Exploring enclosed environments with floating sensors

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Exploring Enclosed Environments with Floating Sensors:
Mapping using Ultrasound

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op maandag 11 november 2019 om 13:30 uur

door

Hendrik Albert Duisterwinkel

geboren te Groningen

Dit proefschrift is goedgekeurd door de promotor en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. A.B. Smolders
1^e promotor:	prof.dr.ir. J.W.M. Bergmans
2^e promotor:	prof.dr.rer.nat. H.J. Wörtche
copromotoren:	dr. G. Dubbelman
	dr.rer.nat. E. Talnishnikh
leden:	Univ.-Prof. Dr.-Ing G. Ascheid (RWTH Aachen University)
	prof.dr.ir. M.K. de Kreuk (Technische Universiteit Delft)
	prof.dr.ir. J.A.M. Kuipers

Exploring Enclosed Environments with Floating Sensors:
Mapping using Ultrasound

Hendrik Albert Duisterwinkel

Summary

Exploring Enclosed Environments with Floating Sensors: Mapping using Ultrasound

This thesis is a contribution to a new field of study in which miniaturized sensor nodes are used to explore and inspect enclosed or otherwise difficult to access environments. Examples of such environments are pipelines, industrial mixing tanks and reactors, and underground channels and formations like reservoirs and geothermal sources. These man-made or natural environments have in common that they contain a (semi-)liquid medium and are enclosed and difficult to access. That is, the interior can hardly or not at all be accessed by remote probing using e.g. electromagnetic (EM) signals due to the shielding of metallic containment or surrounding ground layer(s) and the salinity of the liquid medium.

Essentially no techniques are currently available that deliver in situ measurements of structure and integrity in such environments. This thesis is concerned with a unique and entirely new 'go-with-the-flow' approach to perform such measurements. In our approach, sensor nodes are inserted into and later extracted from the environments. Access to the environments is obtained by using the liquid medium as carrier medium and adapting the sensor size to the characteristic dimensions of the environment.

The 'go-with-the-flow' approach requires that the nodes are small and lightweight. This has drastic effects on the allowable payload (battery, processing, sensors, etc.) that a node can carry. Furthermore, for a variety of reasons, communication in these environments is only realistically possible using ultrasound. Localizing the nodes is of key importance but hard to achieve. External position aids like the Global Navigation Satellite System (GNSS) or a custom network of beacons, will not work in these enclosed environments.

In this thesis we explore a novel approach in which localization is performed cooperatively by using a swarm of sensor nodes, consisting of many (e.g. hundreds or thousands) individual autonomous sensor nodes. Distances between node are measured by the nodes themselves, and an aggregation of these measurements is then used to estimate the locations of the node within the swarm. The use of ultrasound comes with its own set of challenges, like a low data rate and highly time-variant communication channels, making it hard to perform these distance measurements with high accuracy. The thesis attempts to address these challenges and to develop robust methods for cooperative localization.

Sensor nodes are localized by reconstructing their positions relative to each other once the nodes are extracted from the environment and their data read out. A novel reconstruction algorithm is developed aimed at being robust against highly corrupted distance measurements, using a Random Sampling Consensus method. Furthermore it has to deal with potentially ambiguous identification. In order to measure the distances between nodes, a ranging protocol is required to control the sequence of events needed for determining distances. A simulation chain is set up to simulate the ranging protocol for a number of realistic scenarios where the movement of the nodes is simulated in a variety of environments. The output of the simulated ranging measurements is then used as input for the reconstruction algorithm that estimates the positions of the majority of the nodes. In the considered scenarios, the protocol permits reliable reconstruction of the nodes positions.

The results of this study can serve as a basis for defining specifications for real-world functional nodes prototypes. The possible trade-offs among parameters of the ranging protocol are further studied for one particular environmental use case. The parameters that can be adjusted are e.g. the signal strength (communication range) and the number of identifying bits per message. Insight is gained on the effect that they have on e.g. the ranging latency, signal collisions, the fraction of distances that are successfully measured and the total number of messages transmitted to achieve this. Combining this knowledge with the effects that these parameters then have on the reconstruction of the nodes positions, allows one to properly adjust parameters depending on the different environments and application scenarios.

To investigate what other pieces of positional information can be gathered to aid in localizing the nodes, sensor nodes with a diameter of 4 cm and equipped with an inertial measurement unit (IMU) are constructed and used in a field test, floating through an underground pressurized water pipeline. The nodes log the sensed acceleration, rotation and the local magnetic field. The IMU data is used – unlike dead-reckoning techniques – to reveal structural information of the environment by studying the interaction between the nodes and the environment, and to obtain positional information based on the measured magnetic signature. Pipe components like junctions and hydrants are discovered and types of material of which the pipe is made are identified. Positional information is obtained by comparing the data with an incomplete map of the pipeline that was given beforehand.

Contents

Glossary

List of acronyms

3-D	three dimensional
ACK	acknowledgement (signal)
AOA	angle of arrival
BPSK	binary phase shift keying
CDMA	code division multiple access
CHOPS	cold heavy oil production with sand
CID	communication identifier
CO2	carbon dioxide
EM	electromagnetic
EU	european unions
G2O	¡software package for non-linear optimization¿
GNSS	global navigation satellite system
GPS	global position systems
GRE	glass-reinforced epoxy
IMU	inertial measurement unit
LIDAR	light detection and ranging
MWR	multi-way ranging
NFC	near-field communication
OMNeT++	network simulator software package
RANSAC	random sampling consensus
REQ	request (signal)
RSS	received signal strength
SLAM	simultaneous localization and mapping
TDOA	time difference of arrival
TOA	time of arrival
TOF	time of flight
TWR	two-way ranging
UID	unique identifier
WGS84	world geodetic system 1984

XWM   xploring wisemotes™

List of textual abbreviations

e.g.   exempli gratia: for example
i.e.   id est: that means, in other words

1
Introduction

This thesis focuses on a new approach for exploration of enclosed environments that contain a (semi-)fluid medium. Examples of these are man-made environments such as pipelines, industrial mixing tanks and reactors, and natural environments like underground channels, reservoirs and geothermal sources. Structural and in-situ information of these environments is requested by industry e.g. for guaranteeing operations, for safety purposes, or for optimizing processes. Current state-of-the-art technology does not provide this information at all, or not in a cost-effective manner. The main limiting factors are the remoteness and physical isolation of the system. We propose to explore these environments by means of a "go-with-the-flow" approach that uses small buoyant autonomous sensor nodes in large quantities. While traversing the environment, the sensor nodes perform measurements for parameters of interest, and to determine their relative positions. Measurements are stored in memory and read out when the nodes are extracted from the environment. After extraction, the stored data is analyzed in order to gain the desired knowledge.

1.1 Introducing the problem field
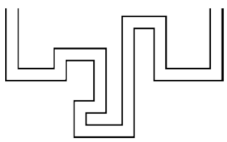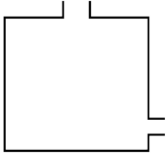
1.1.1 Enclosed environments

Our modern society depends in part on an extensive industrial infrastructure for extracting, distributing and processing fluids such as oil and water. In many cases the environments in which these fluids are contained are difficult to access. Examples of such environments are pipeline distribution systems, underground formations like reservoirs, and industrial tanks and reactors.

Pipeline distribution systems are a backbone technology for industries and society, ranging from the transport of oil from deep reservoirs to chemical plants, to the continuous distribution of fresh water from water purification plants to households in cities and villages.

Table 1.1: Generalized overview of categories of enclosed fluidic environments. Each with their own aspects and information needs.

	pipeline	reservoir	industrial tank
category	pipeline	reservoir	industrial tank
structure	known	internal structure unknown	known
integrity	unknown	-	known
dynamics	unknown	unknown	unknown
dimensions	1000×1×1 m	1000×0.1×0.1 m	10×10×10 m
examples	pipelines, sewage system	oil reservoir, potash mines, geothermal sources	mixing tank, reactor

Underground formations like oil and gas reservoirs have been the source of most of our energy production since the 19th century and both oil and gas serve as a basis for many products on which we rely every day. Not only do these formations hold resources that we use, they can also serve as (temporary) storage for liquids and gasses like hot and cold water and CO_2.

Industrial tanks and reactors are at the core of many chemical industries, and generally process specific input resources to desired (semi-finished) products. Examples of processes occurring inside these enclosed environments are boiling, sifting, mixing, etc.

1.1.2 Industrial exploration needs

The environments mentioned above are difficult to access either because they are designed to keep what is inside in and what is outside out, or because they owe their importance to their inaccessibility, like underground formations that trap valuable carbohydrates throughout the ages, forming a reservoir.

The combination of being important and being enclosed leads to the situation in which we seek more knowledge about them but are hard pressed to do so because they are not easily accessible. Each of the above mentioned examples belongs to a different category of enclosed environments with specific aspects, risks and exploration goals. These categories are summarized in Table 1.1.

Pipeline infrastructures transport large amounts of gaseous and (semi-)fluid media from one place to another and failures can have large effects on human life, industrial processes and environmental safety. Being able to estimate when pipelines are likely going to fail allows for timely preventive repairs [1, 2]. The original structure and locations of the pipelines are commonly known. The specific exploration goals for pipelines relate to measuring whether the structure is still the same and still intact, i.e. to measuring locally the integrity and system conditions. Key parameters of interest for assessing pipeline status are: leakages, remaining wall-thickness, corrosion, and sludge accumulation. In some cases, not even the exact location, the used materials or pipe-elements like flanges, bends, and junctions are known. Obtaining this information aids in predicting failures, performing proper maintenance, and preventing disasters. Regular inspection of the pipe conditions and integrity is essential to ensure a long service life, reduce risk of failures, and help in prioritizing the replacement order of pipe sections and components.

Extraction from and injection into **underground reservoirs** happens with little knowledge of the enclosed environment. Knowledge of the environment is often limited to macro-structural estimates like location, size and shape, obtained using radar probing, and possibly the chemical composition of the material/liquid inside, obtained using e.g. spectroscopic analysis. When the reservoirs are found and confirmed to be worthwhile to extract, boreholes are drilled into the reservoirs and the resources are pumped out until it is no longer profitable to do so. The lack of knowledge on the detailed internal structure makes it harder to efficiently extract resources, or to optimize storage of injected fluids. This is for example clearly seen in cold heavy oil production with sand (CHOPS) applications. In these CHOPS reservoirs the extraction rate of oil does not exceed 10% of the total amount; it is limited by the formation of open-channels in the sand-stone layer between production and injection well, causing a pressure drop [11]. The lack of detailed structural information here is preventing further extraction of oil, potentially valued at billions of euros. Ideally, a sensor system is sent down there, 600 meters deep, to measure in the harsh conditions the parameters of interest, like pressure profile, passage diameters, etc.

Industrial processes like mixing in **tanks and reactors** are often monitored solely based on measuring parameters in the input flow and in the output flow. The dynamics inside these enclosed environments, like flow conditions and interactions, and the state of the (multi-phase) fluids, like temperature distribution, can often only be inferred from the input, the output and their differences. The lack of direct information on what exactly is happening inside makes it harder for process engineers to optimize the intended processes, like increasing the yield or uniformity, and reducing the required energy and time for the process to complete [3].

The need to explore each of these enclosed environments is clear, whether it is to monitor and inspect pipeline infrastructure, to assess the detailed internal structure of underground reservoirs, or to study the inside of industrial tanks. However, their exploration is not trivial as they are shielded by a (metallic) containment to separate the inside from the outside, and/or are buried under several layers of soil. This makes it a hard challenge to inspect the integrity and potential failure mechanisms and to study the dynamics or detailed internal structure.
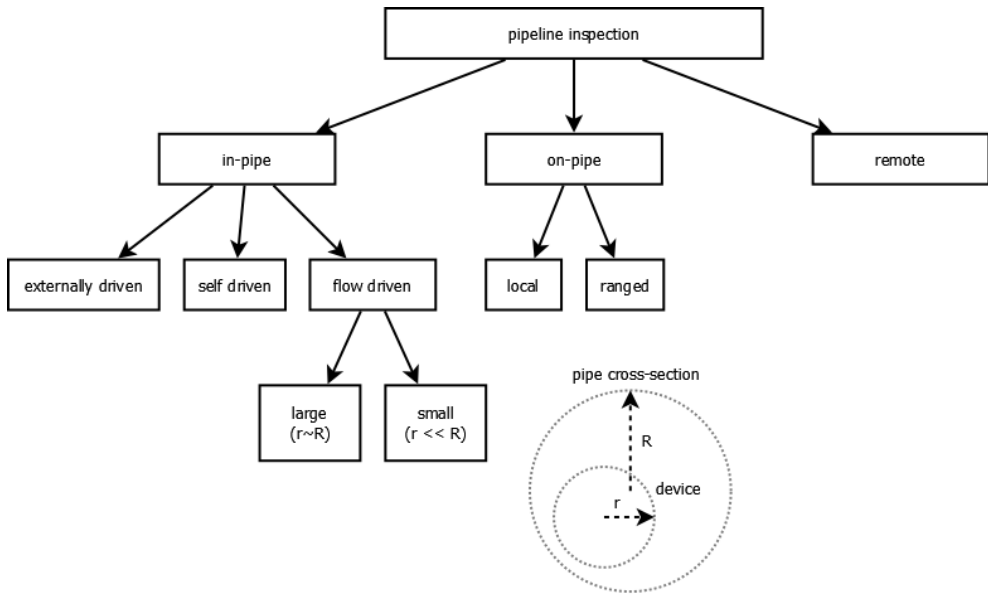
Figure 1.1: Categorizing pipeline inspection techniques.

1.1.3 Existing exploration techniques

The technology and tools available for the above exploration goals have advanced significantly in the last decades. In the following paragraphs available techniques are discussed per category of enclosed environment.

1.1.3.1 Pipelines

For pipelines, a large variety of techniques are available that can be separated into remote, on-pipe, and in-pipe techniques [2, 4, 5]. These are schematically shown in Fig. 1.1.

With **remote techniques** attempts are made to inspect the pipeline from the outside. Examples of these are ground penetrating radar [6] and acoustical surveyance. The knowledge gained from these techniques is often limited to macroscopic parameters like presence, location and type of material or object.

On-pipe inspection techniques have sensor systems physically attached to the pipeline or very close to the pipeline, often placed on the outside of it [2]. A distinction can be made between local systems and ranged systems. Local systems are e.g. ultrasound imaging to find defects in the containment with a very high resolution and on a very local scale, flow measuring systems using e.g. Doppler-based or induction-based sensors. Ranged systems include e.g. passively listening to the creation and development of leaks, (micro-)cracks and corrosion using a technique called acoustical emission [7], or active emitting sound- or EM-pulses and listening to the response to identify and locate defects along the pipeline. These systems can have ranges up to one or several kilometers. On-pipe inspection techniques require physical access to the pipelines, which prevents the use of these techniques for remote

or deep underground pipelines.

With **in-pipe inspection techniques**, devices are inserted into the pipeline and move through it for inspection along the [2, 4, 5]. Categorization of these techniques can be done in many ways. In this work we separate techniques by how the device moves through the pipeline. Externally driven or tethered devices require a direct line of access from an accessible location to the device itself. The range is limited to the length, weight and strength of the tether or driver. Examples of self-driven devices are robotic devices that crawl through the pipeline, or devices with rolling or fish-like actuators. The limiting factor for these devices is often the power required to manoeuvre through environments containing high viscosity fluids (relative to gaseous media) and with high flow-speeds.

Flow-driven devices use the flow of medium to propagate through the environment. A large variety of pipe inspection gauges have been developed that have a diameter close to that of the pipeline itself [2, 4, 5]. The devices remains relatively stable and the sensors can be placed close to the surface of the pipe to inspect it with high resolution. These large devices are generally limited to use in straight pipe sections with no pipe elements like valves and junctions. Often they get stuck due to (unexpected) accumulation of sludge. For many other pipeline systems the diameter is not fixed and the system can have (potentially unknown) pipe-elements like junctions, valves and sharp bends that prevent physical passage.

Pipe inspection gauges are often used in the oil-industry. Distribution oil pipelines often have long straight sections with a fixed diameter, non-sharp corners and barely blocking pipe-elements. Downtime can be kept low due to efficient in- and outlets. For this industry, these devices, although relatively expensive, are very adequate and cost efficient.

In other industries, however, this may be different. For example, a significant fraction of water distribution networks had ad-hoc repairs in the past that are only partially documented or not documented at all. Physical passage of any type of large in-pipe inspection devices cannot be guaranteed due to the lack of knowledge on the location and the type of pipes and pipe components, bends, flanges, and valves. The cost of failure of such pipelines is also significantly lower than that in the oil-industry, making the available solutions not cost-effective.

For environments of this type, smaller and cheaper solutions are required, such as small flow-driven inspection devices. The SmartBall [8], designed specifically for larger pipelines, is an example of a successful product that goes with the flow through pressurized pipelines to detect leaks. While traversing the pipeline, it registers sound and listens to the acoustical signature of hissing leaks or trapped air. The positions of the device from which these events are measured is estimated based on interpolation of the timing of the registered event, relative to the passage of the device by externally applied beacons that the device senses.

1.1.3.2   Underground reservoirs

For underground reservoirs the discovery and initial exploration is performed using techniques like radar, acoustical surveyal, and remote sensing (e.g. gravity, spectroscopy) [9]. Resulting findings are limited to macroscopic parameters like size, location, volume and possibly the type of resource. Test-drilling then serves to identify chemical composition of the resource and physical properties of surrounding ground layers. The depth and difficulty to

get close prevents gaining knowledge on the detailed internal structure.

Existing inspection techniques for obtaining further details about the internal structure are limited to devices placed inside the borehole, and not in the reservoir itself as the dimensions are too small and the reservoir is located too deep underground. Typical borehole inspection techniques can sense in the near vicinity around the borehole with a range of around one meter.

Using cross-borehole tomography, cross-sections of the reservoir between two adjacent boreholes can be measured [10]. However, the resolution of these measurements is limited to not less than around one meter.

1.1.3.3   Industrial tanks and reactors

Inspection of industrial tanks and reactors from the inside is often very difficult due to chemicals that are inside, the limited accessibility and the strict regulations for human entry in these systems. Using a technique called acoustic emission it is possible to listen to processes like e.g. corrosion and crack-forming in the enclosure of the environment.

Monitoring and studying the processes that happen inside industrial tanks and reactors is partially possible by installing systems like Doppler flow profile, liquid level sensor and temperature sensor. However, this does not provide information on e.g. the distribution of parameters of interest, and interactions happening inside.

1.2   Thesis objective

The above survey shows that the technology for inspecting enclosed environments has progressed significantly over the past decades, but still has strong limitations in terms of applicability and cost. Additionally, there is a clear need to obtain additional knowledge on the dynamics and state of the interior of enclosed environments, and to assess/characterize the detailed internal structure.

The aim of this thesis, therefore, is to contribute **to the development of cost effective and broadly applicable inspection tools for enclosed fluidic environments**.

1.3   The go-with-the-flow approach

The closest technique that can potentially be used to tackle the exploration needs of the three different enclosed fluidic environment uses devices that are small relative to the diameter of the environment and that are driven by the flow of the medium. In enclosed environments there seems to be only one existing example of this technique, the SmartBall, and this product is only meant for detecting leaks in pressurized pipelines. We can strongly extend the functionality so as to enhance flexibility and deliver inspection results in a wide variety of applications in a cost-effective manner. Specifically, instead of a single device we use a large quantity of devices, operating in a *swarm*. The individual devices are then called *sensor nodes*, or in short *nodes*.
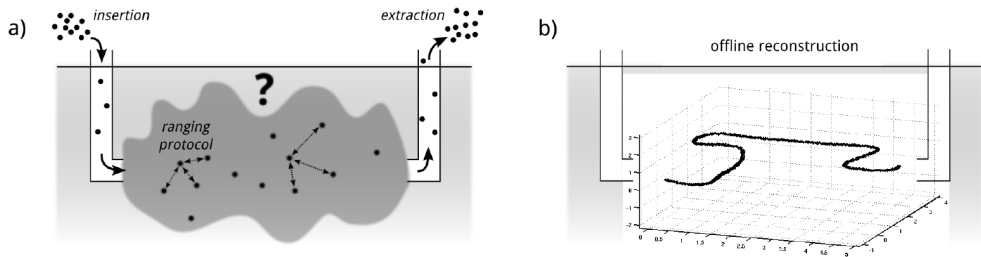
Figure 1.2: Go-with-the-flow approach: a) insertion of large quantities of sensor nodes; nodes go-with-the-flow of the medium and perform distance measurements using a ranging protocol and store their measurements in memory; upon extraction the data of the nodes is read out for further analysis; b) the measured distances between nodes is used to reconstruct the positions of the nodes relative to each other. The resulting shape of the swarm reveals structural information of the environment. The 3-D map can be to visualize additional sensor information.

A typical operational approach of sensor nodes operating in a swarm is visualized in Fig. 1.2 and involves the following steps:

1. insert sensor nodes into the environment;

2. nodes traverse the environment with the flow of the medium;

3. nodes perform and store measurements, not only of parameters of interest, but also for being to able to determine positions of the nodes;

4. nodes are extracted from the environment;

5. data from the nodes is analyzed offline to obtain as much information as possible on the enclosed environment.

The go-with-the-flow approach using swarms of small and autonomous sensor nodes has the potential to provide information about the flow dynamics and the state and distribution of the medium as long as the density of the nodes is matched with that of the medium and the nodes are small enough not to interfere with the dynamics. The nodes effectively serve as tracer particles; the movement and distribution of the nodes can be used to infer the dynamics of the flow of the medium.

This approach also has the potential to go where no sensor has gone before, for example in deep underground reservoirs where the sensor nodes can measure parameters of interest that help in discovering the internal structure and conditions.

For pipelines, the technique has the potential to provide low to medium resolution information and to do that for large sections of pipeline. It promises to allow for a quick and cheap inspection method that can assess integrity for prioritizing pipe replacement and to detect large faults like leakages and unwanted constrictions.

Using the flow of the medium to traverse the medium and to get the nodes out of the enclosed environment implies that the approach cannot be used in systems with stagnant flow or without a flow to a controllable exit point. For example, in fresh water distribution networks

in urban areas, the pipeline system is highly ramified and has various places of stagnant water and no easily accessible exit point; inserted nodes will get stuck here. However, in many instances entry and exit points will or can be made available so that the approach is, in principle, widely applicable.

Accordingly the specific research objective addressed in this thesis can be formulated as:

> "Can we design an efficient and effective 'go-with-the-flow' system for characterization of enclosed fluidic environments?"

1.4   Key constraints and choices

1.4.1   Key characteristics of the go-with-the-flow approach

In the go-with-the-flow approach all sensor nodes are equipped with sensors that continuously or intermittently assess the environmental conditions of interest. To be able to characterize the environment as a whole, it is essential that the acquired sensor information can be related to sensor position, so that after extraction of the sensors a parametric picture of the entire environment can be composed.

Without information on the positions only macroscopic parameters can be deduced. Examples of findings without position information can be found in [11]: e.g. "the narrowest passage is X since nodes with dimensions larger than it do not make it to the exit point", and "a minimum pump-rate of Y is needed for nodes to traverse the entire system, at lower pumping-rates the nodes do not get out".

Assessing the position of sensor nodes over time is far from trivial given the required small size (and hence small power dissipation) of sensor nodes, and given that classical localization techniques (e.g. beacons, GPS etc) are inapplicable due to the inaccessibility of the environment and the strong attenuation of acoustic and electromagnetic waves that could be used for localization.

Accordingly, localization is a key challenge and a major focus of this thesis. Note that the use of the word 'localization' is used to indicate the process or method of determining positions of nodes. The word 'reconstruction' is used to indicate the process of estimating the position of nodes based on measurements that were performed, i.e. to answer the question 'where was the node when the measurements were taken'.

We proceed to discuss the major constraints, design choices and research questions in more detail.

1.4.1.1   No direct communication across the boundaries of the environment

Communicating to and from enclosed environments that contain fluid media and are enclosed by (e.g. metal) containment or are located deeply underground is a hard challenge. Signals with larger wavelengths experience less absorption but are less efficiently produced and de-

tected by small nodes. In specific cases, communication might be possible, but to keep the approach/solution as general as possible we assume that no communication is possible to and from the environment. As a consequence, nodes need to operate autonomously as they cannot be controlled from the outside, and nodes need to store measured data on-board. Data can only be extracted from the nodes once they are extracted from the environment.

1.4.1.2   Size constraint on nodes

Nodes need to traverse the environment in order to make in-situ measurements. Since the measured data also cannot be sent to the outside world, nodes need to be extracted after traversing the environment in order to obtain this data. This means that they need to be smaller than the narrowest parts of the environment in order to be extracted. Furthermore, the weight of the node needs to be small enough to allow the nodes to move with the flow. In some applications the nodes even need to be neutrally buoyant in the medium, lowering the weight limit further.

1.4.1.3   Power limitation

Probably the most significant and overarching constraint of highly miniaturized sensor nodes is their power limitation. Due to volume and weight constraints of the nodes, battery capacity will be extremely limited and this influences all other design choices related to sensing, computation, and communication. In fact, in millimeter-sized nodes batteries may only provide enough energy to perform one set of measurements instead of continuously or intermittently during the passage through the environment.

1.4.2   Options for localization

Fig. 1.3 provides a survey of possible approaches for localization in the go-with-the-flow approach. Although position is commonly expressed in coordinates relative to an absolute and shared frame of reference, this might not be feasible for the go-with-the-flow approach. In the next subsections we discuss the options indicated in Fig. 1.3. Subsequently, in Sec. 1.4.3, we explain our choices and approach.

1.4.2.1   Coordinate-based localization

The most common way to localize is to use a coordinate-based approach. Positions are defined using coordinates relative to a frame of reference. Possible frames of reference include:

- **the Earth frame of reference**: It is a frame of reference that is fixed to the Earth and uniquely defines positions within it. A commonly known example of this is the World Geodetic System 1984 (WGS84), used in e.g. cartography and satellite navigation.

- **the (static) environmental enclosure**: Positions are defined relative to the enclosure of the environment. The environment itself can be used as a continuum of reference points.

- **neighbouring nodes**: positions can be defined relative to the positions of neighbouring nodes, independent of whether they are static or not.
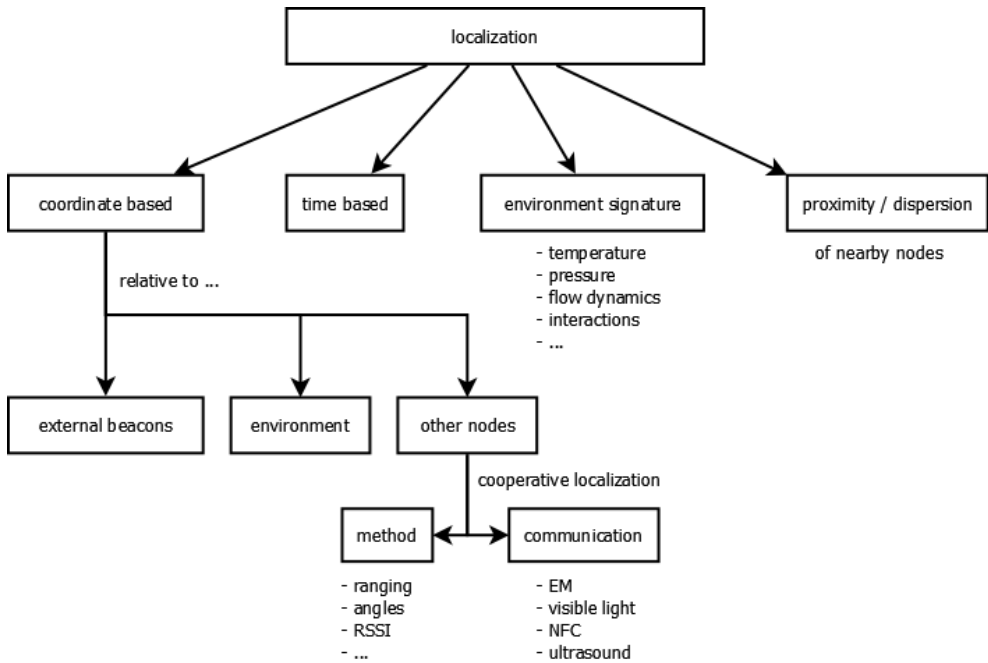
Figure 1.3: Options for obtaining information related to position.

Direct localization using beacons located outside of the environment that define positions in the **Earth frame of reference** is not feasible as no direct communication is possible from and to the environment.

Indirectly localizing nodes relative to the Earth-frame is possible only through referencing to the known entry and exit points of the enclosed environment. A variety of localization methods using entry and/or exit point have been proposed, studied or used. Examples are:

- using the enclosed environment and/or medium as wave-guide for communication channels for time-of-flight, or related, localization methods [12, 13];

- dead-reckoning: using inertial measurements to calculate incrementally the position of the nodes, starting and ending with the known positions of the entry and exit points [14];

- pervading the environment with nodes in which incrementally positions of nodes are calculated, starting close to entry/exit point and incrementally localizing neighboring nodes which can then be used as a reference point, thereby forming a chain through the environment [15, 16].

