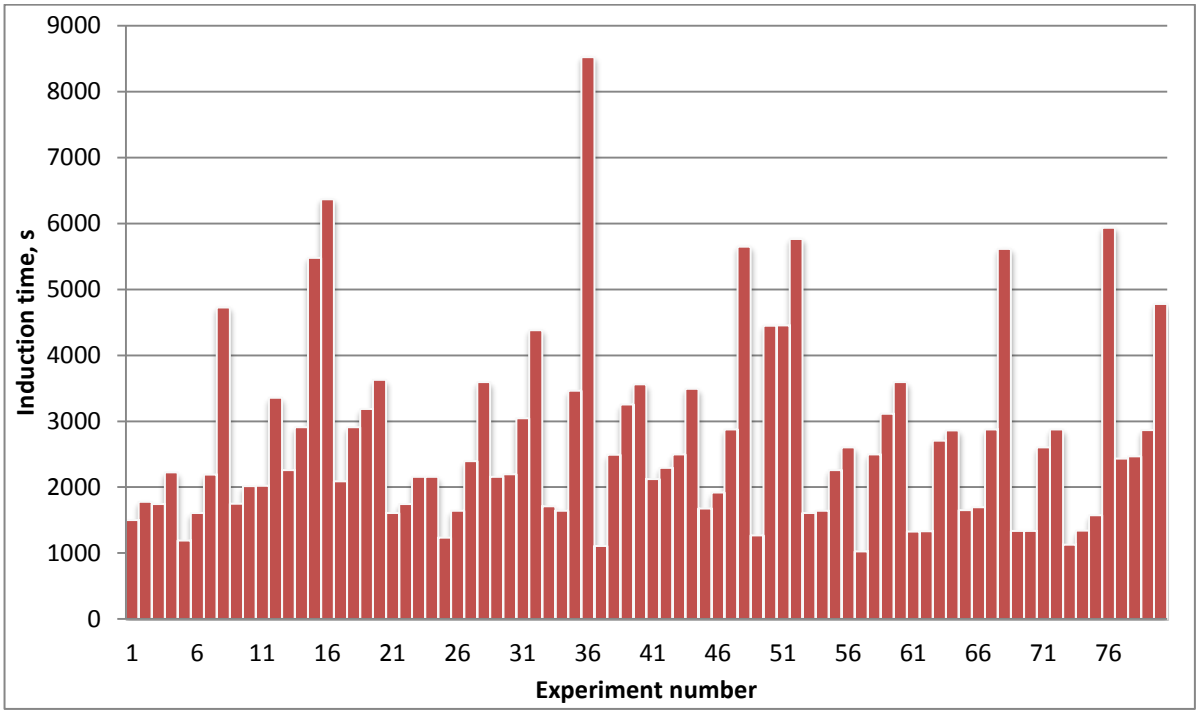

Figure B1 – The induction time in 80 measurements for m-aminobenzoic acid in water/ethanol mixture at the supersaturation ratio S = 1.96 and a temperature of 25 Cº for data from Jiang and ter Horst study [modified from 10]

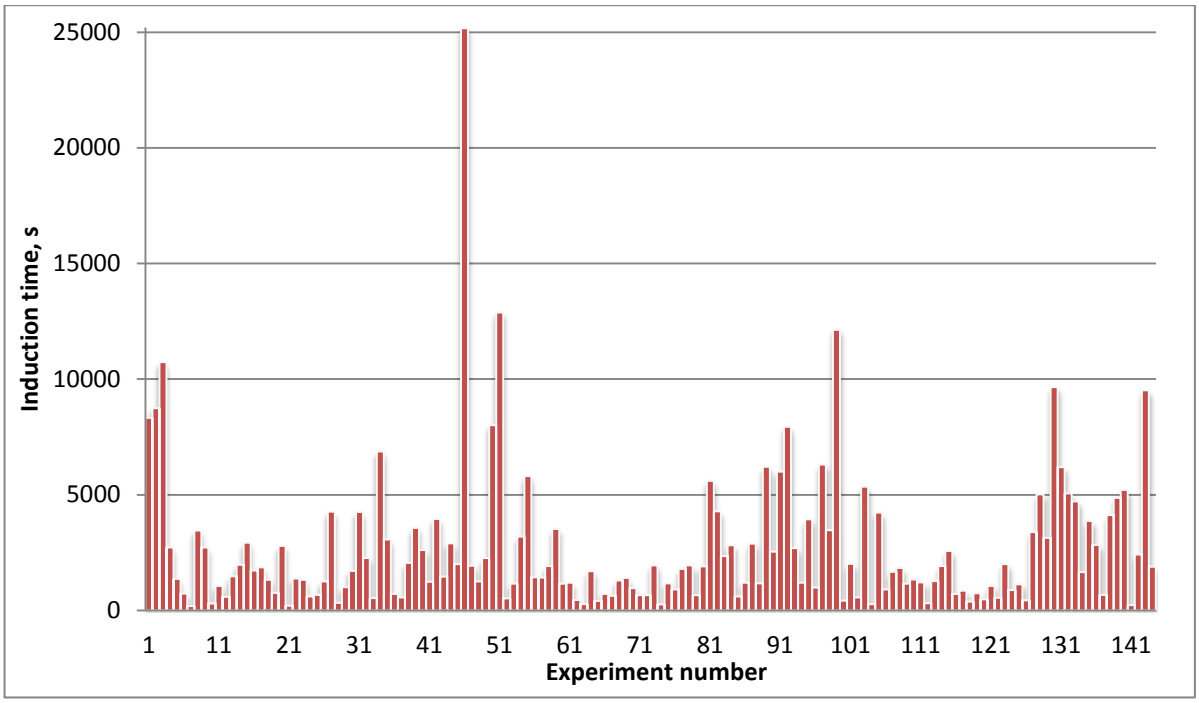

Figure B2 – The induction time in 144 experiments for isonicotinamide in ethanol at a supersaturation ratio S = 1.40 and a temperature of 25ºC for data from Kulkarni et al. study [modified from 11]