Using the **enclosure of the environment** as frame of reference requires a method to identify this enclosure. A commonly used approach is using cameras and/or LIDAR with a technique called Simultaneous Localization and Mapping (SLAM). With SLAM gradually a map of the enclosure is built and the positions of the nodes therein are estimated. This is performed by stitching together detailed images of the enclosure and determining from which
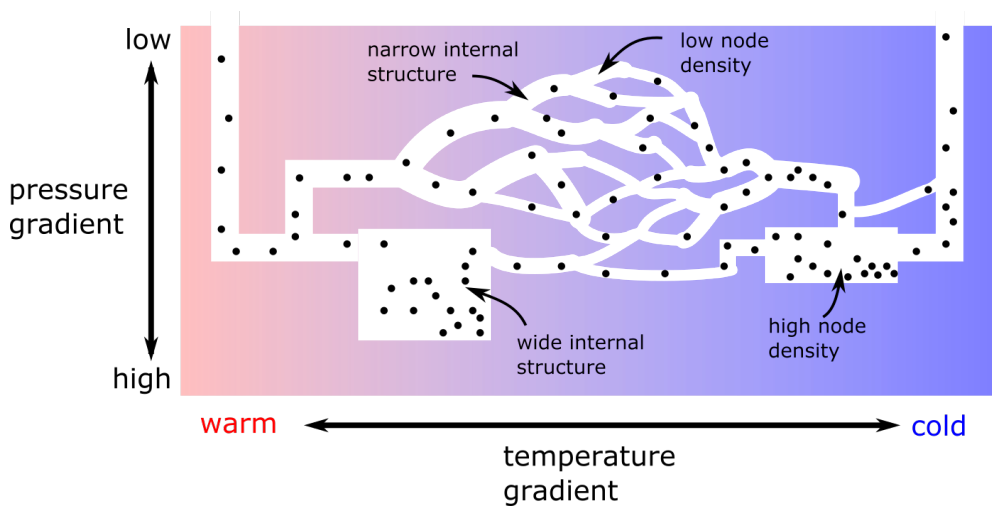
Figure 1.4: Obtaining geometric information in enclosed environments does not need to be with coordinates. Tracking parameters over time can distinguish locations, paths and internal structure. Examples: temperature gradient, pressure gradient, passage width, node density, etc.

position the image is taken to make the measurements fit [17, 18].

Using **neighbouring nodes** to define positions relative to them is known as *cooperative localization*. This can be done by SLAM-like approaches, but also by performing ranging measurements, i.e. by measuring the distances between nodes to form a point-cloud that represents the swarm of nodes and defines the positions of the nodes within the swarm [19].

1.4.2.2 Time-based localization

In time-based approaches, the travel time of the nodes through the environment is used as estimator for node positions. Measurements are stored together with the time-stamp of the measurements. After extraction, the recorded time can be used as a measure of position between entry and exit point.

When the flow of the medium is fairly constant, it can act as a coarse (initial) estimate of the position. When the environment is traversed multiple times, either with the same node multiple times, or with multiple nodes at the same time, a statistical analysis can be performed in an attempt to match the nodes individual recorded times and measurements with a coarse estimation of the position in the environment. Using a larger amount of nodes might yield a more accurate estimate of the position.

1.4.2.3 Environment signature-based localization

In Fig. 1.4 an enclosed environment is schematically visualized. It shows that environmental parameters like temperature, pressure and dimension of internal structure can vary throughout the environment. For example in geothermal-reservoirs and in CHOPS-reservoirs there is a

clear temperature gradient. In principle, this position-dependency allows for a (statistical) mapping of measured parameters to a position in the environment.

Combining knowledge of measured position-dependent parameters with time-based position estimates, as described in the previous section, has the potential of yielding useful information on the position of nodes and simultaneously on the parameters of interest of the enclosed environment.

1.4.2.4 Proximity-based localization

As also seen in Fig. 1.4, not only environmental parameters like temperature and pressure can be used to study positions, also node proximity is a parameter that can be used for this. Depending on the environment (size, diameter, branches, etc), the flow of the medium, and the nodes itself, the nodes move through the environment in a specific time-dependent trajectory. The quantity of nodes per section of the enclosed environment is variable through the environment and changes depending on the insertion rate of the nodes into the environment and the aforementioned aspects.

When the nodes are measuring the proximity of other nodes over time, a time-based dispersion analysis can potentially be performed offline to study environmental aspects like geometry, diameter and internal structure [20]. More specifically, when nodes also measure *which* nodes where nearby and at what time, a more thorough analysis can then be performed: branches and different paths through the environment can also be identified.

1.4.3 Localization approach adopted in this thesis

One of the largest challenges for developing sensor nodes capable of providing position information under the constraints outlined above is the limit in size the nodes can have. The smaller we can make the sensor nodes, the larger the amount of applications and environments in which they can be used. Reaching those levels of miniaturization requires significant hardware development efforts and costs. We choose for the strategy 'minimum effort online, maximum effort offline' in which we rigorously limit and simplify what nodes are doing while traversing the environment, and after extraction we compensate for this by a maximum of off-line effort where ample computation power is available.

To minimize online effort, we assume that nodes can only perform ranging measurements to nodes that are sufficiently close by, i.e. we consider a constrained form of cooperative localization.

Ranging can e.g. be done acoustically or with ultrasound. Since sound and ultrasound can efficiently traverse fluidic environments, acoustic or ultrasound ranging can be done at low power levels and (especially for ultrasound) in small nodes. Other measurement approaches (e.g. cameras, LIDAR) are not applicable to opaque and turbid fluidic environments, and moreover tend to be more bulky and power-hungry.

Our ranging-based cooperative localization approach is schematically depicted in Figure 1.5. The major part of this thesis is concerned with this approach (Chapters 2-5). In Chapter 6 we additionally explore the possibilities of time-based position determination in combination with using the environmental signature to simultaneously localize the nodes and
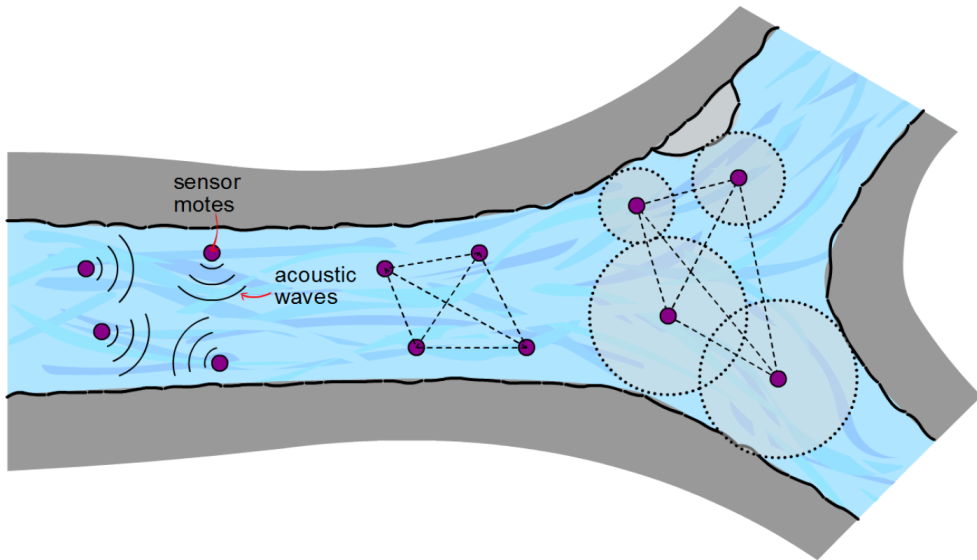
Figure 1.5: A swarm of highly miniaturized sensor nodes is injected into an enclosed fluidic environment and flows through it, along with the medium. Meanwhile, sensor nodes monitor their environment, not only for parameters of interest but also to reconstruct their position relative to other nodes without relying on external communication.

map the environment.

1.4.3.1   Severely limited communication within the environment

Communication within the enclosed environment is also limited by absorption by the medium and by the boundaries of the enclosure.

Depending on the typical dimensions of the environment, the size of the nodes and their spatial distribution, a required communication range of 1 meter appears reasonable for most of our applications. Using electromagnetic (EM) communication between nodes is not attractive, as (sub-)centimeter node dimensions prohibit large antennas. Moreover, at high frequencies suitable to be used with small antenna's, the liquid (and possible highly saline) media have large attenuation. This effectively prevents EM communication for distances larger than, say, a centimeter. Using inductive communication such as near-field communication (NFC) is limited to the centimeter range in saline environments [22]. The (liquid) media is also not guaranteed to be clear enough for the use of (visible) light communication.

In our approach we will study the use of ultrasound signals as means of determining distance to neighbouring nodes. Even though this seems feasible, there are significant challenges in using ultrasound as a communication medium within the environments.

1.4.3.2   Protocols to obtain distance measurements

In order to determine distances between sensor nodes, a *ranging protocol* is required that stipulates how signals need to be transmitted and exchanged between nodes in order to estimate their mutual distance. A variety of ranging protocols is available in other fields of study [23, 24]. However, in our go-with-the-flow applications with the strict constraints outlined above, there is none readily applicable for us.

1.4.3.3   Reconstruction of node positions

Once distances are determined by the nodes while traversing the environment and the data is extracted from the nodes, the positions of the nodes relative to each other need to be reconstructed. Reconstruction algorithms from other application domains can be used for the range-based measurements our nodes are performing. However, in our approach to minimize online effort we are dealing with potentially severely corrupted or altogether missing distance measurements. A maximum offline effort means that the reconstruction algorithm needs to be robust against these corrupted measurements. The more robust we can make the reconstruction algorithm, the more corrupted we can allow our online measurements to be.

1.4.4   Research questions

Based on the above considerations, we devote specific attention to the following 3 specific research questions in this thesis.

1. Can we develop a robust algorithm for the reconstruction of node positions using distance-based measurements that are obtained in the go-with-the-flow approach?

2. Can we develop a ranging protocol for nodes to determine distances while traversing the environment, such that nodes need to perform as little effort as possible during the go-with-the-flow phase?

3. Can we apply other methods to perform localization?

We proceed to elaborate how these research questions are addressed in the various chapters of this thesis.

1.5   Structure of the thesis

Our main focus in this thesis is to obtain reliable position coordinates of the nodes relative to each other while traversing the enclosed environment. Based on the design philosophy 'minimum effort online, maximum effort offline' we investigate what are the possibilities to reconstruct positions in offline analysis in order to determine what is needed as minimum requirements for hardware specification and measurement accuracy.

Therefore, in Chapter 2 we start off developing a robust reconstruction algorithm that has to deal with severely corrupted ranging measurements. We use extensive simulations with increasingly more limitations on the node's capabilities to study their effect on the reconstruction performance. We then see down to what specifications the reconstruction algorithm still yields useful results. The limits and trade-offs can then be used to fix the minimum

hardware design specifications. These simulations show, for example, that a large degree of connectivity – i.e. a large number of nodes within communication range of each other – is highly favoured over unique identifiability of the nodes.

In Chapter 3 we perform a more detailed analysis of trade-offs in node's capabilities and what effect they have on the final reconstruction result. Specifically, we explore the trade-off between connectivity and identifiability. Our analysis demonstrates, for example, that there is a minimum required degree of connectivity for each specific measurement noise level for the reconstruction to be successful.

(Chapter 4) In order to measure the distances between nodes, a ranging protocol is required to control the sequence of events needed for determining distances. Accounting for the limitations and constraints determined in the reconstruction process, a ranging protocol is developed for swarms of highly resource-limited nodes in extremely time-variant communication channels. A simulation chain is set up to simulate the ranging protocol for a number of realistic scenarios where the movement of the nodes is simulated in a variety of environments. The output of the simulated ranging measurements is then used as input for the reconstruction algorithm that estimates the positions of the majority of the nodes. In the considered scenarios, the protocol permits reliable reconstruction of the node positions. The results of this study can serve as a basis for defining specifications for real-world functional sensor node prototypes.

(Chapter 5) The possible trade-offs among parameters of the ranging protocol is further studied for one particular use case. The parameters that can be adjusted are e.g. the signal strength (communication range) and the number of identifying bits per message. Insight is gained on the effect that these parameters have on e.g. the ranging latency, signal collisions, the fraction of distances that are successfully measured and the total number of messages transmitted to achieve this. Combining this knowledge with the effects that these parameters then have on the reconstruction of the node positions, allows one to properly adjust parameters depending on the different environments and application scenarios.

(Chapter 6) To investigate what other pieces of positional information can be gathered to aid in localizing the nodes, sensor nodes with a diameter of 4 cm and equipped with an inertial measurement unit (IMU) are constructed and used in a field test, floating through an underground pressurized water pipeline. The nodes log the sensed acceleration, rotation and the local magnetic field. The IMU data is used – unlike dead-reckoning techniques – to reveal structural information of the environment by studying the interaction between the node and the environment, and to obtain positional information based on the measured magnetic signature. Pipe components like junctions and hydrants are discovered and types of material of which the pipe is made are identified. Positional information is obtained by comparing the data with an incomplete map of the pipeline that was given beforehand.

1.6   Scientific contributions / Table of contents

The peer-reviewed publications incorporated in this thesis are listed in Table 1.2.

Ch.	Short description	Orginal title	Publication
1	Introduction		
2	Reconstruction of nodes using highly corrupted distance measurements; stress-testing the limits	Robust Reconstruction of Sensor Swarms Floating through Enclosed Environments	Wireless Sensor Networks (journal, SCIRP, 2018)
3	Reconstruction of nodes; looking deeper into specific trade-offs	Mapping swarms of resource-limited sensor motes: solely using distance measurements and non-unique identifiers	Symposium Series on Computational Intelligence (conference, IEEE, 2016)
4	Simulation chain of flow-model, ranging protocol and reconstruction algorithm	Go-with-the-Flow Swarm Sensing in Inaccessible Viscous Media	Accepted with minor revisions for Sensors (journal, IEEE, 2019)
5	Development of ranging protocol to obtain distance measurements in our application	Asymmetric multi-way ranging for resource-limited nodes	Ad Hoc Networks (conference, Springer, 2016)
6	Experiment with IMU-nodes in pipeline and analyzing data for obtaining pipeline characteristics	Sensor Motes for the Exploration and Monitoring of Operational Pipelines	Transactions on Instrumentation and Measurement (journal, IEEE, 2018)
7	Conclusion		

Table 1.2: Chapters of thesis

1.6.1   Other peer-reviewed publications

- E. Duisterwinkel, L. Demi, G. Dubbelman, E. Talnishnikh, H. Wörtche, and J. Bergmans. "Environment mapping and localization with an uncontrolled swarm of ultrasound sensor motes", Proceedings of Meetings on Acoustics, 20(1), 2014.

- Gijs Dubbelman, Erik Duisterwinkel, Libertario Demi, Elena Talnishnikh, Heinrich J. Wörtche, Jan W.M Bergmans, "Robust sensor cloud localization from range measurements", 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Chicago, USA, 14-18 Sept. 2014

1.6.2   Patents

- Patent (US 9397732, CA 2872086, NL 2011892), "Motes for environment mapping". Inventors: Elena Talnishnikh, Hendrik Albert Duisterwinkel, Heinrich Johannes Wörtche, Johannes Hubertus Gerardus Van Pol

References

[1] Henk de Kater, *et al.*, "Inspectietechnieken voor rationeel saneringsbeleid van leidingnetten", KWR Research Insititute, 2010.

[2] Z. Liu and Y. Kleiner, "State of the art review of inspection technologies for condition assessment of water pipes," Measurement, vol. 46, no. 1, pp. 1–15, Jan. 2013.

[3] H. J. Wörtche, "Advanced analysis: Motes dynamics in a mixing tank," INGU Solutions, Calgary, AB, Canada, Tech. Rep., 2014.

[4] J. M. M. Tur and W. Garthwaite, "Robotic devices for water main inpipe inspection: A survey," J. Field Robot., vol. 27, no. 4, pp. 491–508, 2010.

[5] T. Hao et al., "Condition assessment of the buried utility service infrastructure," Tunneling Underground Space Technol., vol. 28, pp. 331–344, Mar. 2012.

[6] O. Hunaidi, P. Giamou, Ground-penetrating radar for detection of leaks in buried plastic water distribution pipes, in: Proceedings of 7th International Conference on Ground-Penetrating Radar, Lawrence, Kansas, USA, 1998.

[7] C.U. Grosse, M. Ohtsu, "Acoustic emission testing", Springer-Verlag Berlin Heidelberg, 2008

[8] R. Fletcher, Smartball, "a new approach in pipeline leak detection," in: International Pipeline Conference, Calgary, Alberta, Canada, 2008.

[9] E.A. Cloutis, "Spectral Reflectance Properties of Hydrocarbons: Remote-Sensing Implications," Science, Vol. 245, Issue 4914, pp. 165-168, 14 Jul 1989

[10] W. Menke, "The resolving power of cross-borehole tomography", Geophysical Research Letters, Volume11, Issue2, Pages 105-108, February 1984

[11] E. Talnishnikh et al., "Micro Motes: A Highly Penetrating Probe for Inaccessible Environments", in Intelligent Environmental Sensing, edt. H. Leung, Springer International Publishing, 2015, pp 33-49.

[12] G Kokossalakis, "Acoustic data communication system for in-pipe wireless sensor networks", Ph.D. thesis MIT, 2006

[13] David C. Swanson, Nicholas C. Nicholas, David Rigsby, "System and method for detecting, localizing, or classifying a disturbance using a waveguide sensor system", US patent, US6731210B2

[14] R. Harle, "A survey of indoor inertial positioning systems for pedestrians", IEEE Communications Surveys & Tutorials, 2013

[15] S. Schlupkothen, G. Dartmann and G. Ascheid, "Exploration With Massive Sensors Swarms", in IEEE GlobalSIP Symposium on Controlled Sensing For Inference: Applications, Theory and Algorithms, Austin, Texas, December 2013.

[16] E. Duisterwinkel, L. Demi, G. Dubbelman, E. Talnishnikh, H. Wörtche, and J. Bergmans, "Environment mapping and localization with an uncontrolled swarm of ultrasound sensor motes". Proceedings of Meetings on Acoustics, 20(1), 2014.

[17] J.-l. Blanco, J.-a. Fernández-Madrigal, and J. González. Efficient Probabilistic Range-Only SLAM. In IEEE/RSJ Conference on Itelligent Robots and Systems, pages 22–26, Nice, France, 2008.

[18] J. Djugash, S. Singh, G. Kantor, and W. Zhang. Range-only SLAM for robots operating cooperatively with sensor networks. In IEEE International Conference on Robotics and Automation, pages 2078–2084, Orlando, Florida, USA, 2006.

[19] H. Wymeersch et al., "Cooperative Localization in Wireless Networks", in Proc. IEEE, 97(2):427450, Feb. 2009.

[20] Liu, K., Wang, S., Zhang, F. "Efficient localized localization algorithm for wireless sensor networks". In Proc. 5th international conference on computer and information technology (pp.21–23), 2005

[21] H. J. Wörtche, "Advanced analysis: Motes dynamics in a mixing tank," INGU Solutions, Calgary, AB, Canada, Tech. Rep., 2014.

[22] Alessandro Pozzebon, "Bringing near field communication under water: short range data exchange in fresh and salt water", 2015 International EURASIP Workshop on RFID Technology (EURFID), IEEE, 2015

[23] G. Han, H. Xu, T.Q. Duong, J. Jiang and T. Hara, "Localization algorithms of Wireless Sensor Networks: a survey", Telecommunication Systems 52(4), 2419-2436 (2011).

[24] M. Erol-Kantarci, H.T. Mouftah and S. Oktug, "Localization Techniques for Underwater Acoustic Sensor Networks", IEEE Communications Magazine 48(12), 152-158 (2010).

[25] E. Menegatti et al., "Range-only SLAM with a Mobile Robot and a Wireless Sensor Network", in IEEE Int. Conf. Robotics and Automation, p.8-14, Kobe, Japan, May 2009.

[26] Akyildiz, I.F., Pompili, D., Melodia, T.: Underwater acoustic sensor networks: research challenges. Ad Hoc Netw. 3, 257–279 (2005)

[27] Patwari, N., et al.: Cooperative localization in wireless sensor networks. IEEE Sig. Process. Mag. 22(4), 54–69 (2005)

[28] Thomas Szabo, "Diagnostic Ultrasound Imaging: Inside Out", second edition, Aca-

demic Press, December 2013

[29] Li, H., Jung, K.W., Deng, Z.D.: Piezoelectric transducer design for a miniaturized injectable acoustic transmitter. Smart Mater. Struct. 24 (2015)

[30] J. Lu, et al, "A small long-life acoustic transmitter for studying the behavior of aquatic animals", Review of Scientific Instruments 87, 114902 (2016); doi: 10.1063/1.4967941

[31] Jules S. Jaffe et al., "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics", NAT COMMUN 8, Article number: 14189 (2017)

[32] J.L.A. Franco, *et al.*, "Sonic investigation in and around the borehole", JLA Franco, Oilfield Review, Spring 2006

2

Robust Reconstruction of Sensor Swarms Floating Through Enclosed Environments

authors:
Erik H.A. Duisterwinkel, Gijs Dubbelman, Libertario Demi, Elena Talnishnikh, Heinrich J. Wörtche, Jan W.M. Bergmans

A novel type of application for the exploration of enclosed or otherwise difficult to access environments requires large quantities of miniaturized sensor nodes to perform measurements while they traverse the environment in a 'go-with-the-flow' approach. Examples of these are the exploration of underground cavities and the inspection of industrial pipelines or mixing tanks, all of which have in common that the environments are difficult to access and do not allow position determination using e.g. GPS or similar techniques. The sensor nodes need to be scaled down towards the millimetre range in order to physically fit through the narrowest of parts in the environments and should measure distances between each other in order to enable the reconstruction of their positions relative to each other in offline analysis. Reaching those level of miniaturization and enabling reconstruction functionality requires: 1) novel reconstruction algorithms that can deal with the specific measurement limitations and imperfections of millimetre-sized nodes, and 2) improved understanding of the relation between the highly constraint hardware design space of the sensor nodes and the reconstruction algorithms. To this end, this work provides a novel and highly robust sensor swarm reconstruction algorithm and studies the effect of hardware design trade-offs on its performance. Our findings based on extensive simulations, which push the reconstruction algorithm to its

breaking point, provide important guidelines for the future development of millimetre-sized sensor nodes.

2.1 Introduction

Automatic and detailed reconstruction of environments and objects from sensory data is one of the success stories of modern signal processing. Still there are environments for which detailed reconstruction is not possible with current remote or *in situ* sensing technologies. In this work, we focus on such environments and examples are: the interior of deep underground formations like (oil-)reservoirs, mines and geothermal sources, and industrial infrastructure like pipelines, mixing tanks and reactors. These environments are filled with a (semi-)liquid medium and typically are globally large but locally small (e.g. in diameter). Furthermore, they are difficult-to-access for *in situ* measurements and unsuitable for remote sensing methods. In this paper we consider a 'go-with-the-flow' approach that follows a sensor swarm paradigm, in which large quantities (e.g. hundreds or thousands) of highly miniaturized and redundant sensor nodes are gradually inserted into the environment [1,2]. Carried by the flow of the medium, the swarm fills the environment and by reconstructing the node positions, and therewith the shape of the sensor swarm, the geometric structure and size of the environment can be inferred. Reconstruction of the sensor swarm is performed after motes are extracted from the environment and their data read-out. This process is illustrated in Fig. 2.1.

The individual sensor node positions relative to each other can be reconstructed from sparse inter-node distance measurements [3, 4]. These distance measurements can be obtained using a ranging protocol as in e.g. [5]. As illustrated in Fig. 2.2, reconstructing the positions of the nodes can be seen as a specific kind of graph problem, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertices $\mathcal{V}$ are the sensor node positions and the edges $\mathcal{E}$ are the distances between them. Typically such a problem can be solved using non-linear optimization algorithms [28], however, after obtaining the measured distances (i.e. the edges $\mathcal{E}$ in graph $\mathcal{G}$), the challenge is to obtain a robust initial estimate of the node positions in order for non-linear graph optimization algorithms to succeed. In order to robustly obtain such an initial estimate, our offline[1] reconstruction algorithm uses a Random Sampling Consensus (RANSAC) method applied to general lateration techniques [6, 7]. To the best of our knowledge, the novel reconstruction algorithm presented in this work, is one of the most robust algorithms that is able to perform 3-D reconstruction of large sensor swarms. Our applications require such a robust reconstruction algorithm since the distance measurements are performed under very adverse sensing conditions, as explained in Sec. 2.2. The details of our novel method are provided in Sec. 2.3 and Sec. 2.4.

We use this reconstruction algorithm to study trade-offs in the hardware design space, as many hardware challenges still need to be addressed before large swarms of millimetre-sized functional sensor nodes can be developed and deployed effectively. Trade-offs must be made in the hardware design with respect to sensing capabilities and size. Improved capabilities, e.g. larger communication radius or better signal-to-noise ratios, will improve reconstruction performance, but make cost-effective miniaturization of nodes more challenging. Therefore, in Sec. 2.5, we use extensive simulations with increasingly more limitations on the hardware capabilities to study their effect on the reconstruction performance. These limitations include increased noise in the measurements and not being able to uniquely identify other nodes.

---

[1]In this paper we use the term *online* to indicate the period where motes are in the process of performing measurements and while going through the environment, and the term *offline* for after the motes are retrieved and data is extracted from the motes.
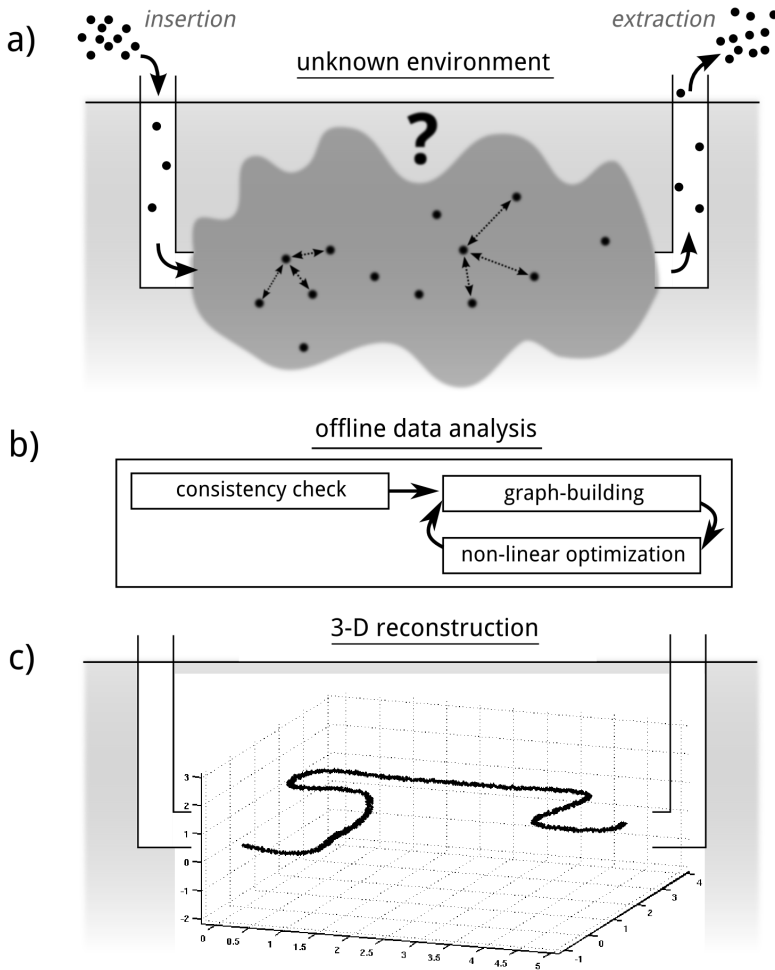
Figure 2.1: A swarm of sensor nodes is inserted into the environment of interest; the sensor nodes traverse the environment using the flow of the medium. Once the sensor nodes are distributed in the environment, distance measurements between neighbouring sensor nodes are performed and stored locally within the nodes. Finally, the sensor nodes and their data are retrieved. All data is analysed offline to obtain the 3-D reconstruction of the sensor swarm from the inter-node distance measurements.

The assessment of the performance of the reconstruction depends on which information is favoured from the swarm and different performance metrics are introduced.

The simulations in this work do not guarantee that robust reconstruction will work in reality, but as also discussed in Sec. 2.6, they do show that advances in offline reconstruction algorithms, can, to a large extent, compensate for severely limited hardware capabilities. Understanding the trade-offs between hardware design and the reconstruction performance can guide the hardware design of yet-to-be-developed highly miniaturized sensor nodes of

future applications [2].

Several applications and research studies already exist that use a sensor swarm approaches where sensor nodes move with a flow through specific environments. In [8], 13cm-sized sensor nodes were developed and deployed in the ocean to flow with the ocean currents and allows to study e.g. plankton behaviour. These nodes can actively change their density to adjust their depth while collecting a variety of sensor data. The nodes receive ultrasound ping signals from nearby buoys (at kilometre distance) such that afterwards their positions can be reconstructed using general lateration. The researchers in [9] developed millimetre-sized sensor nodes that can be injected into living fish to study how e.g. dams and ocean energy devices affect their behaviour. These sensor nodes measure temperature and emit a unique identifier using ultrasound such that they can be remotely tracked by an array of external hydrophones on the shoreline (within hundreds of metres distance). Both of the mentioned studies are characterized by the fact that external localization hardware (buoys or hydrophones with a known location) are used to reconstruct the node positions.

In our earlier work, targeting enclosed and difficult to access environments, we developed 4cm-sized sensor nodes that measure the node's inertial data (i.e. linear acceleration and rotation) and measure the local magnetic field. These nodes with variable density have been used e.g. to explore and inspect an underground pipeline of 15 centimetre in diameter and 260 metre in length [10] and to study the dynamics of multi-phase fluids in a mixing tank of several cubic metres [11]. Another of our attempts to work towards the mentioned future applications is described in [1]. It describes a field test that uses dummy (non-functional) sensor nodes to explore a deep underground sandstone oil-reservoir. This field test revealed that in order for the sensor nodes to pass through the 300 metre sandstone oil-reservoir, the size of sensor nodes cannot exceed 9 mm as the local dimensions are not larger than that.

The mentioned studies show that a 'go-with-the-flow' sensor approach can yield valuable insights into yet undiscovered environments and dynamics. They also show that some applications have clear constraints on e.g. the allowed maximum size of the used sensor nodes in order to pass through the environment. The objective of this paper is to see how far we can limit the capabilities of sensor nodes and how far they can be scaled down while still getting good reconstruction performance. In the following section (Sec. 2.2) we discuss the implications of scaling down sensor nodes to the millimetre size and operating them in enclosed and difficult to access environments.

Even though the used constraints in this paper represent a very challenging application scenario and might seem excessive in light of current applications, we believe they are valid constraints for future miniaturized nodes [2]. We think that providing a solution to the swarm reconstruction problem under these adverse sensing conditions can also be beneficial for current applications and problems. To summarize our contributions: 1) a novel and robust reconstruction algorithms is developed that deals with severe sensing conditions and imperfections, even including non-unique identification of distance measurements; 2) sensor node system design principles are derived from extensive simulations by sweeping input parameters until the reconstruction algorithm fails to reconstruct the sensor swarm; 3) different methods on how to assess the performance of the swarm reconstruction are provided, depending on which information is favoured from the swarm.
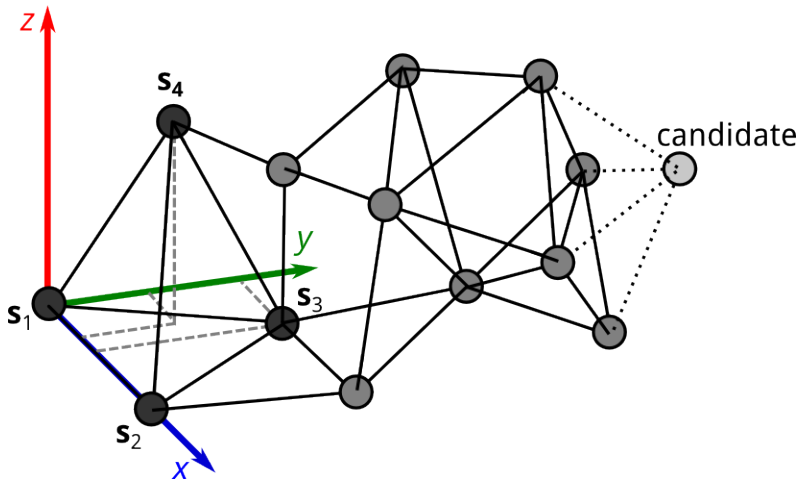
Figure 2.2: Reconstructing nodes within a swarm can be seen as a graph problem, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with the node positions as vertices $\mathcal{V}$ (circles) and the distances between nodes the edges $\mathcal{E}$ (lines). The initial four node positions $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4\}$ are chosen to define the coordinate system to solve the global reflection and rotation ambiguity. Using general lateration techniques, additional nodes, like the candidate node in the figure, can be added when distances to at least four already reconstructed nodes are known. For clarity of the figure, not all (required) edges are drawn.

2.2   The implications of application constraints

This section discusses existing work and shows that the applications which we target pose specific challenges that have not been addressed earlier in a holistic manner. Reconstructing the position of nodes in (wireless) sensor and robot networks has been subject of study for an extensive period of time [3, 4]. Obtaining simultaneously the location of nodes as well as obtaining the map of the yet-unknown environment, is often referred to as Simultaneous Localization and Mapping (SLAM) [12]. Our research also falls under this denominator but has to deal with several distinctive constraints relative to other work. These constraints will be discussed in this section and can be summarized as:

1. no data is communicated between nodes (other than needed for distance measurements), measurements are stored in memory for offline analysis;

2. sparse connectivity in a large swarm;

3. not depending on external communication to e.g. fixed beacons;

4. using solely distance measurements, i.e. no direction information or additional sensor information, like bearing, odometry or inertia;

5. nodes are not identified uniquely.

Constraints 1, 3, and 4 are *hard* constraints and depending on the choice of applications, these either exist or not and make the reconstruction problem harder. Constraint 2 and 5 are

soft constraints and can vary in severity, depending on e.g. parameter choices that can be tweaked. For example, considering constraint 2: the connectivity between nodes depends on the transmission power, the distance between nodes and signal-to-noise ratio at the receiving side. Considering constraint 5: the identifiability of nodes depends on the number of identifiers chosen relative to the number of nodes used. Increasing the severity of these soft constraints makes the reconstruction problem gradually harder. Therefore, in our study, the parameters influencing connectivity and identifiability are swept over a large enough range in order to study the consequences of this on the reconstruction performance and therewith also to find the limits of current reconstruction algorithms.

Although each separate constraint has been considered earlier; this work is, to the best of our knowledge, the first to research reconstruction of sensor swarms when considering all mentioned constraints together and under realistic conditions.

2.2.1   Power consumption

Probably the most significant and overarching constraint of highly miniaturized sensor nodes is their power limitation. Due to volume and weight constraints in the small sensor nodes, battery capacity will be extremely limited and this influences all other design choices related to sensing, computation, and communication. In fact, we envision levels of miniaturization in which nodes have just enough energy to perform only a single inter-node distance measurement.

Furthermore, communication requires relatively large amounts of energy and is therefore only used to realize the inter-node distance measurements, but for example not for online distributed processing of the measurements (hence constraint 1). Instead of distributing the measured data or sending it to a central *sink node*, we use an approach in which the measured data is stored locally in the sensor nodes, and afterwards retrieved and processed centrally when power consumption is no longer a limitation.

Communication transmission power and the signal to noise ratio of receiving electronics are limited, reducing the communication range between nodes. Therefore, the inter-node connectivity between sensor nodes, by means of inter-node distance measurements, will be sparse (hence constraint 2) and dependent on the energy budget.

To the best of our knowledge, there is no other research that aims to perform swarm reconstruction under such extreme low-power conditions.

2.2.2   Beacons and landmarks

A distinction that can be made in SLAM related research is whether or not *beacons* are being used, that is, fixed points in space that serve as reference points where individual nodes can relate to. In our applications, we can not rely on the presence of beacons, as opposed to e.g. [13,14] and the earlier mentioned work in [8,9]. In our applications the environments are generally difficult-to-access and block long-range communication signals (hence constraint 3).

In applications with a static environment, the environment could potentially be used as a continuum of beacons, or in this case better described as *landmarks*. Opposed to the research

mentioned in e.g. [15] where high-end cameras and/or laser range-finders are used, measuring clear fixed landmarks is infeasible as such detection devices have a high power consumption and need continuous measuring; it is something that doesn't scale well to miniaturized nodes. Furthermore, detection of the environment using such sensors as e.g. in [15,16] would require clear line-of-sight in a transparent medium, which cannot be guaranteed in our applications.

Efforts to map node positions and the environments based on acoustic reflections can be found e.g. in [17, 18]. This, however, requires constraints on e.g. the node positions or the environment shape (e.g. straight walls), both of which cannot be guaranteed in our applications.

Given the aforementioned, the nodes are forced to cooperate with each other in order to obtain information on their positions, this also called cooperative localization [3].

2.2.3 Distance measurements

A common method for cooperative localization is to reconstruct node positions based on distance measurements between the nodes. Often such methods are amended with e.g. inertial, bearing, or odometric information from the nodes, like in e.g. [19, 20]. In our research however, we explicitly rely only on inter-node distance measurements, as adding such devices or sensors would add significantly to the size and energy budget (hence constraint 4). For example, an inertial sensor would need to be active continuously to track the absolute position and orientation. Furthermore, current miniaturized inertial sensors are not accurate enough for absolute position determination for longer than e.g. one second. Alternatively, the absolute orientation of the node could be estimated using e.g. small and low-budget sensors that measure the Earth magnetic field and the Earth gravitational field. But unfortunately, in our applications, no guarantees can be given on the external fields as the unknown environment can block or distort these.

Although one can argue that, in the future, advances in sensors will lift some of the mentioned limitations, we show that with only inter-node distance measurements, the reconstruction of the sensor swarm is possible.

2.2.4 Communication and identification

Motes with (sub-)centimetre dimensions only fit antennas that effectively produce EM radiation at frequencies that have a large absorption in the liquid media. The (liquid) media can have a high salinity and are also not guaranteed to be clear enough for the use of (visible) light communication. This effectively prevents EM communication for distances larger than e.g. a centimetre.

Ultrasound transducers at these scales do provide larger communication ranges within the boundaries of the environments, but yield other challenges for stable and fast communication between dynamic nodes in enclosed environments [21]. The information that can be communicated in a unit of time is therefore very limited, typically in the order of kilobytes per second or less.

Furthermore, the relatively slow speed of sound, in combination with the relatively small bandwidth, reflective environments and the fact that nodes are non-static, leads to highly time-

varying communication channels [5,22]. Distance measurements based on round-trip time of flight need to be completed within just a tiny fraction of a second as otherwise the nodes may have moved significantly relative to each other, invalidating the measured distance.

The low data rate and the need to perform distance measurements in a short period of time creates the incentive to reduce the content of the messages used for distance measurements (i.e. reducing the number of bits in the message). For example, in the ranging protocol as studied in [5, 22], at a data rate of 40 kbit/sec, the transmission of every additional bit for the identifiers to indicate the sender and receiver of the message adds 3 percent to the measurement noise caused by the movement of the nodes. This effect becomes larger when considering lower data rates. Furthermore, increasing the message length increases the probability of overlapping messages due to (time-varying) multi-path, thereby making it harder to correctly decode the messages.

In applications where many sensor nodes are used, ideally each node has a unique identifier by which it can unambiguously communicate with other nodes. However, for the envisioned application cases mentioned in this paper, it is worthwhile to explore the possibility of reducing the number of identifying bits as they make up the majority of the ultrasound messages (hence constraint 5). Reducing the number of identifying bits, and thus only having non-unique identification makes communication ambiguous. The inter-node distance measurements then also become ambiguous, but can be resolved using the robust reconstruction algorithm presented in this paper. Non-unique identification of distance measurements makes most existing reconstruction efforts not applicable as they depend on well-defined, non-ambiguous identifications [23–25]. Earlier efforts to first resolve these ambiguities before the reconstruction of node positions can be found in [26], but does not include actual reconstruction results.

Our reconstruction algorithm is, to the best our knowledge, the first that can successfully cope with identification ambiguities and with all the other mentioned hardware constraints.

2.3   Problem description

In this section we discuss the specific problem formulation and introduce our notation. All symbols used in this paper are summarized in Table 2.1. In this paper, whenever we use the bar-indicator, e.g. $\bar{d}$, it is to stress the true value of the parameter, to differentiate it from an estimated or measured value, e.g. d.

As mentioned in the introduction, the reconstruction of the node positions can be seen as a graph-optimization problem where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, in which the vertices $\mathcal{V}$ are the sensor node positions and the edges $\mathcal{E}$ are the distances between them. This is also visualized in Fig. 2.2. However, the edges of this graph can only be based on the *measured distances* that are noisy representations of the *actual distances*. In order to estimate an accurate graph of the nodes' positions we attempt to reduce the least squares error:

$$\mathrm{argmin}_{\mathcal{V} \in \mathbb{R}^3} \sum_{\mathcal{E}} \left(\|\mathcal{V}_i - \mathcal{V}_j\| - \mathcal{E}_{i,j} \right)^2 \tag{2.1}$$

Solving these kind of problems is often performed using non-linear optimization methods

Table 2.1: Legend of symbols used throughout the paper.

symbol	description
N	Number of nodes in swarm
n_{CID}	Number of communication identifiers in use
n_{comm}	Number of nodes within communication radius
r_{comm}	Communication radius of nodes
$\mathcal{G}(\mathcal{V}, \mathcal{E})$	Graph with reconstructed nodes indicating positions as vertex $\mathcal{V}$ and distance between nodes as edges $\mathcal{E}$
i, j	Unique (hardware) identification number (UID) of nodes i and j, other letters are also used to indicate other nodes. An additional bar in the indication stresses these are the true values (opposed to measured or hypothesised)
$\mathrm{CID}(i)$	Communication identification number (CID) used by node i
$\mathrm{d}_{i,j}$	Distance between node i and node j. Additional bars in the indication stress that these are true values (opposed to measured of hypothesised).
$\mathbf{s}_i$	Position of node i
$\mathcal{C}$	Consistency function used for associating measured distances from two nodes into a mutual connection
ϵ_r	Threshold value for maximum allowed distance deviation in consistency condition $\mathcal{C}$
$h_{i,j}$	Set of hypothesised distances between node i and j
$\mathbb{A}_I$	Number of additional identity ambiguities
$\mathbb{A}_D$	Number of additional distance ambiguities
$a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}$	Individual nodes (small letter) in a set of nodes (capital) selected based on some criteria in reconstruction algorithm
$\mathcal{Q}, \mathcal{Q}(b)$	Cliques in already reconstructed graph
$\mathbb{X}, \mathbb{Y}, \mathbb{Z}$	Recall percentages of nodes that have been reconstructed; $\mathbb{Y}$ and $\mathbb{Z}$ have an additional condition on the accuracy of the reconstruction
$\mathbb{E}_{\mathrm{abs}}, \mathbb{E}_{\mathrm{glob}}, \mathbb{E}_{\mathrm{loc}}$	Error metrics of a reconstruction indicating the mean absolute squared error of the reconstructed positions
$\mathbb{E}_{\mathrm{rel}}$	Error metric of a reconstruction indicating the relative error of the reconstructed distances
ν	Different noise parameters of distance measurements, expressed in percentages. Index indicates type of noise.

like e.g. a Levenberg-Marquardt algorithm or a Gauss-Newton algorithm. As with most non-linear optimization problems the key is to provide a proper initial estimate, in this case an initial estimate of the node positions. This initial estimate is then further refined using non-linear optimization. Obtaining the initial estimate for the node positions is performed by our novel robust reconstruction algorithm that is discussed in Sec. 2.4.

This section will first discuss the method on how to measure distances between the nodes (i.e. the edges $\mathcal{E}$) and the difficulties herein. In Sec. 2.3.1 we discuss how the distance measurements are performed. Sec. 2.3.2 discusses the different types of measurement noise to be expected in these measurements and the issues related to identification are discussed in Sec. 2.3.3. The consequences of these on the reconstruction algorithm are discussed in Sec. 2.3.4.

2.3.1   Distance measurement protocol

Following the operational procedure as depicted in Fig. 2.1, once the nodes are well distributed within the environment of interest, the nodes perform a single distance measurement to neighbouring nodes. This distribution can be achieved by e.g. gradually inserting new nodes into the environment until a steady outflow of nodes is reached. How the actual measurements are performed is defined by the *ranging protocol* that is used. Many of these protocols already exist and the one that is designed specifically for our applications is described in [5, 22].

Omnidirectional emission and reception is considered (as opposed to directional), because there is no a priori or online knowledge of the positions of neighbouring nodes. Whether or not messages are received by neighbouring nodes is dependent on the transmitted power and, among other factors, the received signal-to-noise ratio.

For convenience, but without loss of generality, in this paper we assume that distances can and are measured between all neighbouring nodes that are in line-of-sight and are within a fixed and known communication radius, r_{comm}.

Details of these protocols, as e.g. discussed in [22], are outside the scope of this paper, however, it is important to know that the messages that are transmitted and received between the nodes for measuring the distance between them need to contain identifiers. The identifiers are used to identify the sender of the message, and depending on the protocol also to verify whether the message is actually addressed to the receiving nodes. These identifiers make up the largest part (in bits) of the message.

The result of the ranging protocol is a set of measurements where every node i measures the distance $\mathrm{d}_{i,j}$ to every neighbouring node j, and node j having measured the distance $\mathrm{d}_{j,i}$ to node i. This is visualized in Fig. 2.3-a. The measurements are individually stored in the nodes' memory.

2.3.2   Distance uncertainty

The measured distance between nodes will be affected by different imperfections in the system. As the dynamics of the nodes and the environment they are in are unknown, a range of possible imperfections should be accounted for as we do not know the effect on the re-

(a) unique identification

(b) non-unique identification
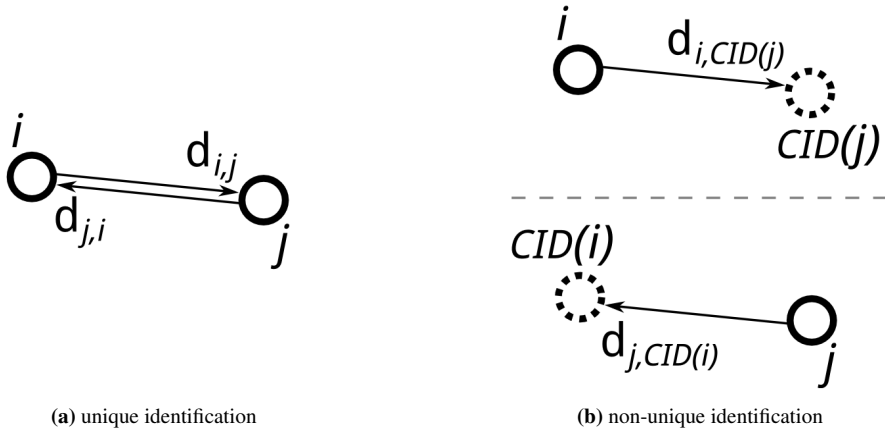
Figure 2.3: Mutual connections from distance measurements (indicated with arrows) with unique and non-unique identification. Distance measurements $d_{i,j}$ as measured by node i and distance measurement $d_{j,i}$ as measured by node j using unique identification can unambiguously be associated with each other. Distance measurements $d_{i,\mathrm{CID}(j)}$ as measured by node i and distance measurement $d_{j,\mathrm{CID}(i)}$ as measured by node j using non-unique identification can not be unambiguously associated with each other as more nodes use the same CID.

construction result. To simulate the measured distances, in this paper we consider a single snapshot of the actual distances and introduce all types of measurement noise to account for the existing imperfections. The types of noise in the distance measurements that we are modelling can be divided in inlier-type noise and outlier-type noise. The inlier-type measurement noise being:

- additive Gaussian noise $\mathcal{N}_a$; to account for e.g. offset in ranging timer or variable delay in electronics, and also to account for the errors due to movement of the nodes while performing the ranging measurements,

- multiplicative Gaussian noise $\mathcal{N}_m$; to account for e.g. different clock frequencies or inhomogeneous medium,

And the outlier-type measurement noise being:

- identification noise; to account for erroneous detection of identifiers,

- outlier noise; to account for burst-like / spiky / intermittent measurement noise.

The environment also gives rise to additional outlier-types of 'noise' that can influence the reconstruction:

- obstruction of signal paths,

- reflection of signals,

- loss of nodes or inability to retrieve them.

How exactly these are simulated is described later, in Sec. 2.5.2.

2.3.3 Identification uncertainty

Due to the constraints described in Sec. 2.2, the identifiers used in the ranging messages are reduced to non-unique identifiers, similar to the research in [5]. Each of the N nodes has a unique (hardware) *identification number*, UID, and is assigned a non-unique *communication identifier*, CID, that is used in the ranging message to identify the sender and addressee of the message. The number of available CIDs is indicated with n_{CID} and can be set as desired, $n_{\mathrm{CID}} \leq N$. The available CIDs are distributed uniformly among the nodes.

Now, instead of every node i measuring distance $\mathrm{d}_{i,j}$ to node j and vice versa, node i is measuring distance $\mathrm{d}_{i,\mathrm{CID}(j)}$ to a node that has a communication identifier $\mathrm{CID}(j)$. And vice versa, node j measures distance $\mathrm{d}_{j,\mathrm{CID}(i)}$. This is illustrated in Fig. 2.3-b. The mapping from UID to CID is known and well defined, however, the mapping from CID to UID is ambiguous as more nodes share the same CID.

2.3.4 Mutual connections

When the nodes are retrieved after the experiment and their data read out, offline analysis can be performed. Normally, with unique identification, mutual connections can be established by associating measurements $\mathrm{d}_{i,j}$ and $\mathrm{d}_{j,i}$ with each other (Fig. 2.3-a). These mutual connections can be used as consistency check of the individual measurements and quantify the uncertainty of the measured distance. When the two individual measurements differ less than a specific threshold, the measurements can be considered *inliers* and a mutual agreed-upon distance can be established, otherwise they are considered *outliers* and cannot be used to establish a mutual connection.

However, in the case of non-unique identification, measurement $\mathrm{d}_{i,\mathrm{CID}(j)}$ (Fig. 2.3-b) can only be used correctly when the measured non-unique identifier $\mathrm{CID}(j)$ is associated with node j. Obtaining the mutual connection where $\mathrm{d}_{i,\mathrm{CID}(j)}$ and $\mathrm{d}_{j,\mathrm{CID}(i)}$ are associated with the correct nodes j and i, respectively (similar to Fig. 2.3-a), is not trivial as the mapping from CID to UID is ambiguous.

Mutual connections, where $\mathrm{d}_{i,\mathrm{CID}(j)}$ is associated with $\mathrm{d}_{j,\mathrm{CID}(i)}$, are obtained by applying a consistency check on each of the possible node pairs that have CIDs consistent with the measurements. A mutual connection is hypothesised when two distance measurements agree on two grounds: 1) the measured CIDs agree with the CID of the other, and; 2) the corresponding measured distance is approximately similar to the other (i.e. inlier measurement). The exact implementation is described in Sec. 2.4.1.

This consistency check is affected by the different types of measurement noise in the individual measurements, specifically the outlier-type of noise that prevent finding the correct mutual connections. Fig. 2.4 illustrates how the different types of outlier-type noise affect finding correct mutual connections. Fig. 2.4-a shows a fully connected reference graph in which all mutual connections are present and correct. Outlier noise in the distance measurements, as shown in Fig. 2.4-b, leads to a set of $\mathrm{d}_{i,\mathrm{CID}(j)}$ and $\mathrm{d}_{j,\mathrm{CID}(i)}$ measurements that cannot be associated with each other as the difference in measured distance is larger than the inlier threshold. Identification noise can be seen in a similar way with the difference being that instead of erroneously measured distances the identifier CID is decoded erroneously. Obstruction of signal paths (Fig. 2.4-c) and the inability to retrieve nodes (Fig. 2.4-d) also
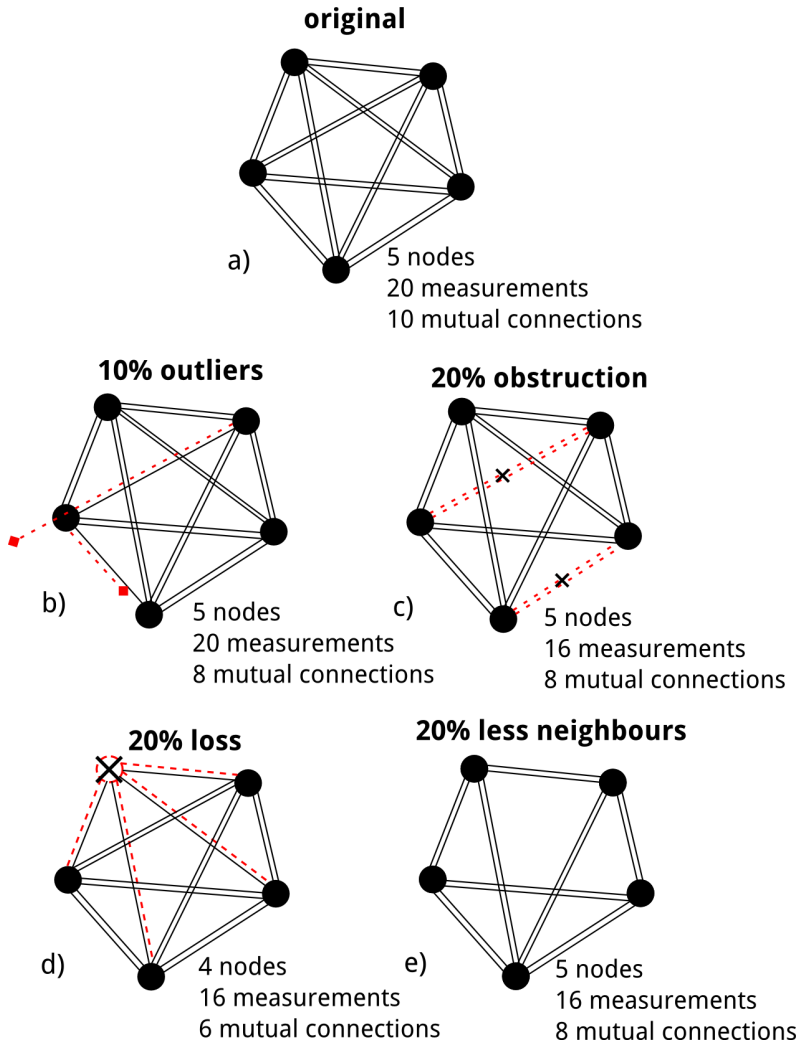
Figure 2.4: An example of (a) fully connected graph with five nodes; (b) outlier noise for two measurements; (c) obstruction between two nodes; (d) loss of one nodes; and (e) reduction of communication radius. Identification noise can be seen as outlier noise but on identification instead of distance. Explanation: Nodes are indicated as dark circles; a line indicates a distance measurement between two nodes. A too long/short line indicates a measurement with outlier noise. A dashed line indicates the distance measurement has never existed (obstruction) or this data is not retrieved (loss).

prevents establishing mutual connections. From a reconstruction perspective, not being able to establish mutual connections due to the presence of noise can to some extend be compared to reducing the number of neighbouring nodes (reducing the communication radius), which is visualized in Fig. 2.4-e.

2.4   Swarm reconstruction algorithm

In this section we will describe the novel reconstruction algorithm. It reconstructs node positions based on inter-node distances with significant measurement imperfections and the ambiguities introduced by using non-unique identifiers.

The reconstruction algorithm can be summarized as follows:

1. Hypothesise mutual connections based on the consistency (distance and identity) between all distance measurements, obtained from all retrieved nodes (Sec. 2.4.1);

2. Find initial four nodes to fix coordinate reference system (Sec. 2.4.2);

3. Robustly add additional nodes to the graph using a guided-RANSAC algorithm and general lateration, exploiting the geometric consistency between true neighbouring nodes (Sec. 2.4.3);

4. Perform non-linear optimization to reduce build-up of errors from previous step; go back to step 3 (Sec. 2.4.4);

5. Loop closing, if needed (Sec. 2.4.5).

An overview of the algorithmic steps of the reconstruction method is provided in 2.9.

2.4.1   Identity and consistency check

Before the reconstruction process, mutual connections should be established between nodes. The decoded non-unique communication identifier, CID, from every received ranging message should be associated with a unique (hardware) identification number, UID. Initially, this can be approached using a combinatoric method in which all possible nodes are considered that use the respective CID. The range of options can be narrowed down, as the distance measurement $d_{i,\text{CID}(j)}$ of node i to node j, when performed correctly and without noise, should be similar to the 'reverse' distance measurement $d_{j,\text{CID}(i)}$ of node j to node i. This means that searching the measured data for the consistency where $d_{i,\text{CID}(j)}$ is similar to $d_{j,\text{CID}(i)}$ should give most of the times an unambiguous mutual correct connection, since it is unlikely that another set of nodes with similar CID pair have the same matching distance.

However, the distance measurements are imperfect and noisy. Consequently, the distance measurements and the 'reverse' distance measurement will not be equal. Therefore, potential mutual connections are hypothesised when the measured distances lie within a specific threshold value ϵ_r of each other and the measured identifiers are consistent with each others

CID. This condition, denoted with $d_{a,b} \simeq d_{c,d}$, then becomes:

$$d_{a,b} \simeq d_{c,d} \quad \text{when} \quad \begin{cases} |d_{a,b} - d_{c,d}| \leq \epsilon_r & \text{and} \\ \text{CID}(a) = \text{CID}(d) & \text{and} \\ \text{CID}(b) = \text{CID}(c) \end{cases} \tag{2.2}$$

For all pairs that are considered consistent with each other, i.e. $d_{a,b} \simeq d_{c,d}$, a mutual connection is hypothesised and the average distance is taken as mutually agreed upon distance between them. All other pairs are not considered a mutual connection. The final consistency function $\mathcal{C}$ is expressed as:

$$\mathcal{C}(d_{a,b}, d_{c,d}) = \begin{cases} \frac{d_{a,b} + d_{c,d}}{2} & \text{when } d_{a,b} \simeq d_{c,d} \\ \varnothing & \text{otherwise} \end{cases} \tag{2.3}$$

where $\varnothing$ indicates that no mutual connection is considered, i.e. an empty entry.

Threshold value ϵ_r is the maximum allowed difference in distance measurement for which two measurements are considered inliers. This threshold should be related to the noise. It can e.g. be obtained using trial and error if the error model is unknown. Note that due to the measurement noise, the threshold, and the non-unique identification not all pairs that are considered consistent are correct, they are only hypothesised as being correct connections.

The obtained mutual connections between node $\bar{i}$ and node $\bar{j}$ and their hypothesised distance is not unambiguous. As illustrated in Fig. 2.5, the hypothesised mutual connections $h_{i,j}$ for measurement $d_{i,\text{CID}(j)}$ can consist of three types of contributions:

$$h_{i,j} = \begin{cases} \mathcal{C}(d_{\bar{i},\bar{j}}, d_{\bar{j},\bar{i}}) & (a) \\ \mathcal{C}(d_{\bar{i},\bar{j}}, d_{p,q}) & (b) \\ \mathcal{C}(d_{\bar{i},k}, d_{j,m}) & (c) \end{cases} \tag{2.4}$$

a) the measurement between the real nodes $\bar{i}$ and $\bar{j}$;

b) *identity ambiguity*: a mutual connection to a node p which is believed to be at similar distance as node $\bar{j}$ but not necessarily within communication radius of $\bar{i}$ (see Fig. 2.5);

c) *distance ambiguity*: a mutual connection with the correct node $\bar{j}$ but with a distance belonging to the measured distance to another node k within its communication radius (see Fig. 2.5).

The identity ambiguities are based on the statistical likelihood that somewhere in the swarm, a node pair with similar CIDs and distance is present such that the measurement pair obeys Eq. 2.2. Assuming a uniform spatial distribution of the nodes, the average number of identity ambiguities in $h_{i,j}$ scales with $\mathbb{A}_I \propto \epsilon_r n_{\text{comm}} N / n_{\text{CID}}^2$. Equally, the distance ambiguities are based on the statistical likelihood that neighbouring nodes give rise to confusion. The average number of distance ambiguities per $d_{\bar{i},\text{CID}(j)}$ then scales with $\mathbb{A}_D \propto \epsilon_r n_{\text{comm}}^2 / n_{\text{CID}}^2$.

Both type of ambiguities, $\mathbb{A}_I$ and $\mathbb{A}_D$, increase with an increasing number of neighbouring nodes n_{comm} and a lower n_{CID}, leading to a more challenging task deciding which are correct mutual connections and which are false connections. These ambiguities are resolved in our RANSAC graph-growing algorithm as detailed in Sec. 2.4.3.
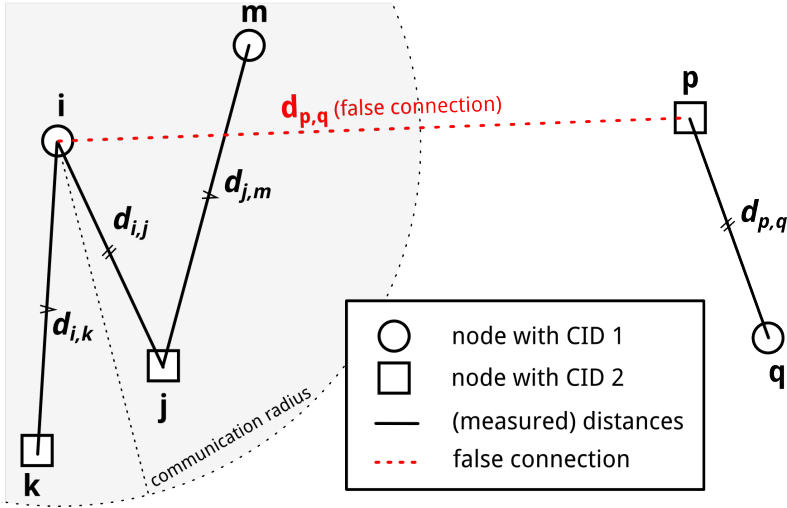
Figure 2.5: Nodes with UID $= \{i, j, k, m, p, q\}$ using only two distinct CIDs (indicated by $\bigcirc$ and $\square$) cause ambiguities in the consistency check. Identity ambiguity arises due to *similar* distance between pairs with similar CID in the swarm (Eq. 2.2). Distance ambiguities arise due to a plurality of nodes with similar CIDs within communication radius.

2.4.2 Initial seed selection

In this work we solve the reconstruction problem by incrementally growing the graph using a robust RANSAC-based method based on general lateration principles. Given an initial graph $\mathcal{G}$, new nodes can be added as illustrated in Fig. 2.2 and detailed in [6]. In an ideal case (without noise), every candidate node, c, of which the distances to four non-coplanar nodes with known positions $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4$ are known, can be added to the graph with as position the intersection of the spheres with radii $\bar{d}_{1,c}, \bar{d}_{2,c}, \bar{d}_{3,c}, \bar{d}_{4,c}$ and centers at $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4$. Here $\bar{d}_{n,c}$ denotes the true distance from node $n = \{1, 2, 3, 4\}$ to node c.

Since the graph growing is performed in $\mathbb{R}^3$ and the distance measurements are performed in $\mathbb{R}$ and without external beacons acting as fixed reference points in space, first a coordinate reference system has to be defined. The initial four node positions are chosen such that they define the coordinate system, therewith resolving the general reflection and rotation ambiguity. Their positions, $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4$, define the coordinate system as follows, and also illustrated in Fig. 2.2: $\mathbf{s}_1 \in \{0, 0, 0\}$, $\mathbf{s}_2 \in \{\mathbb{R}^+, 0, 0\}$, $\mathbf{s}_3 \in \{\mathbb{R}, \mathbb{R}^+, 0\}$ and $\mathbf{s}_4 \in \{\mathbb{R}, \mathbb{R}, \mathbb{R}^+\}$. These nodes are selected based on their connectivity and the stability of their geometric configuration which is obtained by general lateration techniques.

Since $h_{i,j}$ contains ambiguities, the initial seed selection is conditional until the graph growing has successfully added several nodes to the graph $\mathcal{G}$. If this is not possible, a new initial seed is selected and the process is repeated.
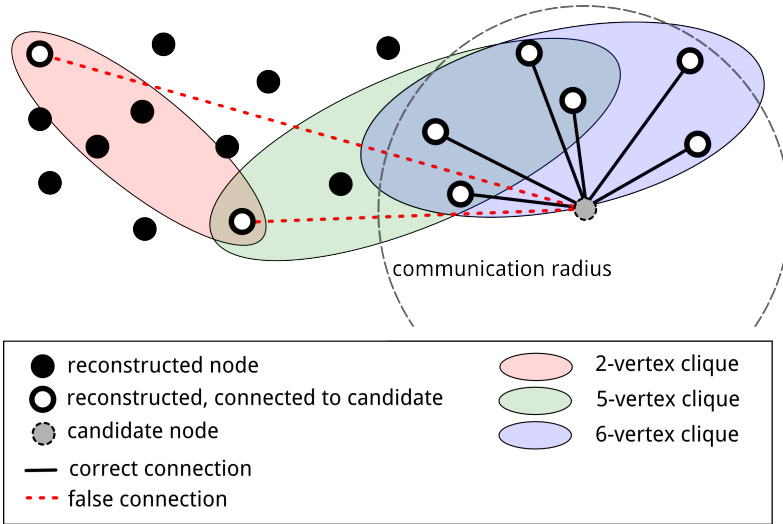
Figure 2.6: RANSAC graph growing algorithm uses an inlier-outlier voting system to filter out outlier distance measurements and proposes a position for candidate nodes to add to the graph. Using cliques of nodes in the graph that are within twice the communication radius helps in reducing the probability that nodes are positioned based on false connections.

2.4.3   RANSAC graph growing

After the initial seed is chosen, additional nodes can be added to the graph when the nodes have at least four connections to already reconstructed nodes. The set of candidate nodes that have at least four connections to already reconstructed nodes is denoted as $\mathcal{C}$. Ideally, $\mathcal{C}$ only consists of nodes that are true neighbouring nodes, but due to the ambiguities in the hypothesised connections $h_{i,j}$ this is not the case. Figure 2.6 illustrates a simplified 2-D situation in which a candidate node $c \in \mathcal{C}$ has connections to its true neighbouring nodes in $\mathcal{V}$, but also false connections to nodes somewhere else in the graph due to identity ambiguities. Whether the connections are true or false is not known at this point.

A Random Sampling Consensus (RANSAC) method is used to attempt to correctly position the candidate nodes. RANSAC is a general robust estimation technique, for more details we refer to [27]. RANSAC attempts to fit a model on data points that contain both inliers (points that satisfy the model) and outliers (points that do not satisfy the model). It does this by randomly selecting a small set of data points, estimating a model based on this small set, and counting the number of other data points that agree with this model. This process is repeated until a model is found that has maximum or sufficient support in the whole set of data points. We developed a specific RANSAC algorithm to solve the task of graph-growing under severe outlier noise and identification ambiguity. An overview of the RANSAC algorithm is provided in 2.10.

When a candidate node c is considered, all nodes that are already reconstructed in the graph and that are connected with this c are selected. This set of nodes is denoted with $A = \{a : a \in \mathcal{V}, \ h_{a,c} \neq \varnothing\}$. A subset of three of these nodes is selected, $A_p = \{a_p \in A\}$,

to *propose* – using general lateration – a position for c up to a reflection ambiguity (three node are always coplanar). However, no guarantees can be given whether the entries in $h_{a_p,c}$ are correct connections and the proposed positions do not need to be viable or close to the true position. Therefore, each of the other connected nodes $A_v = \{a_v : (A_v \cup A_p) = A, (a_v \cap a_p) = \varnothing\}$ are used to *vote* for the proposed position. Consensus is reached when the majority of the voting nodes, including the three proposing nodes, have at least a 50% majority. The reflection ambiguity is resolved by choosing which of the two positions received more supporting votes. When consensus is reached, the number of supporting votes is called the *RANSAC-score* for the specific proposal.

This step in the RANSAC algorithm is repeated with each time a different set of three proposers a_p, until all possibilities are exhausted or until a proposal received a specific threshold in RANSAC-score. In the latter case it is then considered 'sufficiently supported' and is added to the graph. When no 'sufficiently supported' condition is reached, the RANSAC procedure is repeated for a next candidate node until all candidates are considered or a 'sufficiently supported' condition is found for a specific candidate. The candidate with the highest RANSAC-score is added to the graph in $\mathcal{V}$ with edges ($\mathcal{E}$) only to the supporting nodes and corresponding $h_{a,c}$ entry.

The 'sufficiently supported' condition for the RANSAC can be based on the expected percentage of outlier measurements or can be chosen heuristically by trial-and-error, as is done in this work.

2.4.3.1   Guiding RANSAC

The number of false connections from the candidate nodes to the already reconstructed nodes, grows on average linearly with the number of nodes already in the graph, due to $\mathbb{A}_I$, as explained in Sec. 2.4.1. Consequently, when a candidate node is considered that does not have enough true neighbouring nodes in the graph yet but does have a large number of false connections to nodes already in the graph, this candidate node might receive a large RANSAC-score but based on only false connections. The true neighbouring nodes have in that case not enough voting power over the number of false neighbouring nodes.

To reduce the probability of this to happen, the sampling procedure of the candidate nodes is guided using a sorting order of candidate nodes before the RANSAC algorithm is performed on them. The sorting is based on the likelihood that candidate nodes have enough true neighbouring nodes already in the graph. As such, candidate nodes that are more prone to be reconstructed based on false connections, are considered later. This gives the probability that more of its true neighbouring nodes will be added to the graph first. The algorithm is provided in 2.11 and described next.

For each candidate node $\{b : b \in \mathcal{C}\}$, a list is made of connected nodes that have already been reconstructed in the graph $\{a : a \in \mathcal{V}, h_{a,b} \neq \varnothing\}$. A new graph $\mathcal{G}' = (\mathcal{V}(a), \mathcal{E}')$ is made with nodes a as vertices. These vertices are connected with edges $\mathcal{E}'$ when their positions are less than twice the communication radius, $2r_{\mathrm{comm}}$, apart. In this new graph, maximal cliques are listed as $\mathcal{Q}(b)$, and indicate the groups of nodes that are potentially within the communication radius of a proposed candidate node position. This means that on geometrical reasons, all of these nodes can potentially agree on candidate node position. In other words, nodes from outside this clique could never agree with all nodes within the clique

on a candidate node position. The example in Fig. 2.6 shows these cliques in coloured oval areas. The number of nodes in these cliques are registered and are used as sorting order in which RANSAC is performed. Nodes with larger sized cliques in $\mathcal{Q}(b)$ are more likely to have more connections with true neighbouring nodes in the graph.

This sorting has two effects on the reconstruction, it guides the RANSAC procedure and has as a result that: 1) it increases the probability that a 'sufficiently supported' candidate is found quickly, and 2) it increases the probabilitys that from all 'sufficiently supported' candidates, the best one is chosen for addition to the graph.

Guiding the RANSAC procedure with this sorting requires additional computational power (calculating the cliques), relative to unguided RANSAC. This increase in required computational power is compensated for by the quicker finding of a 'sufficiently supported' candidate. The reconstruction can take between 0.1-10 seconds per node, depending mainly on the connectivity and ambiguities.

2.4.4   Robust non-linear refinement

The stepwise addition of new nodes to the graph introduces build-up of errors. These errors in positions can prevent other nodes from being added. In order to reduce this error build-up, a global non-linear optimizer algorithm is executed as described in [6] to minimize the cost function Eq. 2.1.

For this we use the efficient general graph optimization software package g^{2}o [28]. It uses sparse methods to solve the normal equations at the core of the non-linear optimization techniques like Levenberg-Marquardt or a Gauss-Newton. It can easily solve graph problems consisting of thousands of nodes and edges.

In our algorithm, after every m newly added nodes, or failure to add a new node this non-linear optimization is performed. There is a clear trade-off between m and the processing time required for the reconstruction algorithm. In our work, m is chosen arbitrarily to 10. In future work, this can e.g. be adjusted dynamically based on uncertainty of previously added nodes.

2.4.5   Loop closing

The algorithm described above only considers edges between nodes in $\mathcal{G}$ when the specific connections supported the winning proposal in the RANSAC voting. As can later be seen in the results section, Sec. 2.5, this is effective against any type of outlier noise. But when considering loops in the environment, these loops might not be closed due to the previously mentioned error build-up, as schematically illustrated in Fig. 2.7.

Nodes that are supposed to connect both ends of the loop, will only be placed at one of both ends, depending on which has the largest number of supporters. Placement on the other end of the loop could also have yielded successful reconstruction, however, the node will not automatically be added to both ends to close the loop. An additional step should be performed in order to detect such loops and optimize the whole graph such that the loop is closed. This section describes the method how these loops are detected and closed.

Conceptually, the detection of loops is performed by searching for a set of nodes that could
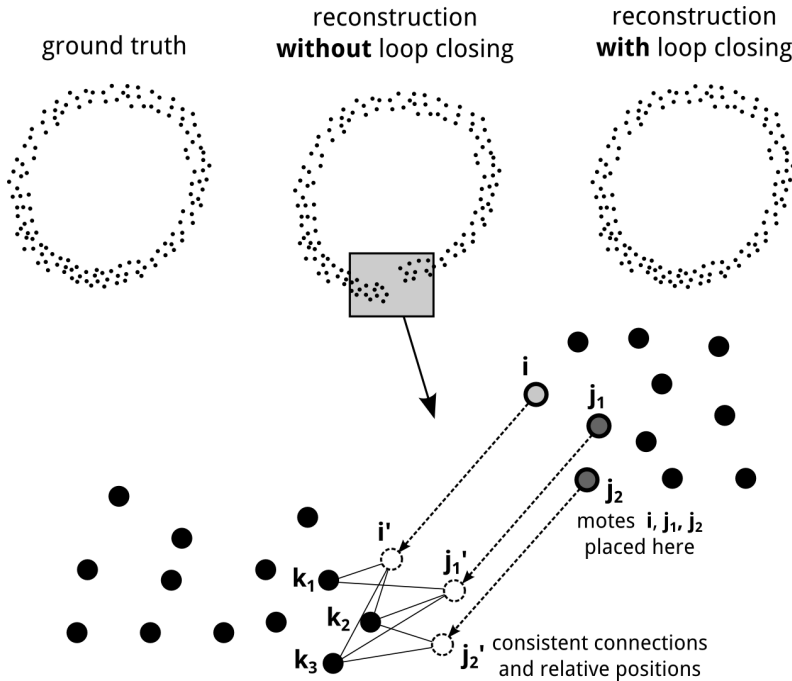
Figure 2.7: A 2-D example of an environment with a loop (ground truth). Due to error build-up in the RANSAC graph growing, loops are not closed (reconstruction withouth loop closing). Loop closing candidates can be detected by comparing local geometric consistency of excluded connections with reconstructed positions. After detection, the loop can be closed (reconstruction with loop closing).

have been reconstructed in two different, distant, positions in the graph. This starts by looking at all hypothesised connections between nodes that are not used in the reconstruction: i.e. all connections $h_{i,j}$ that did not support a winning RANSAC proposal, we call these *excluded connections*. The connections that did get included are called *included connections*.

Let node i have excluded connections to nodes k_{ex}^i and included connections to reconstructed neighbouring nodes j_{in}^i, as illustrated in Fig. 2.7. In order to successfully detect and close a loop, sufficient connectivity should exist between nodes j_{in}^i and k_{ex}^i.

After the reconstruction has stopped or halted, for every node i it is checked how many connections are excluded from the reconstruction, and how many of the neighbouring nodes j_{in}^i also have excluded connections to nodes k_{ex}^i. When several neighbouring nodes agree that nodes are excluded, the node is said to be *commonly excluded*. Nodes k_{ex}^i that are commonly excluded by nodes j_{in}^i are likely to be on an other end of a loop than nodes j_{in}^i.

Final determination of the loop closing can be performed by redoing the RANSAC graph-growing on this set of connections by forcing the nodes to be added to the graph on the other side of the loop. When this yields a successful reconstruction, and the local relative positions of the considered nodes is equivalent to the local relative positions on the original side of the

loop; then the loop can be closed. The loop closing itself can be performed by including the specific loop-closing edges that were initially excluded to the original graph and running the non-linear graph optimizer on the newly obtained graph.

2.5   Numerical simulations

In this section we will use numerical simulations to study the performance of the reconstruction algorithm, as well as how the different soft constraints (connectivity and identification ambiguity) influence this performance. The simulations are performed using two vastly different environment geometries, nodes are given a position in a pipeline environment and in a spherical environment (Sec. 2.5.1). The distance measurements are generated based on the distances between these positions and a variety of noise types is added to account for different measurement imperfections (Sec.2.5.2). These different noise types can influence the distance measurements such that the number of neighbouring nodes to which inlier measurements are available is reduced (Sec. 2.5.3). Assessing the performance of the reconstruction algorithm can be performed in different ways, depending on which qualities are favoured of the reconstruction. Therefore, several performance metrics are introduced to address these different qualities (Sec. 2.5.4). All parameters related to the soft constraints are swept over a broad range until the reconstruction algorithm will fail to reconstruct the node positions (Sec. 2.5.5). Knowing the failure point of our reconstruction algorithm and understanding the trade-offs between the underlying constraints helps in guiding the hardware design for the development of future miniaturized nodes. The findings are summed up by plotting the performance against the calculated effective number of neighbouring nodes to see the most important trade-off (Sec. 2.5.6). In the last part, examples are given to show the loop closing detection and actual closing in environments that require this (Sec. 2.5.7).

2.5.1   Environment model

The envisioned node swarms can be deployed in a variety of environments. In this paper we will work with two types of environments as schematically illustrated in Fig. 2.8. One resembles in abstract terms a mixing tank as the environmental dimensions are approximately the same in all directions. For convenience, the tank-like environment is chosen to be a bounded spherical environment. The second environment is a long pipeline, where the diameter of the pipeline is still large enough such that nodes can be positioned all around one another (as opposed to in one line along the pipe axis). The smooth pipeline has a fixed diameter of 8 cm and 400 nodes are placed in a section of 4 m pipe-length. The pipeline environment will be considered in both a loop and a loop-less fashion.

In the spherical environment, 400 nodes are given a random position within the boundary based on a uniform probability density function. The positions in the pipeline environment are chosen such that the nodes are spread out uniformly over the axis of the pipeline and the off-axis positions are chosen randomly based on a uniform probability density function over the cross-section. A total of 50 different spherical and 50 different pipeline environments are generated for the simulations. Examples can be seen in Fig. 2.9.

In both environments, the communication radius r_{comm} of the nodes is chosen such that each node has on average n_{comm} neighbouring nodes. This parameter n_{comm} is one of the
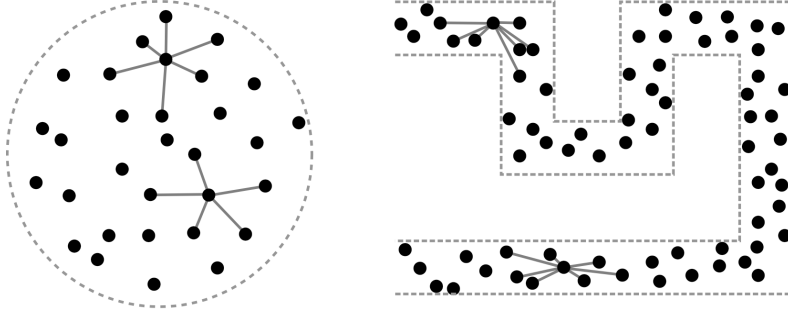
Figure 2.8: A 2-D interpretation of the 3-D spherical environment and 3-D pipeline environment. The lines between the nodes indicate connectivity for two different nodes.
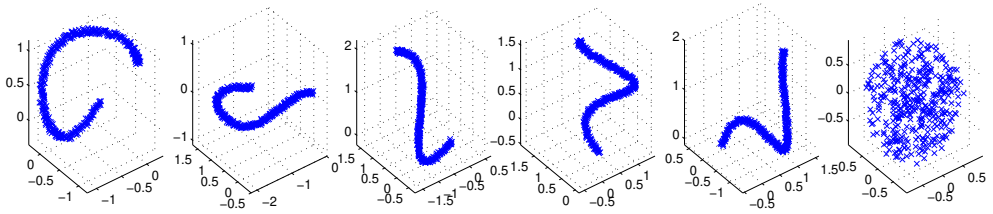
Figure 2.9: The ground truth positions of 400 nodes in several different experiments of randomly generated environments that are used for the simulations. Five pipeline environments, and one tank-like (spherical) environment.

parameters that will be swept. The nodes communication radius r_{comm}, in which ranging measurements are possible, is for all nodes within the swarm the same and ranges throughout the experiments from $10\mathrm{cm} \leq r_{\mathrm{comm}} \leq 15\mathrm{cm}$ (corresponding to $16 \leq n_{\mathrm{comm}} \leq 28$).

2.5.2 Distance measurement generation

Distance measurements are generated based on the generated node positions. The distances between all nodes i and j as measured by nodes i are indicated by $\mathrm{d}_{\bar{i},j}$ and are modelled by the true distances between the nodes $\bar{\mathrm{d}}_{\bar{i},\bar{j}}$ (the bar indicating true values).

Additive and multiplicative Gaussian are added to $\bar{\mathrm{d}}_{\bar{i},\bar{j}}$ according to:

$$\mathrm{d}_{\bar{i},\bar{j}} = \bar{\mathrm{d}}_{\bar{i},\bar{j}} + \bar{\mathrm{d}}_{\bar{i},\bar{j}}\mathcal{N}_{\mathrm{m}} + \mathcal{N}_{\mathrm{a}} \tag{2.5}$$

where $\mathcal{N}_{\mathrm{m}}$ and $\mathcal{N}_{\mathrm{a}}$ being perturbations from the zero-mean Gaussian distributions $\mathcal{N}(0, \sigma_{\mathrm{m}}^2)$ and $\mathcal{N}(0, \sigma_{\mathrm{a}}^2)$ respectively.

Furthermore, a variety of operations are executed on the measurements to account for different outlier noise types, each with their own statistical likelihood, ν, for this type of noise to happen to the measurement.

$$\mathrm{d}_{\bar{i},j} = \mathcal{F}_{\text{identification}}(\mathcal{F}_{\text{outlier}}(\mathcal{F}_{\text{obstruction}}(\mathcal{F}_{\text{loss}}(\mathrm{d}_{\bar{i},\bar{j}})))) \tag{2.6}$$

with

$$\mathcal{F}_{\text{identification}}(\mathrm{d}_{a,b}) = \mathrm{d}_{a,k} \qquad \text{when identification noise (probability } \nu_{\text{identification}})$$
$$\mathcal{F}_{\text{outlier}}(\mathrm{d}_{a,b}) = \mathcal{U} \qquad \text{when outlier noise (probability } \nu_{\text{outlier}})$$
$$\mathcal{F}_{\text{obstruction}}(\mathrm{d}_{a,b}) = \varnothing \qquad \text{when obstruction (probability } \nu_{\text{obstruction}})$$
$$\mathcal{F}_{\text{loss}}(\mathrm{d}_{a,b}) = \varnothing \qquad \text{when node } a \text{ is lost (probability } \nu_{\text{loss}})$$

and all operations return $\mathcal{F}(\mathrm{d}_{a,b}) = \mathrm{d}_{a,b}$ when the specific types of noise do not occur. Identification and outlier noise occur randomly to each measurement, independently of each other. Obstruction and loss results in non-existing measurements, indicated with $\varnothing$. Obstruction happens to both $\mathrm{d}_{\bar{i},j}$ and $\mathrm{d}_{\bar{j},i}$ simultaneously; and loss happens to individual nodes, independently of each other.

Identification noise results in an erroneously decoded identifier, i.e. node j is measured as being node k. Outlier noise $\mathcal{U}$ is chosen to be randomly drawn from a uniform distribution in the range $(0, r_{\text{comm}}]$.

Additionally, signal multipath can result in additional measurements of specific node pairs: besides the distance based on the direct signal path, also distances based on reflected, longer, signal paths. In order to account for this we add an additional distance measurement $\mathrm{d}'_{\bar{i},j}$ to the measurement dataset:

$$\mathrm{d}'_{\bar{i},j} = \bar{\mathrm{d}}_{\bar{i},\bar{j}} + \mathcal{W} \tag{2.7}$$

with $\mathcal{W}$ a perturbation drawn from a uniform distribution in the range $(0, r_{\text{comm}})$. The probability of this reflection happening is defined as $\nu_{\text{reflection}}$. For this reflection to be measured by node i, this distance should fall within the communication radius r_{comm}. A similar entry $\mathrm{d}'_{\bar{j},i}$ is added to account for the inverse measurement. Any reflection $\mathrm{d}'_{\bar{i},j}$ will also be subject to other discussed noise types.

This elaborate process generates realistic distance measurements and is significantly more extensive than that of previous research [6, 26, 29].

2.5.3   Effective connectivity

The discussed noise not only affects the accuracy of the distance measurements but also determines whether correct mutual connections can be estimated (Sec. 2.3.4). Due to noise, the effective number of neighbouring nodes to which correct distance measurements are made is therefore lower than the number of neighbouring nodes n_{comm} that are within communication radius.

The *effective connectivity*, EC, between nodes is defined by the number of neighbouring nodes to which inlier measurements (and thus correct mutual connections) are available for the reconstruction algorithm. An estimation of the effective number of neighbouring nodes can be made based on the theoretical likelihood that inlier measurements can be established in the presence of the mentioned outlier noise types:

$$\text{EC} = n_{\text{comm}}(1\text{-}\nu_{\text{outlier}})^2(1\text{-}\nu_{\text{id}})^2(1\text{-}\nu_{\text{loss}})(1\text{-}\nu_{\text{obstruction}}) \tag{2.8}$$

Table 2.2: Performance metrics of reconstructions shown in Fig. 2.10.

	$\mathbb{X}(\%)$	$\mathbb{Y}(\%)$	$\mathbb{Z}(\%)$	$\mathbb{E}_{abs}(m^2)$	$\mathbb{E}_{rel}$	$\mathbb{E}_{glob}(m^2)$	$\mathbb{E}_{loc}(m^2)$
Fig. 2.10-a	97	97	97	$5 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	$5 \cdot 10^{-5}$
Fig. 2.10-b	95	93	93	$1 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$2 \cdot 10^{-2}$
Fig. 2.10.c	88	64	0	$9 \cdot 10^{-1}$	$6 \cdot 10^{-2}$	$5 \cdot 10^{-1}$	$7 \cdot 10^{-2}$
Fig. 2.10-d	100	100	100	$1 \cdot 10^{-4}$	$2 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	$4 \cdot 10^{-7}$
Fig. 2.10-e	99	99	99	$2 \cdot 10^{-2}$	$1 \cdot 10^{-4}$	$5 \cdot 10^{-4}$	$4 \cdot 10^{-7}$
Fig. 2.10-f	99	94	94	$6 \cdot 10^{-3}$	$5 \cdot 10^{-3}$	$4 \cdot 10^{-3}$	$2 \cdot 10^{-4}$
Fig. 2.10-g	99	59	0	$2 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$6 \cdot 10^{-4}$
Fig. 2.10-h	100	20	0	$3 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$2 \cdot 10^{-1}$	$1 \cdot 10^{-3}$
Fig. 2.10-i	99	0	0	$1 \cdot 10^{-0}$	$3 \cdot 10^{-1}$	$2 \cdot 10^{-1}$	$5 \cdot 10^{-3}$

with ν indicating the probability that the specific outlier-type of noise are present in the individual measurements as defined in Sec. 2.5.2. The terms for outlier and identification noise are squared as these are effects happening to $d_{i,j}$ and $d_{j,i}$ independently from each other but both influence the ability to establish a correct bidirectional measurement.

The reconstruction algorithm is developed to be able to robustly filter out different noise types by relying on the geometric consistency between true neighbouring nodes and their measured distances to each other. The more inlier measurements that are available, the more robust the algorithm is. The effective connectivity EC is a predictor of the number of inlier measurements and therefore can be used as predictor of reconstruction performance. The usage of EC as such a predictor is verified in Sec. 2.5.6.

2.5.4   Reconstruction performance analysis

In our applications, or more in general for swarm operations like these, it is not established what is the best method to assess the performance of the reconstruction algorithm, i.e. to assess the quality of the reconstructed swarm. It is highly dependent on what information is favoured from such a swarm. Conventional metrics might not be suitable as the goal of the applications might be different. When exploring a yet unknown and difficult-to-access environment, initially one might want to know the overall structure of the environment: e.g. a rough estimation of the local geometry and the overall shape of the total environment. Later, e.g. when also adding additional sensor information, one might be more interested in fine-grained local geometry where the absolute error over the entire swarm is less relevant. We therefore present different performance metrics that serve different goals.

2.5.4.1   Absolute error

The absolute error is the mean squared absolute error of all reconstructed node positions relative to their ground truth positions. It is calculated as $\mathbb{E}_{abs} = \sum_i^N \|\hat{\mathbf{s}}_i - \bar{\mathbf{s}}_i\|^2/N$, the sum over all N nodes where $\hat{\mathbf{s}}$ is the reconstructed position of the node and $\bar{\mathbf{s}}$ its ground truth position. In order to compare the ground truth positions with the reconstructed positions, a linear fit between the 3-D positions of the initial four seed nodes and their ground truth
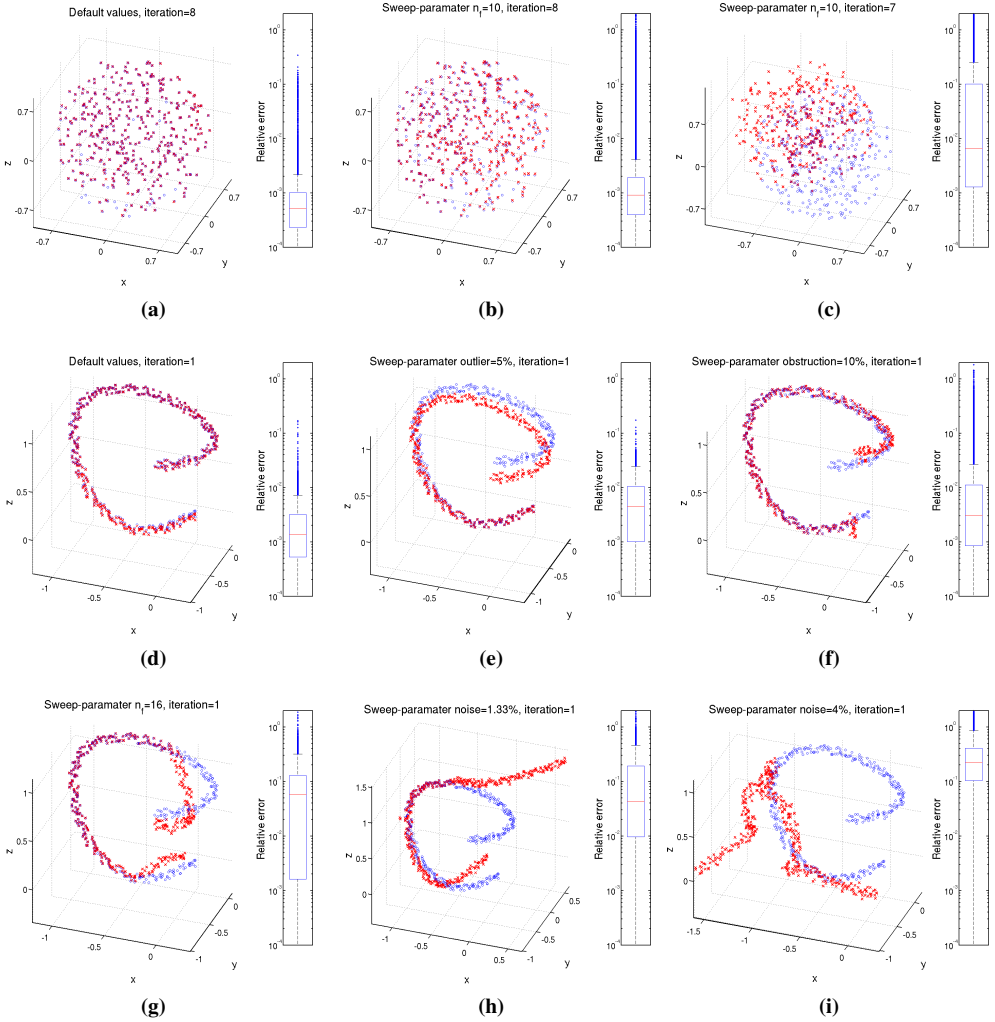
Figure 2.10: Reconstructions of node swarms in a spherical environment (a-c) and in a pipeline environment (d-i) with different parameter sets. Blue circles are the ground truth positions of the nodes and red crosses the reconstructed positions. The recall metrics $\mathbb{X}, \mathbb{Y}, \mathbb{Z}$ and the error metrics $\mathbb{E}_{abs}, \mathbb{E}_{rel}, \mathbb{E}_{glob}, \mathbb{E}_{loc}$ are assess the performance of the reconstruction are listed in Table. 2.2.

positions is performed to resolve the general rotation and reflection ambiguity.

The mean squared absolute error is a commonly used performance metric in reconstruction studies. A disadvantage of this metric is that when a reconstructed swarm exhibits error build-up, the absolute error between one side of the swarm and the other can be large, while the local errors can actually be small. In swarm reconstruction studies like this, this metric should therefore not be used as the only metric to assess reconstruction performance.

2.5.4.2 Relative error

The relative error is the mean squared relative error of the reconstructed distances between nodes, relative to the ground truth distances. It is calculated as $\mathbb{E}_{rel} = \sum_{i,j}^{M}\|\hat{\mathbf{d}}_{i,j}-\bar{\mathbf{d}}_{i,j}\|^2/M$, the sum over all M reconstructed distances, where $\hat{\mathbf{d}}$ is the reconstructed distance and $\bar{\mathbf{d}}$ the ground truth distance.

Unlike $\mathbb{E}_{abs}$, the metric $\mathbb{E}_{rel}$ is not affected by build-up of errors. It assesses the reconstruction performance only on a local and relative scale.

2.5.4.3 Global error

The global error $\mathbb{E}_{glob}$ is defined similar as $\mathbb{E}_{abs}$, but includes the performing of a rigid transform (transformation and rotation) of the entire reconstructed swarm such that $\mathbb{E}_{abs}$ is minimized.

This metric illustrates the reconstruction performance of the swarm better than $\mathbb{E}_{abs}$ as it is not based on a specific chosen seed from which everything is built on. However, it is still sensitive to error build-up.

2.5.4.4 Local error

The local error $\mathbb{E}_{loc}$ is defined as mean squared absolute error, but only after performing a rigid transform of all subsections of 20 connected nodes with their ground truth positions.

This heuristically chosen performance metric considers only the local errors, but other than $\mathbb{E}_{rel}$, it does focus on the (absolute) reconstructed positions of the nodes rather than the (relative) distances between them.

2.5.4.5 Recall

The *recall*, $\mathbb{X}$, is the percentage of nodes reconstructed by the reconstruction algorithm. As this does not take into account whether or not these nodes are reconstructed correctly, we introduce the adjusted recall $\mathbb{Y}$ and $\mathbb{Z}$. These are the percentages of nodes that are reconstructed within a specified error condition that is heuristically chosen. For $\mathbb{Y}$, this condition is when at least 80% of the node's reconstructed distances (or edges in $\mathcal{G}$) have a relative error $<10\%$; or when more than 50% of the nodes reconstructed distances have a relative error of $<1\%$. The adjusted recall $\mathbb{Z}$ is similar to $\mathbb{Y}$, but includes the condition that the neighbouring nodes to which the relative error suffices this condition, should also fall in the category of $\mathbb{Y}$. Note that this categorization of recall into $\mathbb{X}$, $\mathbb{Y}$ and $\mathbb{Z}$ is based on a heuristic, subjective interpretation of the reconstruction result.

Examples of reconstructions are shown in Fig. 2.10 and the corresponding performance metrics are listed in Table 2.2. The distribution of the relative errors of the reconstructed distances is indicated with a boxplot on the right of each reconstruction. Only Fig. 2.10-d and Fig. 2.10-a show reconstructions with a very low $\mathbb{E}_{abs} \leq 1.4e\text{-}4m^2$ and visually seem like a 'perfect' reconstruction. All others seem to exhibit their own erroneous characteristics, but non of them are completely wrong. Fig. 2.10-e has a fairly high $\mathbb{E}_{abs}$ but $\mathbb{E}_{glob} = 4.7e\text{-}4m^2$ with the same order of magnitude as $\mathbb{E}_{abs}$ in Fig. 2.10-d. The reconstruction only seems to suffer slight error build-up that leads to large absolute errors at the pipe end but for the rest has an accurate reconstruction. For the majority of the swarm in Fig. 2.10-f, the positions are accurately reconstructed. Only at the ends of the pipe the reconstruction algorithm seems to have created large errors. In Fig. 2.10-h and Fig. 2.10-g the reconstructions also seems correct on the local scale (e.g. seen in the relative low $\mathbb{E}_{loc}$) but exhibits some local errors that cause the reconstructed graph to deviate from the original axis of the pipeline. These differences illustrate the importance of having different error metrics as the assessment should happen based on which information is favoured from the swarm. As an extreme, even in Fig. 2.10-i the diameter of the pipe can still be estimated from the reconstruction while the shape of the swarm is very inaccurate.

Deciding whether a reconstruction gives satisfactory results or not is therefore not trivial and depends on the application goals. The adjusted recall parameters are an attempt to quantify these goals. In this paper, on the basis of subjectively interpreting how well the shape of the swarms in Fig. 2.10 are reconstructed, recall parameter $\mathbb{Y}$ can be seen as a reasonable metric to assess the performance of the reconstructions. We can define a satisfactory reconstruction to be one with e.g. $\mathbb{Y} > 80\%$.

2.5.5 Parameter sweeps and breaking points

In this section we study the performance of the reconstruction algorithm while sweeping the parameters related to the soft constraints. These parameters are the ones involved with measurement noise as described in Sec. 2.3.2, the number n_{CID} of non-unique identifiers CID, and the number of neighbouring nodes within communication radius n_{comm}. In order to change the number of neighbouring nodes we actually change the sensing radius r_{comm} such that on average, nodes have n_{comm} neighbouring nodes.

For each set of parameters, 50 different experiments are performed, each with randomly generated node positions (but following the environment geometry constraints as defined in Sec. 2.5.1). All experiments are performed with a total number of $N = 400$ nodes. For similar work with a smaller and larger number of nodes, the reader is referred to our earlier work [7] and [6].

The Gaussian noise components in $\mathrm{d}_{\bar{i},\bar{j}}$ are together summarized as ν_{gaussian}, as standard deviation of $(\mathrm{d}_{\bar{i},\bar{j}} - \bar{\mathrm{d}}_{\bar{i},\bar{j}})$. This value is expressed as percentage of a fixed communication radius[2]:

$$\nu_{\text{gaussian}} = \sqrt{\frac{\sum_{i,j}^{M}(\mathrm{d}_{\bar{i},\bar{j}} - \bar{\mathrm{d}}_{\bar{i},\bar{j}})^2}{M-1}} \frac{1}{r_{\text{comm}}[n_{\text{comm}}=20]} \tag{2.9}$$

---

[2]the average communication radius when $n_{\text{comm}} = 20$
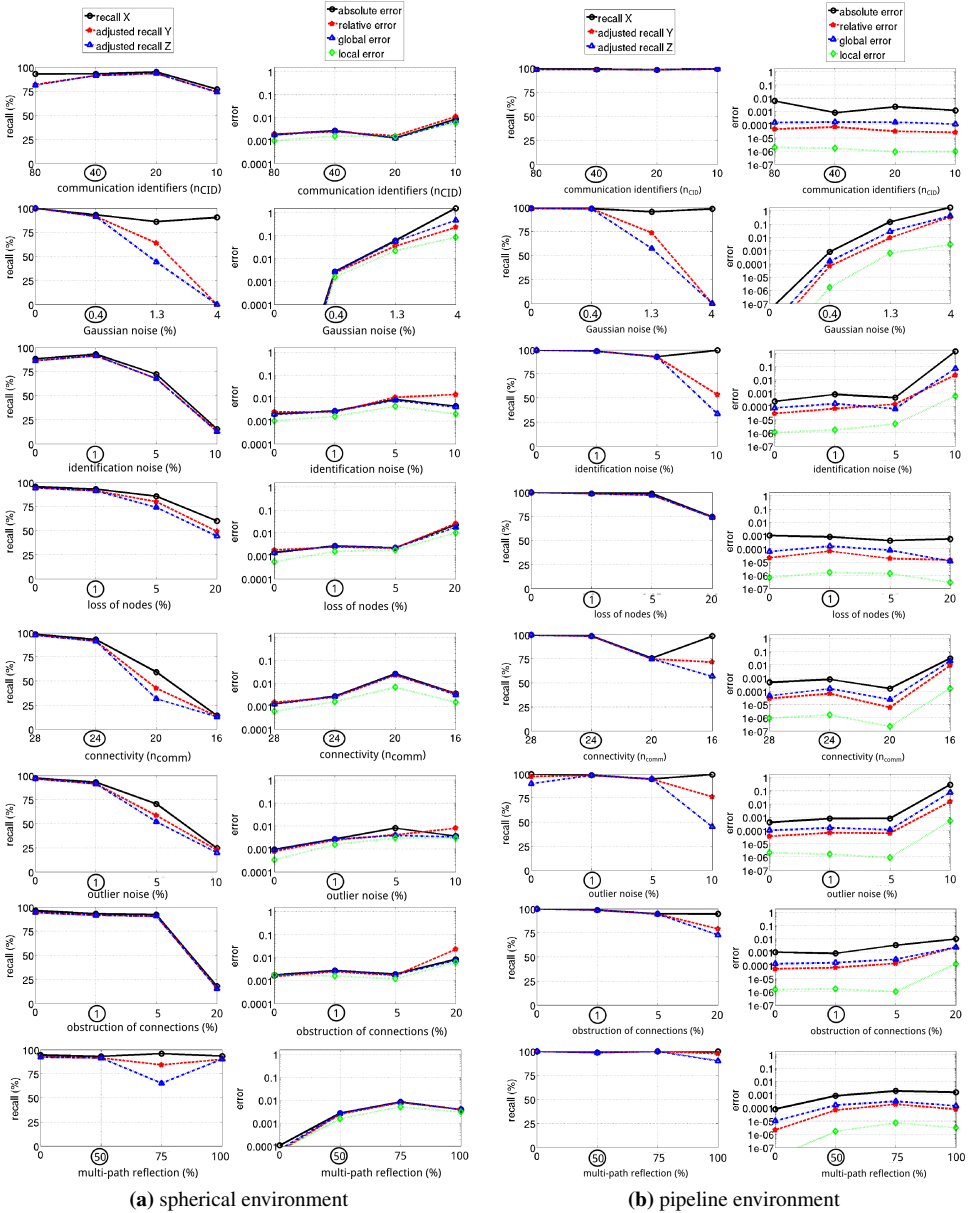
Figure 2.11: Median of the recall result $\mathbb{X}, \mathbb{Y}, \mathbb{Z}$ on the left and error metrics $\mathbb{E}_{abs}$, $\mathbb{E}_{rel}$, $\mathbb{E}_{glob}$, $\mathbb{E}_{loc}$ on the right. Indicated parameters on the x-axis are swept while keeping other parameters on a fixed value (encircled). From top to bottom these are: $\{n_{\text{CID}}, \nu_{\text{gaussian}}, \nu_{\text{identification}}, \nu_{\text{loss}}, n_{\text{comm}}, \nu_{\text{outlier}}, \nu_{\text{onstruction}}, \nu_{\text{reflection}}\}$

Fig. 2.11-a shows for a spherical environment the recall $(\mathbb{X}, \mathbb{Y}, \mathbb{Z})$ and the different error metrics $\mathbb{E}_{abs}$, $\mathbb{E}_{rel}$, $\mathbb{E}_{glob}$, $\mathbb{E}_{loc}$ while sweeping the input parameters. The median values of these metrics of all 50 experiments is visualized. It can clearly be seen that when increasing the measurement noise (Gaussian, outlier, identification, obstruction and loss), the recall drops and the error metrics increase.

Fig. 2.11-b shows the performance metrics using similar input parameters, but in a pipeline environment. It can be seen that the recall metrics in this case are much higher than in the spherical environment. This is due to a different distribution of connectivity among the nodes, as explained next in Sec. 2.5.6.

Furthermore, the error metrics in the pipeline environment show significantly more differences among them than in the spherical case. The local error is orders of magnitude lower than for example the absolute and global error. The environment stretches across larger distances and has smaller local dimensions compared to the spherical case; errors build up easier and cause a larger absolute error, even in cases where the local geometry is reconstructed correctly. This can clearly be seen in e.g. Fig. 2.10-e/h.

The observed dip in recall in the spherical environment in Fig. 2.11-a when reducing the number of identifiers to $n_{\mathrm{CID}} = 10$ is mainly due to a self-imposed time constraint in the reconstruction algorithm (when average time per reconstructed node exceeds 10 seconds). The ambiguities in this dataset are around $\mathbb{A}_{\mathrm{total}} = 200\%$. Reducing the number of identifiers further can still yield successful reconstruction but drastically increases the number of ambiguities and therewith the time required to calculate cliques as described in Sec. 2.4.

The lines in Fig. 2.11-a and Fig. 2.11-b only show the trend of the parameter sweep and due to the nature of the reconstruction process show statistical quirks. The standard deviation between all 50 experiments of the same input parameter sets are large and not shown to preserve readability of graphs. For example, in cases where the majority of experiments yield high recall (e.g. >90%), it is not uncommon that in one or more experiments of the same parameter set the recall does not surpass 10%.

The general trend of these lines clearly show there is a dependency of the input parameters on the reconstruction performance. The reconstruction exhibits graceful degradation when the soft constraints increase in severity up to the point where the recall drops and error increases.

2.5.6   Effective number of neighbours

The recall result $\mathbb{Y}$ of our experiments is plotted against the meta-parameter EC in Fig. 2.12. For both environments, a clear trend is visible that the recall is dependent on the effective number of neighbouring nodes. A minimum number of inlier measurements is required that can be used for the RANSAC voting to guarantee satisfactory reconstruction of the nodes' positions. A minimum (effective) connectivity of around EC $= 20$-22 is required in order to reconstruct the swarm with an adjusted recall $\mathbb{Y} > 80\%$.

Even though the average number of neighbouring nodes are chosen to be similar in both the spherical and the pipeline by setting r_{conn} for all nodes, the variability in the per-node connectivity in the spherical environment is much larger than in the pipeline environment. Fig. 2.13 shows the histogram of the per-node connectivity, averaged over all experiments.

(a) spherical environment   **(b)** pipeline environment
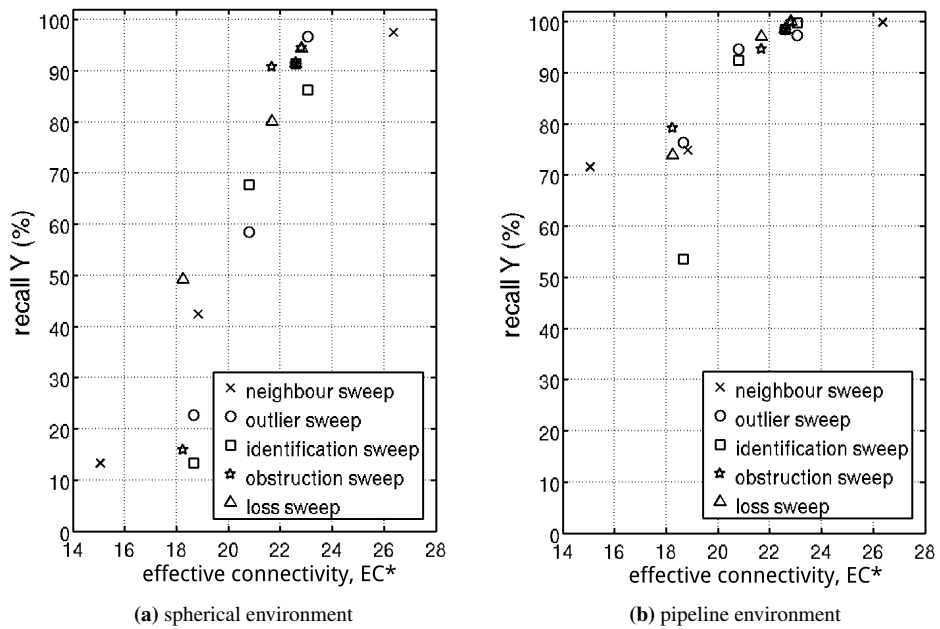
Figure 2.12: Adjusted recall result $\mathbb{Y}$ versus the meta-parameter EC, the effective number of neighbours.
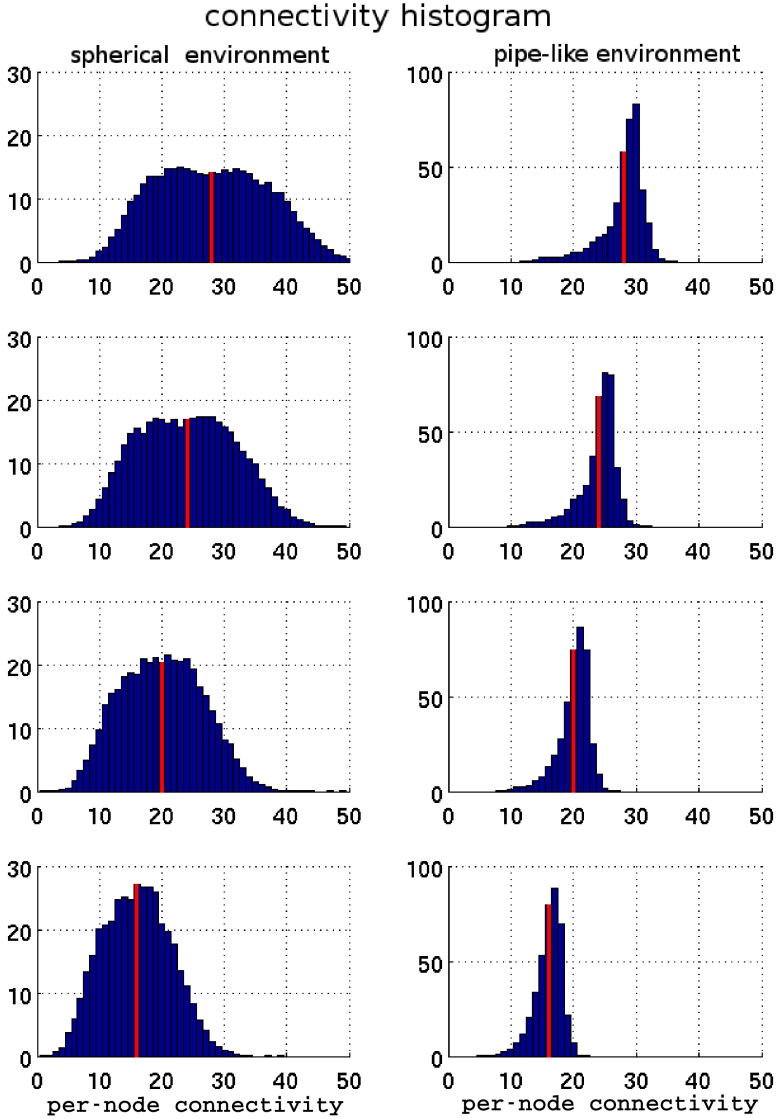
Figure 2.13: Histogram of real neighbouring nodes within communication radius. The average number of neighbours, n_{comm}, is indicated with a red line. From top to bottom $n_{\mathrm{comm}} = \{28, 24, 20, 16\}$; left the spherical environment right the pipeline environment.

(a) single loop environment

(b) single loop environment

(c) multiple loop environment
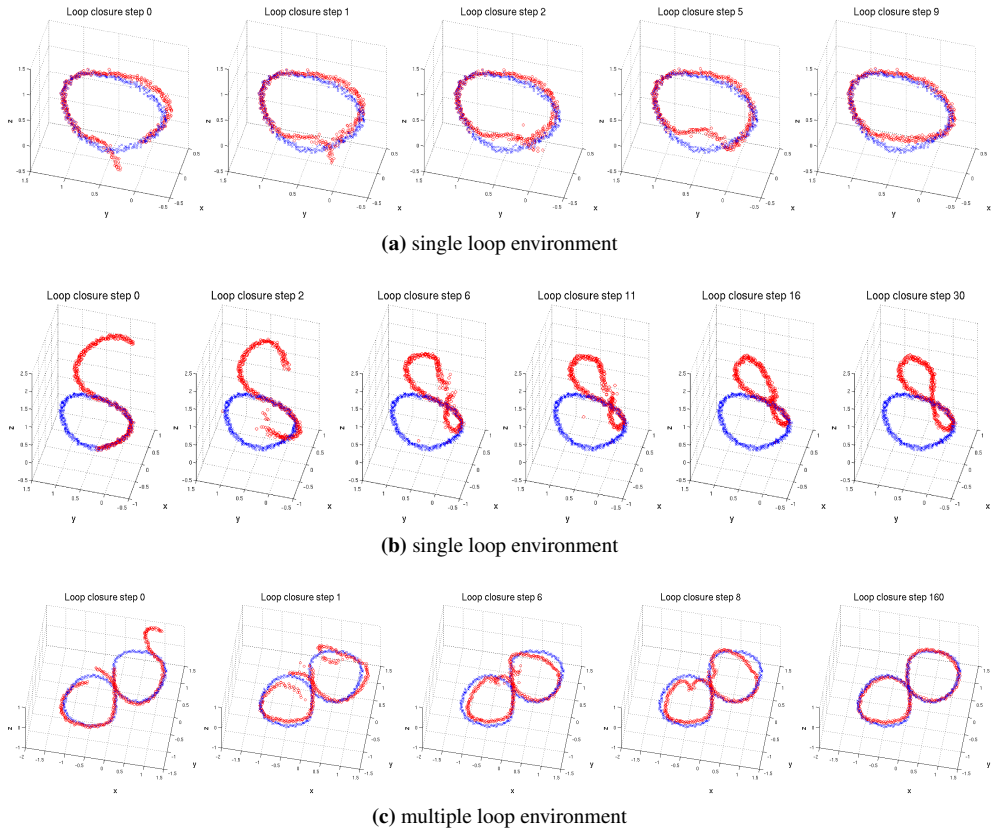
Figure 2.14: Several steps of the loop closing phase using non-linear optimization. Blue is ground truth positions and red is reconstructed positions.

The chosen n_{comm} for each of the experiments is indicated using a red line, this is the average value of the histogram. The nodes with a lower connectivity are much less likely to be reconstructed (correctly) and will influence the recall result of the entire swarm.

It should be noted that the results in Fig. 2.11 and 2.12 show the median values over 50 experiments, each time with a different swarm topology and node positions. The findings presented here concern the reconstruction of the full swarm in a single reconstruction experiment. Smaller sub-swarms can be reconstructed using lower connectivity.

2.5.7 Loop closing

Many environments will contain one or multiple loops; therefore it is important that the reconstruction algorithm can deal with closing these loops. Our loop closing method as described in Sec. 2.4.5 is evaluated in this section.

Examples of reconstructed swarms that require loop closing, and the iterative loop closing in the non-linear optimization process are shown in Fig. 2.14. Build-up of error has caused
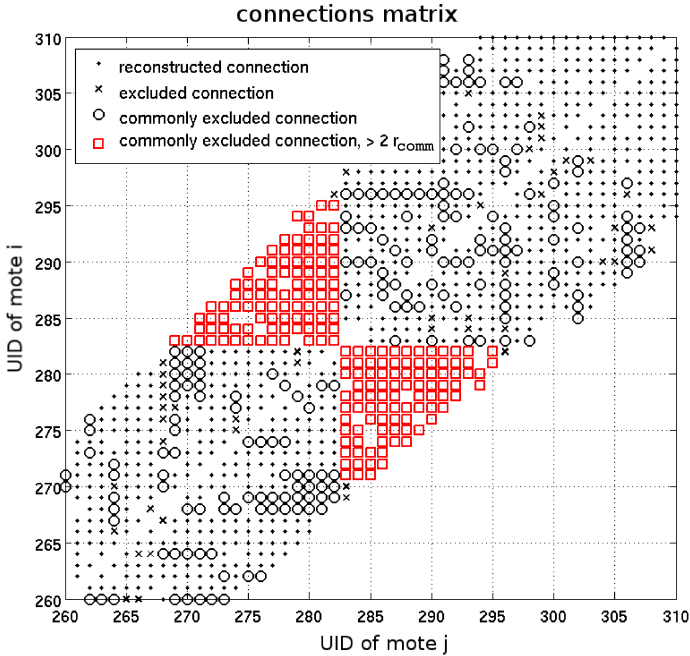
Figure 2.15: This connection matrix shows reconstructed connections (·) and excluded connections (×). It indicates excluded connections that are excluded by a plurality of neighbouring nodes (○) as well as those with a distance $> 2r_{\text{comm}}$ (□). Loop closing candidates can easily be observed using this categorization. The linear ordering of the nodes' UID along the pipe axis is solely used for illustrative purposes. It is not used in the reconstruction as only CIDs are available from measurements.

the two ends of the loop(s) to not end up at the same position. Fig. 2.15 shows the result of the loop closing detection algorithm (Sec. 2.4.5) of the dataset shown in Fig. 2.14-b. The commonly excluded connections of which the reconstructed distance is larger than $2r_{\text{comm}}$ are depicted with red squares and clearly indicate that the two ends of the reconstructed swarm are supposed to be connected. The nodes that are reconstructed at the loop-ends do not receive support from nodes on the other side of the loop in the RANSAC voting algorithm. To close the loop, these detected connections that are originally excluded from the graph are now included in the graph. Performing the non-linear optimization using these new connections, now closes the loop as seen in the consecutive steps in Fig. 2.14.

As seen in Fig. 2.14-b, loop closing can distort the reconstructed swarm when the loop-ends are too far from each other. Closing the loop using this non-linear graph optimization then forces a different swarm geometry in such cases.

These experiments shows that our method can perform loop closing. Multiple loops can be dealt with simultaneously or individually, as seen in Fig. 2.14-c.

2.6   Discussion and future work

The results show that the reconstruction algorithm can robustly deal with a large set of different types of measurement noise. Each measurement imperfection has a different influence on the established mutual connections and on the reconstruction. The inlier-type of measurement noise are present in all measurements and only affect the accuracy of the reconstructed positions. The outlier-type of noise prevents establishing correct mutual connections and on top of that also create false and ambiguous connections due to the non-unique identification. Using the geometric consistency of the positions of true neighbouring nodes, the reconstruction algorithm can robustly filter out the false connections and prevent them from negatively influencing the reconstruction. Reducing the quantity of available CIDs significantly increases the ambiguities in establishing mutual connections. However, it is found that due to the robust nature of the reconstruction algorithm, these ambiguities could be resolved when enough inlier measurements are available.

The robustness of our novel algorithm allows us to see that the key parameter for successful reconstruction is the number of inlier node connections that are available for reconstruction. Initially, the total number of distance measurements is determined by the communication range and the density of the nodes in the swarm. It is due to the outlier-type of measurement noises that not all node connections will be inlier measurements and hence can not be used for reconstruction. The effective connectivity can be predicted by our Eq. 2.8 and is an important predictor for the reconstruction performance. We can therefore conclude that in the sensor node hardware design for these applications, resources should be focussed on increasing the communication radius to increase connectivity and to prevent outlier-types of noise from reducing the effective connectivity. Achieving this can for example go at the cost of (unique) identification.

Depending on which information is favoured from the swarm, the error metrics can be used to study the performance of the reconstruction. There is no single error metric that can summarize the reconstruction performance for all application goals. Using different error metrics and the heuristically chosen adjusted recall metrics $\mathbb{Y}$ and $\mathbb{Z}$ we showed that the subjective interpretation of the quality of the reconstruction can be partially quantified. When the sensor nodes need to be designed for specific applications, a performance condition can be set up, e.g. $\mathbb{Y} > 80\%$, to assess under which input conditions the reconstruction is most likely to achieve the requested performance.

2.7   Conclusion

This work evaluated the feasibility of using sensor swarms, consisting of many highly miniaturized and severely resource limited sensor nodes, to reconstruct difficult-to-access environments. For this, a novel Guided Random Sample Consensus algorithm together with nonlinear graph optimization is proposed. Its main contribution is that the algorithm can handle severe and different types of measurement errors as well as the use of ambiguous non-unique communication identifiers.

Extensive and realistic simulations show that sensor node connectivity is to be favoured over unique identification of nodes. When the number of used communication identifiers is

only 2.5% of the total number of nodes ($n_{\text{CID}} = 10$, $N = 400$), the entire sensor swarm can still be reconstructed successfully. Thereby allowing less identification information to be used, i.e. shorter communication bursts, to trade-off for larger emission power per burst. This, in turn, allows for improved sensor node connectivity, as more nodes can be reached with a single burst. On the contrary, simulations show that when swarm connectivity drops below 20 inter-node connections per node, reliable sensor swarm reconstruction is hampered for this algorithm. Therefore, *favouring connectivity over identification*, is a pivotal finding of our work, as it can and will guide future developments of highly miniaturized sensor nodes.

2.8 Acknowledgement

References

[1] E. Talnishnikh *et al.*, "Micro Motes: A Highly Penetrating Probe for Inaccessible Environments", in: *Intelligent Environmental Sensing*, edt. H. Leung, Springer International Publishing, 2015, pp 33-49.

[2] European Union's Horizon 2020 FET-Open project: PHOENIX (No 665347). www.phoenix-project.eu

[3] N. Patwari *et al.*, "Cooperative localization in wireless sensor networks", IEEE Signal Processing Magazine (54), July 2005

[4] H. Wymeersch *et al.*, "Cooperative Localization in Wireless Networks", in *P IEEE*, 97(2):427–450, Feb. 2009.

[5] H.A. Duisterwinkel *et al.*, "Asymmetric Multi-Way Ranging for Resource-Limited Nodes", 8th EAI International Conference on Ad Hoc Networks (AdHocNets), 26-27 September 2016, Ottawa, Canada

[6] G. Dubbelman *et al.*, "Robust Sensor Cloud Localization from Range Measurements", in *IEEE Int. Conf. Intelligent Robots and Systems*, Chicago, Illinois, USA, Sept. 2014.

[7] H.A. Duisterwinkel *et al.*, "Mapping Swarms of Resource-Limited Sensor Motes: Solely Using Distance Measurements and Non-Unique Identifiers", IEEE Symposium on Computational Intelligence for Engineering Solutions (SSCI 2016), Athens, Greece, 6-9 December 2016

[8] Jules S. Jaffe *et al.*, "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics", NAT COMMUN 8, Article number: 14189 (2017)

[9] J. Lu, et al, "A small long-life acoustic transmitter for studying the behavior of aquatic animals", Review of Scientific Instruments 87, 114902 (2016); doi: 10.1063/1.4967941

[10] E.H.A. Duisterwinkel, et al., "Sensor Motes for the Exploration and Monitoring of Operational Pipelines", IEEE Transaction on Instrumentation and Measurements, DOI: 10.1109/TIM.2017.2775404

[11] H.J. Wörtche, "Advanced Analysis: Motes Dynamics in a Mixing Tank," INGU Solu-

tions, 2014.

[12] H. Durrant-Whyte, and T. Bailey, "Simultaneous localization and mapping: part I", IEEE ROBOT AUTOM MAG, Volume:13 , Issue: 2, p.99, June 2006

[13] J. Djugash *et al.*, "Range-only SLAM for Robots Operating Cooperatively with Sensor Networks", in *IEEE Int. Conf. Robotics and Automation*, p.2078–2084, Orlando, Florida, USA, 2006.

[14] D. Haehnel, "Mapping and Localization with RFID Technology", In: Proc. IEEE International Conference on Robotics and Automation (ICRA), New Orleans, LA (USA)

[15] J. Neira, A. Davison, and J. Leonard, "Guest editorial, special issue in visual slam," IEEE T ROBOT, vol. 24, no. 5, pp.929–931, October 2008

[16] R. Kirkham *et al.*, "PIRAT—A System for Quantitative Sewer Pipe Assessment", INT J ROBOT RES, November 2000 vol. 19 no. 11 1033-1053

[17] F. Ribeiro, D. *et al.*, "Geometrically Constrained Room Modeling with Compact Microphone Arrays", Audio, Speech, IEEE T AUDIO SPEECH, Volume:20 , Issue: 5, p.1449, December 2011

[18] I. Dokmanić, *et al.*, "Acoustic echoes reveal room shape", PNAS, vol 110, no 30, 2013

[19] E. Menegatti *et al.*, "Range-only SLAM with a Mobile Robot and a Wireless Sensor Network", in *IEEE Int. Conf. Robotics and Automation*, p.8–14, Kobe, Japan, May 2009.

[20] A. Franchi *et al.*, "Mutual Localization in Multi-Robot Systems using Anonymous Relative Measurements", INT J ROBOT RES, May 2013

[21] I.F. Akyildiz, D. Pompili and T. Melodia, "Underwater acoustic sensor networks: research challenges", AD HOC NETW 3 (2005) 257–279

[22] E.H.A. Duisterwinkel, PhD dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands, 2018, to be published

[23] I. Borg and P. Groenen (2005). "Modern Multidimensional Scaling: theory and applications" (2nd ed.). New York: Springer-Verlag. pp. 207–212. ISBN 0-387-94845-7.

[24] Roudy Dagher *et al.*, "Localization in Wireless Sensor Networks" Wireless Sensor and Robot Networks From Topology Control to Communication Aspects, Worldscientific, pp.203-247, 2014, 978-981-4551-33-5.

[25] Jonathan Bachrach and Christopher Taylor, "Localization in Sensor Networks", in Handbook of Sensor Networks: Algorithms and Architectures, John Wiley & Sons, Inc., 23 Sept 2005

[26] S. Schlupkothen and G. Ascheid, "Localization of wireless sensor networks with concurrently used identification sequences", in *IEEE Ad Hoc Networking Workshop* (MED-HOC-NET), 2015

[27] "USAC: A universal framework for random sampling consensus", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 2022-2038, Volume 35, Issue 8, 2013

[28] R. Kuemmerle *et al.*, "g2o: A General Framework for Graph Optimization". In *IEEE*

Int. Conf. Robotics and Automation, 2011.

[29] S. Schlupkothen *et al.*, "A Novel Low-Complexity Numerical Localization Method for Dynamic Wireless Sensor Networks", IEEE T SIGNAL PROCES, vol. 63, no. 15, AUGUST 1, 2015

[30] G. Han *et al.*, "Localization algorithms of Wireless Sensor Networks: a survey", *TELECOMMUN SYSTS*, 52(4):2419–2436, August 2011.

[31] M. Erol-kantarci and H. T. Mouftah, "Localization Techniques for Underwater Acoustic Sensor Networks", *IEEE COMMUN MAG*, (December):152–158, 2010.

[32] J. A. Rothermich *et al.*, "Distributed Localization and Mapping with a Robotic Swarm.", *LECT NOTES COMPUT SC*, Springer Berlin Heidelberg, 2005.

[33] J. McLurkin and J. Smith, "Distributed Algorithms for Dispersion in Indoor Environments Using a Swarm of Autonomous Mobile Robots", *Distributed Autonomous Robotic Systems 6*, Springer Japan, p.399-408, 2007.

[34] D. Dardari *et al.*, "Ranging With Ultrawide Bandwidth Signals in Multipath Environments", P IEEE, vol. 97, 2009.

[35] E. Duisterwinkel *et al.*, "Environment mapping and localization with an uncontrolled swarm of ultrasound sensor motes", in *Proc. of Meetings on Acoustics*, 20(1), 2014.

[36] Z.D. Deng, et al, "Design and implementation of a new autonomous sensor fish to support advanced hydropower development", REVIEW OF SCIENTIFIC INSTRUMENTS 85, 115001 (2014)

2.9   APPENDIX: Algorithm: overview

Data: $d_{\bar{i},\mathrm{CID}(j)}$, $\mathrm{CID}(\{1,...,N\})$
Result: graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
Consistency + hypothesise mutual connections $\mathrm{h}_{i,j}$ using $\mathcal{C}$ (Sec. 2.4.1);
Create stable seed from 4 nodes (Sec. 2.4.2);
while *unreconstructed nodes left* **do**
    select all $c \in \mathcal{C}$ that observe ≥ 4 nodes in $\mathcal{V}$;
    sort c according to most likelihood to succeed (Sec. 2.4.3.1);
    perform RANSAC on c (Sec. 2.4.3 & Algo. 2);
    if *nodes c with sufficient RANSAC-score* **then**
        add nodes with highest RANSAC-score to $\mathcal{G}$;
    else if *node added since last non-linear refinement* **then**
        run non-linear refinement (Sec. 2.4.4);
    else
        abort algorithm;
    end
    if m *new nodes have been added to* $\mathcal{G}$ **then**
        run non-linear refinement (Sec. 2.4.4);
    end
    Loop closing (Sec. 2.4.5)
end

Algorithm 1: reconstruction

2.10   APPENDIX: Algorithm: RANSAC

Data: sorted c', $h_{i,j}$, $\mathcal{Q}$
Result: RANSAC-score and proposed positions of c' in $\mathcal{G}$
while *no good enough RANSAC-score* **do**
    select (next) largest clique in $\mathcal{Q}$ with connected candidate b;
    select all possible sets of 3 nodes $\{a : a \in \mathcal{Q}(b)\}$;
    for *all a* **do**
        for *all a_p = possible (3-)sets of distances in $h_{a,b}$* **do**
            hypothesise position for b, based on lateration;
            if *hypothesised position* $\notin \mathbb{R}^3$ **then**
                continue;
            end
            select all voters $\{a_v : (a_v \cup a_p) = a, (a_v \cap a_p) = \varnothing\}$;
            voters support hypothesis when connection $h_{a_v,b}$
              is consistent with hypothesised position;
            voters oppose otherwise;
            if *# of supporters > # of opposers* **then**
                RANSAC-score hypothesis is # of supporters;
                if *RANSAC-score is good enough* **then**
                    break **while**-loop;
                end
            end
        end
    end
end

Algorithm 2: RANSAC

2.11   APPENDIX: Algorithm: guiding RANSAC

Data: c, $h_{i,j}$, $\mathcal{V}$
Result: sorted c' and cliques $\mathcal{Q}(c)$
for *every b in c* **do**

>   select all nodes $a \in \mathcal{V}$ with existing $h_{a,b}$;
>   create graph $\mathcal{G}'(b) = (\mathcal{V}(a), \mathcal{E}')$ in which all vertices $V(a)$ are
>   connected with an edge $\mathcal{E}'$ if their positions are within $2r_{\text{comm}}$
>   of each other;
>   find cliques $\mathcal{Q}(b)$ in $\mathcal{G}'(b)$;

end
sort c according to clique size of largest cliques in $\mathcal{Q}(c)$;

Algorithm 3: Guiding RANSAC by sorting the candidate nodes

3

Mapping Swarms of Resource-Limited Sensor nodes: Solely Using Distance Measurements and Non-Unique Identifiers

authors:
*Erik H.A. Duisterwinkel, Gijs Dubbelman, Libertario Demi,
Elena Talnishnikh, Jan W.M. Bergmans and Heinrich J. Wörtche*

This work is on 3-D localization of sensor nodes in massive swarms based solely on 1-D relative distance measurements between neighbouring nodes. We target applications in remote and difficult-to-access environments such as the exploration and mapping of the interior of oil reservoirs where hundreds or thousands of nodes are used. These applications bring forward the need to use highly miniaturized sensor nodes of less than 1 centimeter, thereby significantly limiting measurement and processing capabilities. These constraints, in combination with additional limitations posed by the environments, impede the communication of unique hardware identifiers, as well as communication with external, fixed beacons.

We propose solving this challenging localization task by a novel RANSAC algorithm that can cope with noisy 1-D relative distance measurements and non-unique communication identifiers. It uses local geometric consistency, to resolve the ambiguity caused by non-unique communication identifiers and outlier measurements, and thereby is able to robustly assign unique hardware identifiers to be used for global non-linear graph optimization.

Extensive simulations show that this novel localization method is able to fully reconstruct the positions of the nodes, in cases when the number of communication identifiers is only 2% of the number of nodes. When the number of communication identifiers is lower, the algorithm exhibits graceful degradation.

3.1   Introduction

The exploration of remote, deep underground or difficult-to-access environments has been subject of study for decades. Current challenges include the mapping and exploration of the interior of narrow subterranean cavities like (oil) reservoirs and piping systems, which have high economic value and societal importance [1], and the investigation of the dynamics in industrial (multi-fluid) mixing tanks. The use of many (e.g. thousands), small and cheap mass-produced micro-sensor systems – from now on referred to as *sensor nodes* which are operated in a *swarm*–is a promising way to approach these challenges. The specific sensor nodes for mapping have many more promising applications, but are currently not available yet. The development for the larger sized sensor nodes (5 cm diameter) for e.g. the mapping and exploration of a mixing tank is under way [2]. Many hardware challenges still need to be addressed before large swarms of miniaturized sensor nodes (centimeter-sized or smaller) can be developed and deployed effectively. Trade-offs must be made in the hardware design with respect to node capabilities and node size. Improved node capabilities, e.g. larger sensing radius or better signal-to-noise ratios, inevitably make cost-effective miniaturization of nodes more challenging. Therefore, in our study, we set the challenge to perform reliable node localization using severely limited node capabilities. Although, these limitations might not be relevant to current nodes, we believe they are relevant for yet-to-be-developed highly miniaturized nodes of the near future [3]. We show that advances in off-line localization algorithms, can, to a large extent, compensate for severely limited node capabilities. This is important as it implies that using large swarms of highly miniaturized nodes in the future, is a realistic scenario. At the same time, our research provides lower bounds on node capabilities that can guide the hardware design of yet-to-be-developed nodes.

Each type of environment that is to be explored brings forward a specific set of constraints; in the application cases we are considering, the environments prevent the use of large, complex or existing sensing and localization systems, due to various issues, discussed in Sec. 3.1.1. Exploration can in that case be performed by injecting swarms of sensor nodes into the flooded environment, as e.g. illustrated in Fig. 3.1. The nodes traverse the environment due to the internal dynamics (e.g. flow) and disperse over the volume of interest. The nodes are extracted from the environment and the measured data that is stored in their internal memory can be analysed offline. The shape, or *map*, of the sensor swarm can be used to infer the geometric structure and size of the environment. The individual node positions relative to each other – from which we can determine such a map– can be estimated from inter-more distance measurements [4], [5]. As illustrated in Fig. 3.2, estimating the positions of the nodes can be seen as a specific kind of graph problem, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertices $\mathcal{V}$ are the node positions and the edges $\mathcal{E}$ the distances between them. The challenge is to obtain a robust initial estimate of the node positions in order for non-linear graph optimization algorithms to succeed [6]. In contrast to earlier work [7] where a similar problem has been considered, in our work [8], an initial estimate is obtained using a novel robust Random Sampling Consensus (RANSAC) algorithm [9], [10], applied to general lateration techniques.

In this work, we explore the limits of swarm-based 3-D localization of extremely resource limited sensor nodes with the specific set of constraints posed by the application and the environment. Our work differs from previous works [7], [9], [10], [12], [13], [14], [15], [16], [17] by,
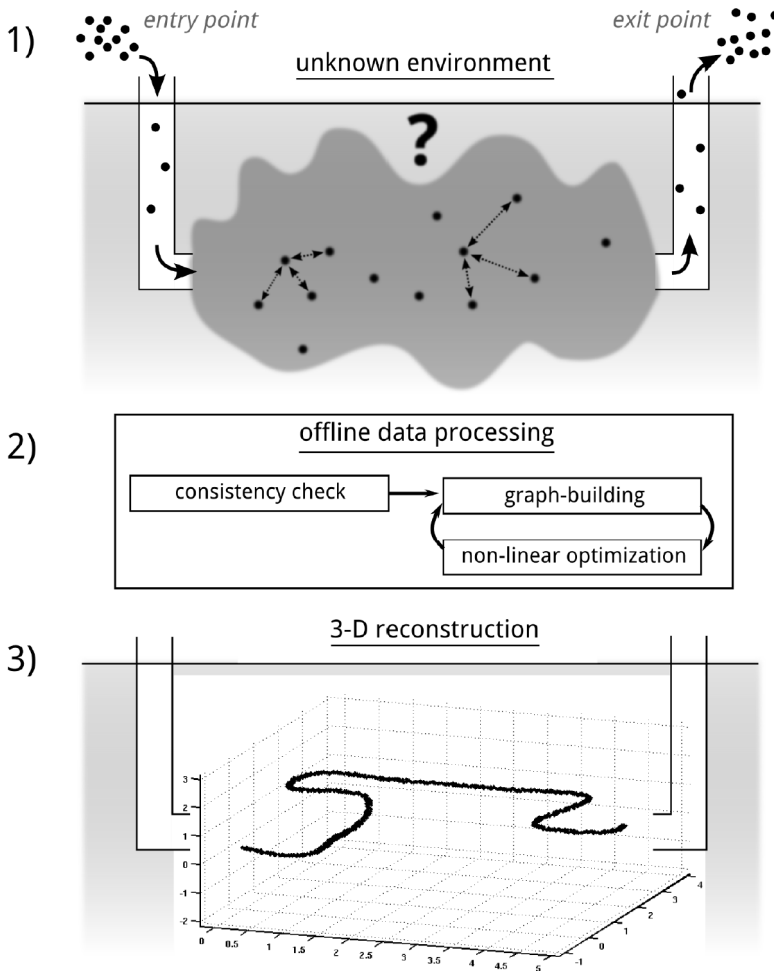
Figure 3.1: Swarms of sensor nodes are inserted in the environment of interest; they traverse the environment using its internal dynamics (e.g. flow ; and are extracted in order to retrieve the data. Once the nodes are distributed in the environment, distance measurements between neighbouring sensor nodes are taken. Offline, this data can be processed to determine the nodes positions relative to each other.
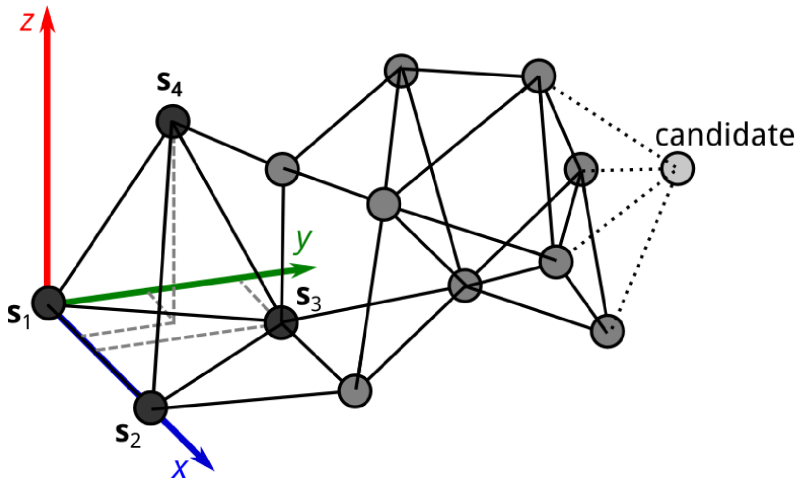
Figure 3.2: Localizing nodes within a swarm as a graph problem, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with the node positions as vertices $\mathcal{V}$ and the distances between nodes as edges $\mathcal{E}$. The initial four node positions $\{s_1, s_2, s_3, s_4\}$ are chosen to define the coordinate system to solve the global reflection and rotation ambiguity. Additional nodes can be added to the graph using general lateration techniques. For clarity of the figure, not all (required) edges are drawn.

1. using solely distance measurements (also called *range-only*), i.e. no direction information or additional sensor information, like bearing, odometry or inertia;

2. the distance measurements are performed using non-unique identifiers;

3. no additional data is exchanged, measurements are stored in memory for offline analysis;

4. not depending on external communication to e.g. fixed beacons; and,

5. sparse connectivity in a large swarm.

Although each separate constraint has been considered earlier; this work is, to the best of our knowledge, the first to show feasibility of reconstructing sensor swarms when considering all mentioned constraints and under realistic conditions. The main difference with our previous work in [8], is using non-unique identifiers (constraint 2) which has far-reaching consequences on the localization [11]. In [12], fully anonymous measurements are considered for robot localization, but cannot be compared to our problem, as the localization algorithm relies on the significantly larger sensing, online processing and communication capabilities of the robots. A similar complete problem as ours is considered in [18], [19] where a different solution is proposed to resolve the ambiguities. In both papers, only Gaussian noise is considered and the effect of ambiguities and noise on the final localization is not shown.

This paper describes in Sec. 3.2 the exact problem that is considered. In Sec. 3.3 we elaborate on the localization algorithm. Sec. 3.4 describes the simulation model we use to test the novel mapping method and localization algorithm in pipe-like environments. In Sec. 3.5, we numerically evaluate the conditions/criteria for the localization algorithm to reliably esti-

mate the node positions in the swarm, while reducing the number of available communication identifiers. Our conclusions are provided in Sec. 3.6.

3.1.1   Constraints on sensing devices

In our application cases, we consider environments that are enclosed and/or deep underground and might have a narrow passage. Examples of these are the sand-free channels in heavy oil containing sandstone formations which are formed during oil extraction (Cold Heavy Oil Production using Sand, CHOPS) [1]. These sand-free channels, also called *wormholes*, are typically around one to several centimeters in diameter; extend over a few hundred meters; and are located 200-400 meter underground. Placing beacons is infeasible since the structure deep underground spans large distances and is not known beforehand. Communication using radio is not feasible due to the extreme high salinity. Acoustic communication however, is possible within the wormholes for communication between neighbouring sensor nodes.

Passage through a wormhole is only possible if the physical dimensions of the nodes are smaller than that of the local structure. For this specific application case it limits the sensor size to less than 1 centimeter in diameter, causing severe limitations on the instrumentation which can be taken aboard of the node. At these scales, the major limitation is the energy storage, which limits the communication range, data rate and packet sizes, and the amount of processing that can be performed on-board. For these reasons we are looking into 1-D relative distance-only measurements which can be performed using ultrasound time of arrival or time of flight. This can for example be achieved using highly miniaturizable ultrasound technology, e.g. [20].

Identifiers can be encoded into the ranging pulse using e.g. BPSK or CDMA. But as ultrasound transducers are generally narrow-banded, data encoded in signal pulses causes signal pulses to be long [21], [20]. As these environments are small and reflective and the speed of sound is five orders of magnitude lower than that of radio waves, signal overlap due to emission of signals by neighbouring nodes and their multi-paths is expected to be significant [21]. Elaborate communication techniques to deal with this, require more complex processing [22], which might not be feasible on the miniaturized nodes. In order to reduce multi-path in these environments, we attempt to limit the length of the ranging pulse, and therefore, have to limit the quantity of identification information encoded in it. The nodes unique hardware identifier (uiD) can therefore not be encoded in the ranging-pulse, but only highly abbreviated and non-unique communication identifiers (CiD) can be communicated. Consequently, received ranging-pulses from neighbouring nodes can only be identified up to a large ambiguity.

3.2   Problem description

Following the operational procedure shown in Fig. 3.1, once a 'steady-state' distribution of the nodes in the volume of interest is achieved, the nodes perform simultaneously distance measurements to neighbouring nodes. This can be achieved using e.g. a preset time after insertion. The ranging measurements are performed using ultrasound time-of-flight (TOF) or time-of-arrival (TOA). Omnidirectional ultrasound emission and reception is considered (as opposed to directional), because there is no a priori or online knowledge of the positions of

neighbouring node. Each node reaches all neighbouring nodes that are in line-of-sight, and within a specific sensing radius r_s.

We assume a homogeneous speed of sound, and that the movement of the nodes during the distance measurements can be ignored, i.e. flow velocity of fluid is small compared to speed of sound in fluid. For now, the nodes are also assumed to be synchronized such that obtaining the distances using TOA or TOF is straightforward. However, achieving this synchronization under the stringent communication constraints in the real application is not straightforward, but believed to be possible [3], [4] ; it is a topic for future research.

3.2.1 Reconstructing node positions

A general method of reconstructing the positions of the sensor nodes based solely on distance measurements relies on general lateration methods. The graph $\mathcal{G}$ can be grown as illustrated in Fig. 3.2 and detailed in [8]. Every candidate node, c, with known distances to four non-coplanar nodes (of which the positions s_1, s_2, s_3, s_4 are known) can be added to the graph. The position of node c will be the intersection of the spheres with radii $\overline{d}_{1,c}, \overline{d}_{2,c}, \overline{d}_{3,c}, \overline{d}_{4,c}$ and centers at s_1, s_2, s_3, s_4. Here $\overline{d}_{n,c}$ denotes the true distance from node $n = \{1, 2, 3, 4\}$ to node c. The initial four node positions are chosen such that they define the coordinate system, therewith fixing the general reflection and rotation

ambiguity: $s_1 \in \{0, 0, 0\}, s_2 \in \{\mathbb{R}, 0, 0\}, s3 \in \{\mathbb{R}, \mathbb{R}^+, 0\}$ and $s_4 \in \{\mathbb{R}, \mathbb{R}, \mathbb{R}^+\}$.

The ranging measurements are performed in twofold, node i measuring distance to node $j, d_{\overline{i},\overline{j}}$, and vice versa, $d_{\overline{j},\overline{i}}$. In the data processing, this allows for the possibility to perform a forward-backward consistency check in determining the distance between the nodes.

3.2.2 Identity and distance ambiguity

The nodes are assigned at random one of the n_f available CIDs; all CIDs are distributed uniformly among the nodes. The nodes only emit this non-unique CID in the ranging pulse, rather than a unique UID. Consequently, the measured data per node i will consist of a set of measured distances $d_{\overline{i},\text{CID}(j)}$ for all neighbouring nodes j within its sensing radius r_s. These distance measurements are not uniquely associated with neighbouring nodes j, but rather only to nodes with $\text{CID}(j)$. Therefore, the forward-backward consistency check, cannot be used to *uniquely define* a hypothesised distance between node pairs, as in [8], but leaves ambiguities.

In order to build the graph, the non-unique distance measurement, $d_{\overline{i},\text{CID}(j1)}$, should be associated with the uniquely identified nodes i and j. This can be attempted by first considering all possible arrangements that are consistent with current measurements; similar as in [12].

With perfect distance measurements, the challenge to associate measurement $d_{\overline{i},\text{CID}(j)}$ with node j and $d_{\overline{j},\text{CID}(i)}$ with node i is straightforward. Searching for all possible arrangements that are consistent with the observation, i.e., $d_{\overline{i},\text{CID}(j)} = d_{\overline{j},\text{CID}(i)}$ will give most of the times an unambiguous correct result, since it is unlikely that another set of nodes with similar CID pair have exactly the same distance. However, distance measurements are subject to noise and result in an erroneous distance determination. The searching condition $\mathcal{F}$ for
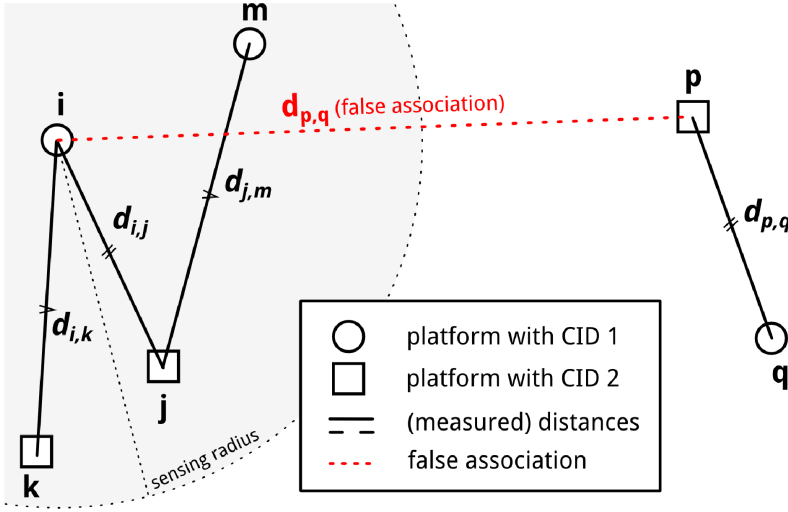
Figure 3.3: Platforms with UID $= \{i,\ j,\ k,\ m, p,\ q\}$ using only two distinct CIDs (indicated by O and □) cause ambiguities in associating UID with measured distances. Identity ambiguity arises due to *similar* distance between pairs with similar CID within the swarm (Eq. 3.2). Distance ambiguities arise due to a plurality of nodes with similar CIDs within a sensing radius.

associating nodes with measurements should be extended to

$$\mathcal{F}(d_{a,b},\ d_{c,d}) = \begin{cases} \frac{d_{a,b}+d_{c,d}}{2} & \text{when } d_{a,b} \simeq d_{c,d} \\ \varnothing & \text{otherwise} \end{cases} \tag{3.1}$$

where $d_{a,b} \simeq d_{c,d}$ means $d_{a,b}$ and $d_{c,d}$ are *similar* to each other. We can define this similarity by the condition

$$d_{a,b} \simeq d_{c,d} \text{ when } \begin{cases} |d_{a,b} - d_{c,d}| \leq \epsilon_r & \text{and} \\ \text{CID}(a) = \text{CID}(d) & \text{and} \\ \text{CID}(b) = \text{CID}(c) \end{cases} \tag{3.2}$$

where ϵ_r is set as threshold value for maximum allowed difference in distance measurement for which two measurements are considered inliers. This threshold should be related to the noise. It can e.g. be obtained using trial and error if the error model is unknown.

The obtained hypothesised distance between node $\bar{i}$ and node j is not unambiguous. As also illustrated in Fig. 3.3, the hypothesised distances and associations $h_{i,j}$ for measurement $d_{\bar{i},\text{CID}(j)}$ can consist of three types of contributions:

$$h_{i,j} = \begin{cases} \mathcal{F}(d_{\bar{i},\bar{j}},\ d_{\bar{j},\bar{i}}) & (a) \\ \mathcal{F}(d_{\bar{i},j},\ d_{p,q}) & (b) \\ \mathcal{F}(d_{\bar{i},k},\ d_{j,m}) & (c) \end{cases} \tag{3.3}$$

a) the measurement between the real nodes $\bar{i}$ and $\bar{j}$;

b) *identity ambiguity*: an association to a node p which is believed to be at similar distance as node $\bar{j}$ but not necessarily within sensing range of $\bar{i}$ (see Fig. 3.3);

c) *distance ambiguity*: an association with the correct node $\bar{j}$ but with a distance belonging to the measured distance to another node k within its sensing radius (see Fig. 3.3).

The identity ambiguities are based on the statistical likelihood that somewhere in the swarm, a node pair with similar CIDs and distance is present. Assuming uniform distribution of the nodes, the average amount of identity ambiguities in $h_{i,j}$ scales with $\mathbb{A}_I \propto \epsilon_r n_s N / n_f^2$, in which n_s is the average amount of neighbouring nodes within the sensing radius r_s and N the total amount of nodes in the swarm. Equally, the distance ambiguities are based on the statistical likelihood that neighbouring nodes give rise to confusion. The average amount of distance ambiguities per $d_{\bar{i},\mathrm{CID}(j)}$ then scales with $\mathbb{A}_D \propto \epsilon_r n_s^2 / n_f^2$.

Both type of ambiguities, $\mathbb{A}_I$ and $\mathbb{A}_D$, increase with an increasing amount of neighbouring platforms n_s, leading to a more challenging task associating measurements with the correct nodes. However, our RANSAC algorithm, which is described in the next section, exploits the fact that measurements of true neighbouring nodes can be used to resolve ambiguities by providing geometric consistency checks. An increase in ambiguities can be compensated for by a larger amount of consistent measurements from neighbouring nodes.

3.3   Swarm localization algorithm

From the ambiguous set of hypothesised distances and associations $h_{i,j}$ in Equation 3.3, the graph $\mathcal{G}$ should be robustly built. We use a similar approach as in [8] but made significant modifications to prune false entries in $h_{i,j}$.

The main steps in our localization method are described next in Algo. 4 and Sec. 3.3.1 through Sec. 3.3.4.

3.3.1   Initial seed selection

Four neighbouring nodes are selected as initial seed to start the graph growing. Their positions, s_1, s_2, s_3, s_4, define the coordinate system as described in Sec. 3.2.1. The nodes are selected based on their connectivity and the stability of their geometric configuration which is obtained by general lateration techniques. Since $h_{i,j}$ contains ambiguities, the initial seed selection is conditional until the graph growing has successfully added several nodes to the graph $\mathcal{G}$. If this is not possible, a new initial seed is selected and the process is repeated.

3.3.2   RANSAC graph growing

Additional nodes can be added to the graph when the nodes have at least four connections to already reconstructed nodes. This set of candidate nodes is denoted as C. Figure 4 illustrates a simplified 2-D situation in which a candidate node $c \in C$ has associations to its true neighbouring nodes in $\mathcal{V}$, but also false associations to nodes somewhere else in the graph due to identity ambiguities.
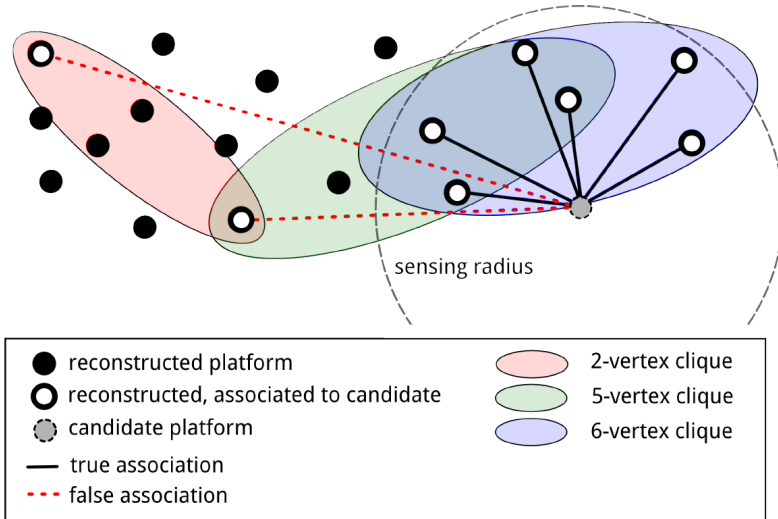
Figure 3.4: RANSAC graph growing algorithm uses an inlier-outlier voting system to filter out outlier distance measurements and proposes a position for candidate nodes to add to the graph. Using cliques of nodes in the graph that are within twice the sensing radius helps in reducing the chance nodes are positioned wrongly.

A RANSAC approach can be used to attempt to correctly position the candidate node based on true associations only. The inlier-outlier voting mechanism is similar as in [8] but is adjusted to handle ambiguities. This is described in Algo. 5. When a candidate node c is considered, all nodes that are already reconstructed in the graph and that are associated with this c are selected, $A = \{a : a \in \mathcal{V}, h_{a,c} \neq \varnothing\}$. A set of three of these nodes is selected, $\{a_p \in A\}$, to *propose* –using general lateration– a position for c up to a reflection ambiguity. But since no guarantees can be given whether the entries in $h_{a_p,c}$ are true distances or true associations, the proposed positions do not need to be viable or close to the true position. Therefore, each of the other associated nodes $\{a_v : (a_v \cup a_p) = A, (a_v \cap a_p) = \varnothing\}$ are used to *vote* for the proposed position. Consensus is reached when the majority of the voting nodes, including the three proposing nodes, have at least a 50% majority. The reflection ambiguity is chosen based on which of the two positions received more supporting votes. When consensus is reached, the amount of supporting votes is called the *RANSAC-score* for the specific proposal.

This step in the RANSAC algorithm is repeated with each time a different set of three proposers a_p, until all possibilities are exhausted or until a proposal received a specific threshold in RANSAC-score. In the latter case, it is then considered *good enough* for addition in the graph. When no 'good enough' condition is reached, the RANSAC procedure is repeated for a next candidate node until all possibilities are exhausted or the 'good enough' condition is found. The candidate with the highest RANSAC-score is added to the graph in $\mathcal{V}$ with edges only to the supporting nodes and corresponding $h_{a,c}$ entry in $\mathcal{E}$. The 'good enough' condition for the RANSAC is chosen heuristically.

Data: $d_{\bar{i},\text{CID}(j)}$,CID($\{1,\ldots,N\}$)
Result: graph $\mathcal{G} = (\mathcal{V},\ \mathcal{E})$
Hypothesise distances and associations $\text{h}_{i,j}$ using $\mathcal{F}$ (Eq. 3.1& 3.3);
Create stable seed from 4 nodes (Sec. 3.3.1);
while *unreconstructed nodes left* **do**
    select all $c \in \mathcal{C}$ that observe ≥ 4 nodes in $\mathcal{V}$;
    sort c according to most likelihood to succeed (Sec. 3.3.3);
    perform RANSAC on c (Sec. 3.3.2 & Algo. 5);
    if *nodes c with sufficient RANSAC-score* **then**
    | add nodes with highest RANSAC-score to $\mathcal{G}$;
    else if *node added since last non-linear refinement* **then**
    | run non-linear refinement (Sec. 3.3.4);
    else
    | abort algorithm;
    end
    if *m new nodes have been added to $\mathcal{G}$* **then**
    | run non-linear refinement (Sec. 3.3.4);
    end
end

Algorithm 4: Localization

Data: sorted c^*, $h_{i,j}$, $\mathcal{Q}$
Result: RANSAC-score and proposed positions of c^* in $\mathcal{G}$
while *no good enough RANSAC-score* **do**
    select (next) largest clique in $\mathcal{Q}$ with associated candidate b;
    select all possible sets of 3 nodes $\{a : a \in \mathcal{Q}(b)\}$;
    for *all a* **do**
        for *all a_p =possible (3-)sets of distances in $h_{a,b}$* **do**
            hypothesise position for b, based on lateration;
            if *hypothesises position $\notin \mathbb{R}^3$* **then**
            | continue;
            end
            select all voters $\{a_v : (a_v \cup a_p) = a, (a_v \cap a_p) = \varnothing\}$;
            voters support hypothesis when observation $h_{a_v,b}$ is consistent with
              hypothesised position;
            voters oppose otherwise;
            if *# of supporters ¿ # of opposers* **then**
                RANSAC-score hypothesis is # of supporters;
                if *RANSAC-score is good enough* **then**
                | break **while**-loop
                end
            end
        end
    end
end

Algorithm 5: RANSAC

Data: c, $h_{i,j}$, $\mathcal{V}$

Result: sorted c^* and cliques $\mathcal{Q}(c)$

for *every b in c* **do**

    select all nodes $a \in \mathcal{V}$ with existing $h_{a,b}$;

    create graph $\mathcal{G}^*(b) = (\mathcal{V}(a), \mathcal{E}^*)$ in which all vertices $\mathcal{V}(a)$ are connected with
     an edge $\mathcal{E}^*$ if their positions are within $2r_s$ of each other;

    find cliques $\mathcal{Q}(b)$ in $\mathcal{G}^*(b)$;

end

sort c according to clique size of largest cliques in $\mathcal{Q}(c)$;

Algorithm 6: Sorting of candidate nodes

3.3.3 Sorting candidate nodes

The amount of false associations from candidate nodes to the nodes in the already recon-structed graph grows on average linearly with the amount of nodes in the graph as explained in Sec. 3.2.2. As a consequence, candidate nodes might get voted to a wrong position in the graph during the RANSAC process in cases where the true neighbouring nodes are outnum-bered.

To reduce the chance of this to happen, a sorting order of candidate nodes is made before the RANSAC algorithm is performed on them. The sorting is based on the likelihood that candidate nodes have enough true neighbouring nodes already in the graph. As such, candi-date nodes that are more prone to be positioned based on false connections, are considered later. This gives the chance that more of its true neighbouring nodes will be added to the graph first. The algorithm is stated in Algo. 6 and described next.

For each candidate node $\{b \ : \ b \ \in \ C\}$, a list is made of associated nodes that have already been positioned in the graph $A = \{a \ : \ a \ \in \ \mathcal{V}, \ h_{a,b} \ \neq \ \varnothing\}$. A new graph $\mathcal{G}^* = (\mathcal{V}(A), \ \mathcal{E}^*)$ is made with these nodes as vertices. These vertices are connected with edges $\mathcal{E}^*$ when their positions are less than twice the sensing radius, $2r_s$, apart. In this new graph, maximal cliques are listed as $\mathcal{Q}(b)$. These cliques indicate the groups of nodes that are within the sensing radius of a proposed candidate node that can potentially be agreed on by all of these nodes. The example in Fig. 3.4 shows these cliques in coloured oval areas. The amounts of nodes in these cliques are registered and are used as sorting order in which RANSAC is performed. Platforms with larger sized cliques in $\mathcal{Q}(b)$ are more likely to have more associations with true neighbouring nodes.

3.3.4 Robust non-linear refinement

The stepwise addition of new nodes to the graph introduces build-up of errors. These errors in positions can prevent other nodes from being added. In order to reduce this error build-up, a global non-linear optimizer algorithm is performed [6], [8], after every m newly added nodes, or failure to add a new node.

There is a clear trade-off between the size of m and the processing time required for the localization algorithm. In our work, m is chosen arbitrarily to 10. In future work, this can e.g. be adjusted dynamically based on uncertainty of previously added nodes.

3.4   Simulation setup

The swarm-based distance-only localization can be used in a variety of environments. In this paper we will be using long, pipe-like structures as an example. The diameter of the

pipe is large enough such that nodes can be positioned all around one another instead of in one line along the pipe axis. Simulations are performed 4-7 times, each time with a different pipe-structure. The course path of the loop-less, non-overlapping pipe without branches is randomly chosen with incremental steps of 1 meter over a single $\{x, y, z\}-$axis. The pipe axis is smoothed out by drawing a spline through these points. The pipe diameter is chosen to be fixed at 8 cm and there are 100 nodes in every meter of pipe-length. The positions of the nodes are assumed to be uniformly distributed within the pipe; r_s is chosen such that each node has on average n_s neighbouring nodes. Parameter n_s is one of the parameters that will be swept. The nodes sensing radius r_s, in which ranging measurements are possible, is for all nodes the same.

In reality, homogeneous distribution of node positions is difficult to achieve, but allows for thorough study of all other parameters involved.

3.4.1   Measurement noise

The measured distance between nodes will be affected by different imperfections in the system. The distance measurement errors are modelled similar to what we have described in [8], which means we add additive Gaussian noise to account for inaccuracies in the timing of TOF and multiplicative Gaussian noise to account for imperfect knowledge of the speed of the (ultrasound) signal through the medium. Let $\overline{d}$ be the true distance between a set of nodes, then the measured distance d is modelled by

$$d_{\overline{i},\overline{j}} = \overline{d}_{\overline{i},\overline{j}} + \overline{d}_{\overline{i},\overline{j}}\mathcal{N}_{m} + \mathcal{N}_{a} + B\mathcal{U} \qquad (3.4)$$

with $\mathcal{N}_{m}$ and $\mathcal{N}_{a}$ being perturbations from the zero-mean Gaussian distributions $\mathcal{N}(0, \sigma_{m}^{2})$ and $\mathcal{N}(0, \sigma_{a}^{2})$ respectively. Outlier noise is added to the measurement with the term $B\mathcal{U}$ in which $\mathcal{U}$ is a perturbation drawn from a uniform distribution in the range $[-r_s, \ r_s]$ and $B \in \{0, 1\}$. The chance that a measurement has outlier noise, i.e. $B = 1$, is determined by a Bernoulli distribution such that on average a percentage ω is an outlier. Any measurement that lies outside of the sensing range $(0, \ r_s]$ is discarded from further processing.

Identification noise, that accounts for the possible errors in determining the senders CiD, is included by assigning a CID that is drawn randomly from the available set of CIDs to a measurement. The chance that a measurement has identification noise is determined by the ratio ϕ.

Initially, in Sec. 3.5, we consider no outlier and identification noise, i.e. $\omega = 0\%$ and $\phi = 0\%$, and we consider three Gaussian noise levels: $\sigma_{m} = \{0.30, 1.0, 3.0\} \cdot 10^{-3}$ and $\sigma_{a} = \{0.36, 1.3, 3.6\} \cdot 10^{-3}m$. This yields an average standard deviation of the distance measurement noise of $\nu \approx \{0.4\%, 1.33\%, 4\%\}$ expressed in percentage of the average sensing radius when $n_s = 20$.

In Sec. 3.5.1 we investigate the robustness against additional outlier and identification noise with noise levels ω and ϕ ranging between 0% and 20%.

3.4.2   Performance parameters

To analyse the performance of the localization algorithm, we investigate the parameter *recall*, $\mathbb{X}$, which is the percentage of nodes reconstructed by the localization algorithm. The recall parameter can only be interpreted correctly when also considering the *accuracy*. The accuracy is described by both the relative error and the absolute error of the node positions. The relative error gives the error of the distance between two neighbouring nodes, expressed as $\hat{d}/\overline{d}$, where $\hat{d}$ is the reconstructed distance between nodes and $\overline{d}$ the true distance between them. The absolute error gives the error between the reconstructed position and the true position of the nodes: $|\hat{s} - \overline{s}|$. Since the reconstruction is performed based on relative position only, i.e. relative to the initial four nodes that determine the coordinate system, the absolute error is only useful when the reconstructed swarm is aligned with its ground truth. In order to solve this ambiguity, a linear fit between the 3-D positions of the initial four nodes and their true positions is performed.

These experiments are designed to evaluate conditions/criteria of the localization algorithm for reliable reconstruction of the sensor swarm with non-unique identifiers. Besides the pipe structure and the nodes positions, the parameters which play a role are the amount of nodes in use (N) , the amount of CID (n_f), the average amount of neighbouring nodes within the sensing radius (n_s) and the distance measurement noise levels (ν, ω, ϕ) .

A parameter set yields successful reconstructions when: a) the amount of added nodes, based mainly on false associations, is below 5% of $\mathbb{X}$; and, b) the local pipe structure (its diameter and its course axial direction) is reconstructed correctly. Build-up of small directional changes over the course of the pipe axis are allowed for a reconstruction to be defined successful. In this work, this is assessed rather subjectively. As addition on b), successful reconstruction is also achieved when there is at one point in the reconstructed swarm a large deviation in the course path of the pipe axis relative to its true direction, but the positions of the nodes after the this point are reconstructed correctly relative to the neighbouring nodes. An example of such a *line-break* is shown in Fig. 3.7. a.

3.5   Numerical simulations

The *recall* $\mathbb{X}$ results of the experiments without outlier and identification noise are illustrated in Figure 5. Indicated in green are the parameter sets which yield successful reconstructions, grouped in: $\mathbb{X} = 100\%$, $\mathbb{X} \geq 80\%$ and $\mathbb{X} \geq 20\%$. Unsuccessful reconstructions are indicated in white, grouped in: $\mathbb{X} \geq 50\%$ (cross) and $\mathbb{X} < 50\%$ (plain). The green hatched areas indicate that either a maximum of one iteration has a significant pipe axis line break or several iterations have a minor pipe axis line break.

It can be seen that at each N and noise level ν, there exists a minimum n_s and $n_f \ll N$ for which successful reconstruction can be achieved with 100% recall. Increasing n_f beyond this point does not significantly improve the recall or accuracy parameters.

Figure 6, shows the relative and absolute accuracy of the reconstructions with the specific
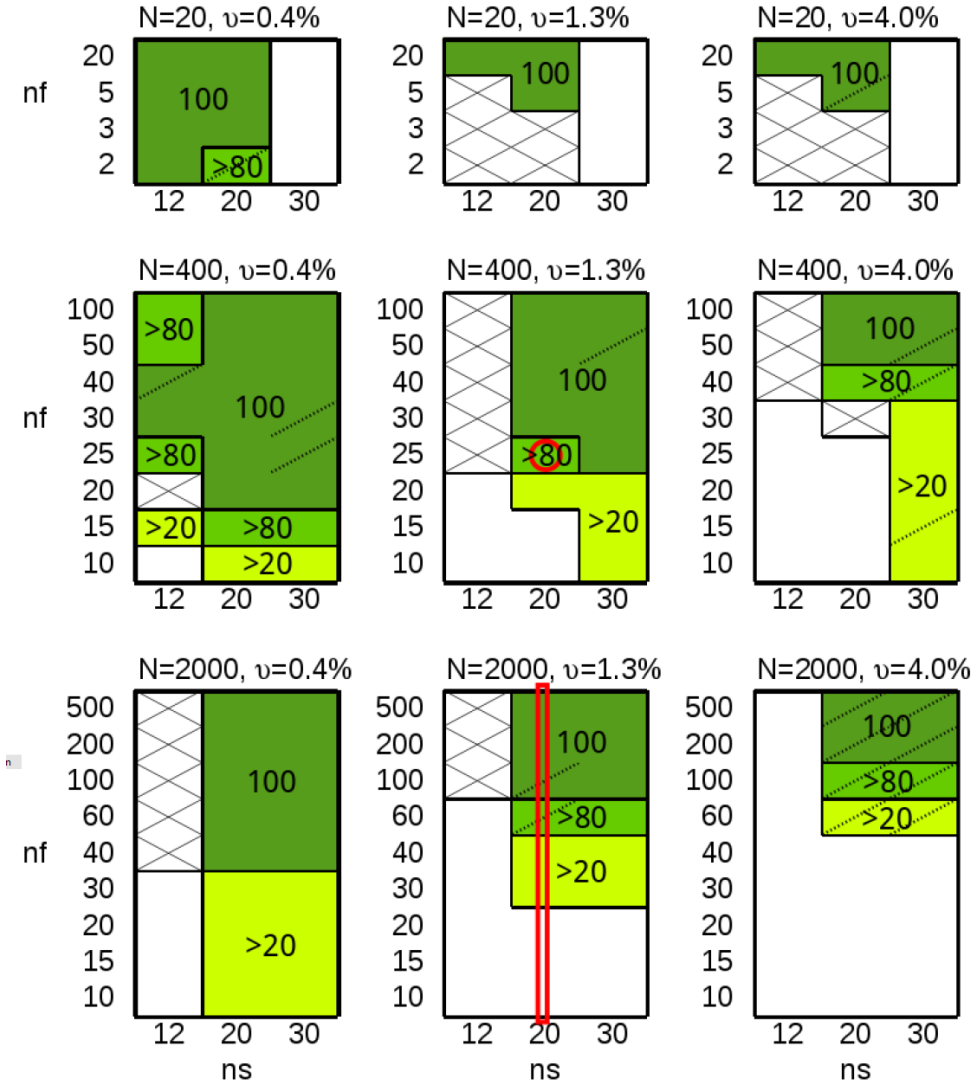
Figure 3.5: Reconstruction of sensor swarm from simulations with different parameter sets N, ν, n_f, n_s. Recall X in five categories: sufficient reconstruction: $\mathbb{X} \geq \{100\%, 80\%, 20\%\}$ (green ; unsuccessful reconstruction $\mathbb{X} \geq 50\%$ (white with cross) and $\mathbb{X} < 50\%$ (plain white). Red rectangle marks dataset used in Fig. 3.6, red cirle marks parameterset used in Sec. 3.5.1.
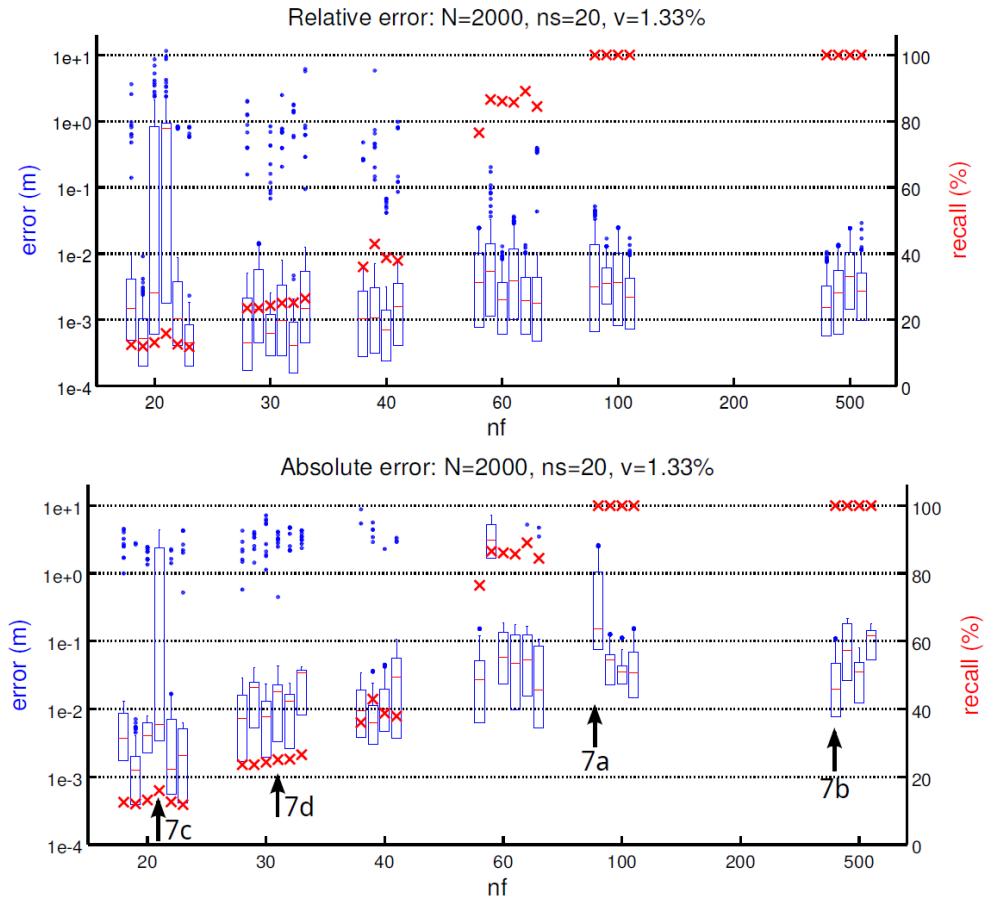
Figure 3.6: Recall (red ×) plotted together with relative error (top) and absolute error (bottom) for all iterations with specified parameter sets (marked in Fig. 3.5 with red rectangle). Reconstructions of four selected runs (arrows) are shown in Fig. 3.7.

set of parameters highlighted in Figure 5 with a red line. The errors are shown in boxplots for all reconstructed sensor nodes. Only the ten highest outliers are plotted ($> 1.5\times$ the interquartile range); the lowest quartile and lower outliers are not plotted. As a reference, the recall is also plotted.

The amount of neighbouring nodes n_s has a crucial role in whether a successful reconstruction can be achieved. In these experiments $n_s = \{12, 20, 30\}$ has been used and the recall results show that in most cases it is not sufficient to choose $n_s = 12$. In most cases, there is a significant chance that somewhere in the graph growing process, there are too little true neighbouring platforms already reconstructed for inlier votes in the RANSAC algorithm. A minimum of $n_s = 20$ is required for reliable reconstruction at larger N. This confirms our findings in [8]. Increasing n_s from 20 to 30, however, does not increase the recall significantly at these noise levels. Figure 7 shows four reconstruction of different iterations with the parameter set from Figure 6. Figure 7. a and 7. b show the reconstruction of pipe structure with equal node placement, both with a recall $\mathbb{X} = 100\%$. Figure 7. a has been made with $n_f = 100$ and shows a significant line break in the reconstructed pipe axis. Figure 7. b with $n_f = 500$ has significantly better reconstruction at this critical point in the swarm.

Figure 7. c and Figure 7. d show reconstructions in an equal pipe structure and node placement, but with a different n_f. The lower amount of n_f increase the amount of ambiguities introduced in the dataset, and due to the coherent nature of these ambiguities, RANSAC can in this particular case not distinguish between correct candidate nodes and false candidate nodes to add to the graph. The reconstruction in Figure 7. c has an average total ambiguity factor (both distance and identity) of $\mathbb{A} = 480\%$ versus $\mathbb{A} = 210\%$ in Figure 7. d.

The reconstructions with parameter set $N = 400, \nu = 4\%, n_s = 30, n_f = 15$ has a dataset with the largest average ambiguity factor for which the reconstruction is still considered successful. With an average ambiguity factor of $\mathbb{A} = 1700\%$ it still has a reliable recall of $\mathbb{X} = 20\%$. The reconstruction of these iterations was terminated due to a self- imposed time constraint. All of the partial reconstructions (with sufficient accuracy) with $n_s = 30$ as well as the highly ambiguous datasets with $n_s = 20$ have been terminated due to this time-constraint. Some other high ambiguous datasets with $\mathbb{A} > 2000\%$ have been preventively terminated due to the large amount of total entries in $h_{i,j}$ (larger than 2×10^5).

3.5.1 Non-Gaussian outlier and identification noise

The results described above are based on measurement datasets without non-Gaussian outlier and identification noise. But as these types of noise are inevitable in realistic scenarios, we also study the localization algorithm on datasets including outlier and identification noise. The unmodified algorithm is performed with ω and ϕ ranging from 0% to 20%. This is performed on the parameter set $N = 400, n_s = 20, n_f = 25, \mu = 1.33\%$ (marked in Fig 5).

with red circle). Recall results are shown in Table 3.1. Outlier and identification noise of up to 10% still yields successful reconstruction with $\mathbb{X} \simeq 100\%$ for the majority of the iterations.
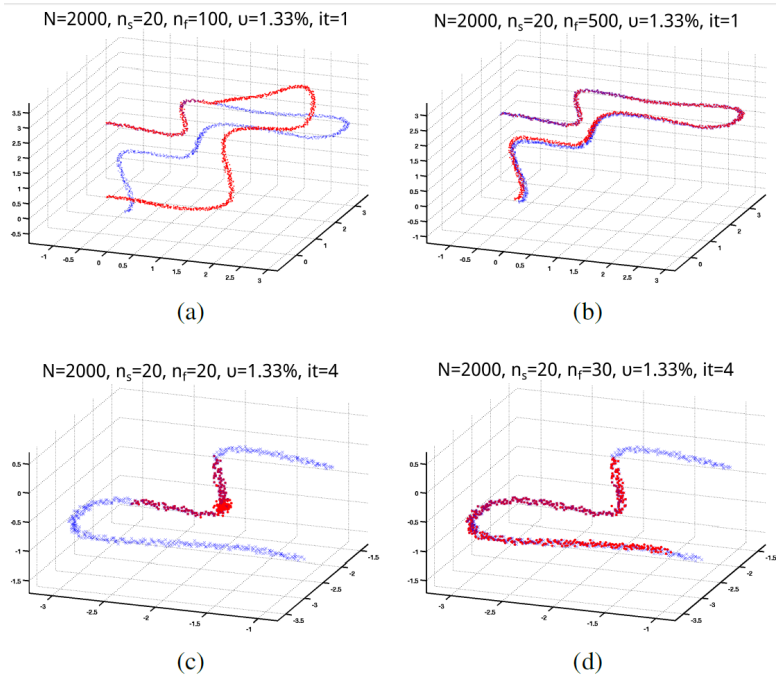
Figure 3.7: (Parts of) reconstructions of four different runs (indicated with arrows in Fig. 3.6). Blue are true positions, red are reconstructed positions.

Table 3.1: Recall $\mathbb{X}$ in percentage with outlier noise ω and identification noise ϕ of dataset $N = 400, n_s = 20, n_f = 25, \mu = 1.33\%$. The median of 10 iterations is taken; the lowest recall of the iterations is between brackets.

Outlier noise, ω	Identification noise, ϕ				
	0%	1%	5%	10%	20%
0%	99 (92)	100 (91)	99 (15)	96 (88)	65 (7)
1%	100 (89)	99 (91)	99 (87)	100 (72)	83 (1)
5%	100 (88)	100 (89)	98 (74)	100 (15)	27 (1)
10%	100 (75)	100 (87)	94 (12)	93 (9)	8 (1)
20%	83 (2)	70 (9)	64 (1)	21 (1)	1 (1)

3.6   Conclusion

The localization algorithm proposed in this paper was found to provide robust reconstructions of the nodes relative positions in the novel swarm-based mapping approach with the use of non-unique communication identifiers. It uses a novel RANSAC method to obtain local geometric consistency of neighbouring nodes for reconstruction. A realistic time-of-flight noise model has been used, which includes imperfect knowledge of the speed and timing of ranging pulses, as well as outlier noise in range measurements and communication identifiers.

Our work shows that depending on the noise levels and the geometric structure of the environment, there is a minimum amount of communication identifiers required for successful reconstruction. This can in our case be as low as 2% of the amount of nodes in the swarm. Under the specific constraints that we are investigating, increasing the amount of identifiers above this minimum, or even using unique identification,

does not significantly improve localization. This relaxation aids the feasibility of sensor swarm mapping using resource-limited sensor nodes.

This work currently only considers a static case in which nodes have enough energy to perform a single distance measurement. When considering cases in which the nodes are able to perform subsequent multiple measurement, additional analysis can be performed: e.g. ambiguities can be filtered out better and the nodes distribution throughout the environment can be less homogeneous; this is subject of future research.

References

[1] E. Talnishnikh *et al.* , "Micro motes: A Highly Penetrating Probe for Inaccessible Environments", in *Intelligent Environmental Sensing*, edt. H. Leung, Springer International Publishing, 2015, pp 33-49.

[2] "Xploring Wisenodes", INCAS3.

[3] EU Horizon 2020 FET-Open project: PHOENIX. www.phoenix-project.eu

[4] E.H.A. Duisterwinkel, "Asymmetric Multi-Way Ranging for Resource-Limited Nodes", in *8th EAI International Conference on Ad Hoc Networks*, Sept. 2016, Ottawa, Canada

[5] H. Wymeersch *et al.* , "Cooperative Localization in Wireless Networks", in *Proc. IEEE*, 972 :427450, Feb. 2009.

[6] R. Kuemmerle *et al.* , g20: A General Framework for Graph Optimization In *IEEE Int. Conf. Robotics and Automation*, 2011.

[7] S. Schlupkothen *et al.* , "A Novel Low-Complexity Numerical Localization Method for Dynamic Wireless Sensor Networks", *IEEE Trans. Signal Processing*, Aug. 2015

[8] G. Dubbelman *et al.* , "Robust Sensor Cloud Localization from Range Measurements", in *IEEE Int. Conf. Intelligent Robots and Systems*, Chicago, Illinois, USA, Sept. 2014.

[9] J. Djugash *et al.* , "Range-only SLAM for Robots Operating Cooperatively with Sensor Networks", in *IEEE Int. Conf. Robotics and Automation*, p.2078-2084, Orlando, Florida, USA, 2006.

[10] E. Menegatti *et al.* , "Range-only SLAM with a Mobile Robot and a Wireless Sensor Network", in *IEEE Int. Conf. Robotics and Automation*, p.8-14, Kobe, Japan, May 2009.

[11] E. Duisterwinkel *et al.* , "Environment mapping and localization with an uncontrolled swarm of ultrasound sensor motes", in *Proc. of Meetings on Acoustics*, 201 , 2014.

[12] A. Franchi *et al.* , "Mutual Localization in Multi-Robot Systems using Anonymous Relative Measurements *Int. Journal of Robotics Research*, vol. 32, no. 11 (2013) : p.1302-1322.

[13] G. Han *et al.* , "Localization algorithms of Wireless Sensor Networks: a survey

Telecommunication Systems, 524):2419-2436, August 2011.

[14] M. Erol-kantarci and H. T. Mouftah, "Localization Techniques for Underwater Acoustic Sensor Networks *IEEE Commun. Mag.* , p.152-158, Dec. 2010.

[15] J. A. Rothermich *et al.* , "Distributed Localization and Mapping with a Robotic Swarm *Swarm Robotics*, Springer Berlin Heidelberg, 2005.

[16] J. McLurkin and J. Smith, "Distributed Algorithms for Dispersion in Indoor Environments Using a Swarm of Autonomous Mobile Robots *Distributed Autonomous Robotic Systems 6*, Springer Japan, p.399- 408, 2007.

[17] D. Hahnel, "Mapping and Localization with RFID Technology In *IEEE Int. Conf. Robotics and Automation* (ICRA , New Orleans, LA (USA , 2004

[18] S. Schlupkothen and G. Ascheid, "Localization of wireless sensor networks with concurrently used identification sequences in *IEEE Ad Hoc Networking Workshop* (MED-HOC-NET) , 2015

[19] S. Schlupkothen *et al.* , "A dynamic programming algorithm for resolving transmit-ambiguities in the localization of WSN in *Mediterranean Ad Hoc Networking Workshop* (Med-Hoc-Net , Vilanova i la Geltru, Spain, June 2016, pp. 1-8

[20] Z.D. Deng *et al.* , "An injectable acoustic transmitter for juvenile salmon *Scientific Reports 5*, 2015.

[21] I.F. Akyildiz *et al.* , "Underwater acoustic sensor networks: research challenges *Ad Hoc Networks 3*, 2005, p.257279.

[22] D. Dardari *et al.* , "Ranging With Ultrawide Bandwidth Signals in Multipath Environments *Proc. IEEE*, vol. 97, 2009.

4

Go-with-the-Flow Swarm Sensing in Inaccessible Viscous Media

authors:
Erik H.A. Duisterwinkel, Gijs Dubbelman, Elena Talnishnikh, Jan J.W.M. Bergmans, Heinrich J. Wörtche, Jean-Paul M.G. Linnartz

This paper extends the 'go-with-the-flow' method to explore enclosed environments, like oil reservoirs, pipe lines that transport liquids, and industrial tanks for processing chemicals, where sensing nodes cannot establish communication with the external world. Nonetheless, large quantities of highly miniaturized, thus power-constrained sensor nodes are injected into these environment and flow through them along with the medium, monitoring their environment but also reconstructing their time-varying position from mutual communication, but without any communication to external base stations or beacons. The relative trajectories of nodes yield essential insights of the fluid flow in the otherwise inaccessible environment. We present a functional implementation of a ranging protocol accommodating size and energy constraints. Our simulation chain models node movement from different types of flow dynamics. It comprehensively assesses not only the performance of the communication and ranging protocols, but also of the reconstruction algorithm. Our assessments cover a wide range of different environments and flow profiles, including highly dynamic ones.
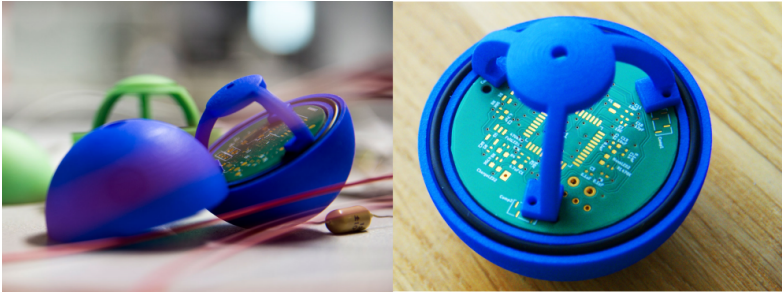
Figure 4.1: Prototype nodes from the Phoenix project that are developed for experimental tests.

4.1   Introduction

Monitoring or exploring the state and conditions of enclosed environments is hard but often economically and environmentally important. The rising need for exploring enclosed environments that are hardly accessible is seen in e.g. the optimization of transporting and processing liquids and chemicals through pipelines and in industrial (mixing) tanks, and the extraction and storage of valuable resources like oil and heat in underground channels and formations like reservoirs and geothermal sources. These man-made or natural environments have in common that they enclose a (semi-)liquid medium. Yet, due to the shielding of metallic containment or surrounding ground layer(s) and the salinity of the liquid medium, the interior can hardly or not at all be accessed by remote probing using e.g. electromagnetic (EM) signals. So, communication to and from these environments is not feasible. Physical probing is only possible for a few of these environments but requires highly specialized and expensive hardware [1–3]. Novel exploration methods for these environments that are versatile and cheap enough are strongly desired.

In the search for new exploration methods we elaborate on the promising 'go-with-the-flow' approach [4, 5]. It is characterized by directly injecting large quantities of miniaturized sensor nodes, i.e. a *swarm* of *sensor nodes*, into the flooded system, and let them go-with-the-flow in order to penetrate the environment and to perform in situ measurements. Examples of such nodes are shown in Fig. 4.1 and the operations schematically visualized in Fig. 4.2a. The nodes store the measurements in their memory. Upon extraction of the nodes, their stored measurements are read out using e.g. USB or NFC, and analyzed off-line where ample computation power is available.

A crucial requirement is to obtain knowledge on the nodes' positions during their trajectory in the flow. Structural information about the environment can be extracted from this and sensor measurements of relevant parameters (e.g. temperature, pressure, salinity) can be visualized on a map as exemplified in Fig. 4.2b.

Obtaining positions from individual nodes in a swarm is not new. Recent studies also show that this is feasible in a go-with-the-flow approach [6, 7]. However, these previous studies relied on external localization hardware in the form of a distributed system of reference points that has at least a one-directional communication channel available towards or from
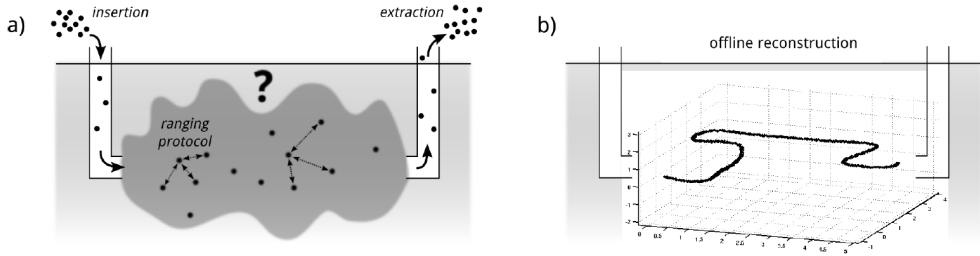
Figure 4.2: Overview of applications: a) insertion of large quantities of sensor nodes; nodes go-with-the-flow of the medium and perform distance measurements using a ranging protocol and store their measurements in memory; upon extraction the data of the nodes is read out for further analysis; b) the measured distances between nodes is used to reconstruct the positions of the nodes relative to each other. The resulting shape of the swarm reveals structural information of the environment. The 3-D map can be used to visualize additional sensor information.
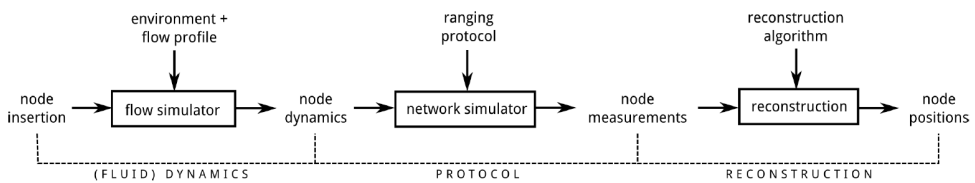
Figure 4.3: Overview of the simulation chain: in a flow simulator, the environment and flow profile determine the trajectory of the node positions after insertion; the new ranging protocol is then simulated on the dynamic swarm using a network simulator; the nodes distance measurements are then used in a robust reconstruction algorithm to estimate their positions relative to each other.

nodes in the swarm. Cooperative localization, as e.g. described in [8], is an approach where nodes only rely on neighboring nodes for reference. It is often used in robotics where one can not rely on e.g. GPS or other beacon-based localization [6–10]. Even though these co-operative localization methods become more energy and computation efficient, they are not adapted to the stringent constraints and adverse conditions that characterize our go-with-the-flow applications. The specific set of constraints in our applications is summarized below, and elaborated further on in Sec. 4.2:

1. the size constraint posed by the environments results in limited hardware possibilities (energy, processing, sensing);

2. the lack of external communication and reference points requires an autonomous and cooperative approach for relative position determination;

3. the unknown and relatively fast movement of the nodes results in highly time-variant acoustic communication channels between nodes.

This set of constraints necessitates a new ranging protocol and a robust reconstruction algorithm. In this paper we present a functional example of such a ranging protocol and reconstruction algorithm and explore the underlying design trade-offs. The distance measurements between nodes are performed while going with the flow (on-line), the obtained data is analyzed after the nodes are retrieved (off-line) and node positions are reconstructed based on the measured distances between them.

Our simulation chain that links all the different aspects together is shown in Fig. 4.3. It consists of three parts:

- simulation of flow profiles and node trajectories through different environments;

- simulation of performing distance measurements in a network simulator, using the ranging protocol;

- the algorithm that robustly reconstructs the relative node positions based on the distance measurements.

In this paper, we extend and incorporate our robust reconstruction algorithm that was shown to be able to deal with highly corrupted distance measurements. In previous work we tested the limits to how severely the distance measurements could be corrupted before reconstruction fails [11–13]. The assumption was made that the nodes are static, i.e., the distance measurements are instantaneous and modeled using a large set of different noise types to account for a variety of imperfections in the measurements. To resolve this unrealistic assumption we extend the simulation chain in [14] by including dynamic behavior of the nodes. In this paper we distinguish and implement three classes of environments and flow profiles to cover a wide range of possible realistic scenarios, as inspired by commonly used models fluid dynamic modes [15].

It is an ongoing activity in the project (European Union's Horizon 2020: Phoenix project) to also confirm this by real-life tests, using nodes as in Fig. 4.1 [5, 16–18]. The goal of this project is to develop sensor nodes to explore difficult-to-access environments, such as the ones presented in this paper. The corner-stone of the approach in our project is the co-evolution between experiments and simulations. Simulations like the one presented in this paper provide insight in the trade-offs in hardware specifications and protocol implementa-

tions on the final reconstruction performance. This guides the hardware development of the nodes to perform the actual experiments.

This paper is structured as follows. In Sec. 4.2 we start off by describing the challenges in our applications and why existing techniques cannot be used. These difficulties are translated into specific design challenges for obtaining distance measurements through the ranging protocol in Sec. 4.3.1. Sec. 4.3.2 provides a functional example of the ranging protocol is our attempt in tackling these challenges and allowing to make proper trade-offs between all parameters involved into getting the final measurement performance. In Sec. 4.4 the individual aspects of the entire simulation chain are described. The result section, Sec. 4.5, shows the metrics that we use to assess the performance of the ranging protocol and the final reconstruction. The conclusion can be found in Sec. 4.6.

4.2   Application challenges

Measuring distances between autonomous sensor nodes and reconstructing their positions (also called *localization*) based on these distances are well studied topics, e.g. [8–10]. The work presented in this paper addresses the combination of stringent application challenges that make existing work insufficient for our application.

4.2.1   Size and energy

While size and energy is a common theme in Internet of Things studies, our sensor nodes need to be small enough to physically pass through the narrowest parts of the environment (typically millimeters or centimeters). Similarly, the node weight needs to be small enough to allow the nodes to move with the flow. In some applications the nodes even need to be neutrally buoyant in the medium, reducing the weight limit further. While we do not explicitly design nodes of a certain weight and size, in this paper we aim at restricting the number of messages that are transmitted, received and analyzed, but also the number of bits in these messages.

4.2.2   Enclosed environment

The environments targeted by this exploration approach are generally enclosed environments, either due to a metal containment (e.g. a pipe) or due to being deep underground. Communication with the outside world is not feasible as the signals will be absorbed. Signals with larger wavelengths experience less absorption but are less efficiently detected and produced by the small nodes. Therefore, unlike in [6, 7], we cannot rely on techniques that use external beacons or anchors (e.g. satellites or ultrasound beacons/receivers), to which the nodes can relate to. The nodes can only see signals from neighboring nodes. Our solution solves this by collecting ranging measurements between nodes during the travel and analyze these afterwards, when the nodes are collected and data read-out.

4.2.3   Required communication range

Depending on the typical dimensions of the environment, the size of the nodes and the spatial distribution of the nodes, a required communication range of 1 meter appeared reasonable.

Using electromagnetic (EM) communication between nodes is not attractive, as (sub-)centimeter node dimensions prohibit large antennas. Yet, at high frequencies suitable to be used with small antenna's, the liquid (and possible highly saline) media have huge attenuation. This effectively prevents EM communication for distances larger than, say, a centimeter. Using inductive communication such as near-field communication (NFC) is also limited to the centimeter range in such environments [19]. The (liquid) media is also not guaranteed to be clear enough for the use of (visible) light communication. Communication between the nodes using ultrasound can be performed in these environments, and will be studied here.

4.2.4   Acoustic channel

The main challenge in underwater acoustic communication are the random and highly time-variant channels [9]. This limits the quantity of data that can be send and the rate at which it is transmitted to ensure packet reception. The ranging latency (defined as the time it takes to finish a ranging transaction) should be small, to reduce the displacement of the nodes relative to each other within this time. In some cases, nodes will move out of range before a ranging transaction can be completed.

The goal of the localization is to obtain snapshots of the positions of the nodes relative to each other. For an accurate representation of this snapshot, the ranging accuracy between nodes need to be within 1-5% of the communication range, with a communication range of 1 meter, this is around 1-5 centimeter.. For the constellation of all relative positions of nodes to be considered quasi-static, the nodes should not move more than this ranging accuracy before all messages are exchanged. This limits the time frame, or the total latency, of the ranging transactions to only 50-250 milliseconds, see Sec. 4.3.1.4.

The ultrasound frequencies that can be used in this application are limited to one or at most a few hundreds of kHz. This is bounded on the lower side by the required ranging accuracy and on the upper side mainly by node size and (omni-) directionality requirement of the transmission and reception. Since the nodes operate in a swarm and their positions relative to each other is not known while measuring, their transmission and reception ability should extend all directions. This limits the ultrasound wavelength to around the size of the transducer [20], which in turn is limited by the size of the node itself (millimeters to centimeters). True omnidirectionallity cannot be achieved in practice, these non-ideal circumstances are tested in our earlier work in [13].

In order for the localization method, presented in this paper and in [11, 21], to reconstruct an entire swarm of 100's-1000's of nodes, each node should have around 20 neighboring nodes within their communication range. Finishing a complete ranging cycle of the swarm within hundreds of milliseconds, considering the propagation time, processing time and the limited ability to deal with signal collisions, limits the maximum message length to around 0.2 ms. Assuming a bandwidth of 50% leaves a maximum of around 10 bits per message. This makes establishing stable and robust communication between the hundreds or thousands of nodes a hard challenge.

4.2.5   Random Channel Access

Every time that a message arrives, its phase/timing/synchronization needs to be established, and compared to the most recent previous message of the same node. Typically, the channel changes significantly between two messages. This requires a trade off between frequent updates from the same node and messages being lost in packet time collision. Since no elaborate signal separation algorithms can be run, colliding messages may result in the loss of one or both messages.

4.3   Ranging protocol

This section describes the design challenges in Sec. 4.3.1, and describes our protocol to solve these in Sec. 4.3.2.

4.3.1   Design challenges

Our protocol must efficiently use the communication channels, avoid overhead, identify neighboring nodes, and reduce the ranging latency.

4.3.1.1   Reduction of transmitted messages

Since battery capacity and channel capacity are scarce resources in the miniaturized nodes, the number of messages need to be restricted.

The concept of multi-way ranging (MWR), initially proposed in [22] as N-Way Time Transfer, is adopted as the base of the protocol because of the significant reduction of messages needed to complete a full ranging cycle (i.e., to perform all possible distance measurements). It exploits the fact that messages are also received by nodes other than the one addressed in the packet. It utilizes the time-difference-of-arrival (TDOA) of messages to help determine the distance between nodes within the communication radius. More common methods such as two-way ranging (TWR) [23] only considers bi-directional communication for determining the range and ignore messages received by nodes that are not the addressee. Even though the quantity of messages received to estimate the distances is within the same order of magnitude, the required number of messages sent in TWR scales quadratic with the number of nodes, where in MWR this number only scales linearly with the number of nodes.

The MWR in existing literature, however, does not describe how to deal with sparse connectivity in swarms as in our applications, and neither does it include a scanning phase where the presence of neighbouring nodes can be checked. As the network topology is non-static, it is also not known which neighboring nodes are within communication range. The simple sequence of events in traditional MWR [22], where node $i+1$ transmits a ranging signal after node i, cannot be easily controlled in these applications. To solve this issue, in our protocol we use a master-slave system, as further described in Sec. 4.3.2.

Besides the energy required to transmit messages, signal overlap of the messages (collisions) also requires a more elaborate signal processing to filter and distinguish signals, and therefore also requires more energy. The system needs to trade off message quantity and intervals with excessive collisions. The ranging protocol in Sec. 4.3.2 uses a time-divided

communication/ranging scheme to reduce signal overlap without the use of an additional component to check for channel occupancy.

4.3.1.2 Elimination of overhead

Due to our severely constrained system, we had to avoid explicit synchronization. Moreover our packets do not carry information other than to allow the nodes to identify the sender and addressee based on an identifier. This information is stored locally for off-line processing.

To gain efficiency we are not using a traditional mac-/network layer, as commonly used in the OSI-model. We are not reattempting transmissions of messages lost in collisions or inhibited because of a busy channel. Instead, any 'retransmission' carries new, fresh data for a new ranging cycle. In fact, messages that are sent later than planned cannot be used to estimate the distance using the T(D)OA information without further communication overhead.

The control of the entire ranging process should only depend on limited information that is locally available to the nodes. Not (properly) receiving messages should not result in a global failure of the measured distances, but be limited to local failures only.

4.3.1.3 Scanning

As it is not known which nodes are within communication range, most ranging protocols initiate a separate scanning phase before the ranging phase. In this phase, nodes determine which neighboring nodes are within communication range to determine which nodes to perform distance measurements to. Not only does such an additional scanning phase weigh in the energy budget, it also faces all the similar challenges in data transfer as the ranging part.

In our protocol described in Sec. 4.3.2 we show that the scanning phase can be omitted and the identification issue solved by addressing all possible nodes that are within communication range by using a limited set of non-unique calling identifiers. In fact, this results in a trade-off between ranging latency and channel occupancy, but we show that appropriate compromises can be chosen that perform well in different scenarios.

4.3.1.4 Ranging latency

To make sure the movement of the nodes does not disturb the ranging measurement significantly we are aiming for quasi-static snapshots of the positions of the nodes relative to each other. This is only feasible when the ranging transactions are completed within a specific time window. We call this time window the ranging latency and can quantify the maximum allowed ranging latency based on the maximum allowed node displacement within a snapshot.

Table 4.1 shows the relationship between the ranging latency, the nodes relative velocity, and the node displacement. These numbers are expressed in percentage of an average distance of 0.67 meter[1].

The maximum latency design goals can be set to e.g. 60 ms for node to node ranging latency, and to e.g. 200 ms for the ranging latency of the entire swarm. When the relative

---

[1]This average is obtained for a communication radius of 1 meter and a uniform distribution of neighboring nodes in the swarm.

Table 4.1: Displacement of nodes relative to each other depending on their relative speed and the duration of ranging transaction (latency). Expressed in percentage of average distance of 0.67 m.

	latency		
relative speed	10 ms	60 ms	200 ms
0.1 m/s	0.2%	0.9%	3%
0.2 m/s	0.3%	1.8%	6%
0.5 m/s	0.8%	4.5%	15%
1.0 m/s	1.5%	9%	30%

node velocities are ≤ 1 m/s, the average distance displacements due to latency are $\leq 9\%$ and $\leq 30\%$, respectively.

4.3.1.5 Error model

The off-line reconstruction algorithm is robust to deal with failed ranging transactions between node pairs as shown in [13], as long as sufficient other ranging transactions did succeed. The priority of the ranging protocol is therefore that individual failures of ranging transactions should not have a significant effect on the rest of the swarm.

The reconstruction algorithm is tested against a wide variety of possible ranging errors (Gaussian noise, outlier noise, bit errors, etc.). However, the mentioned earlier work did not study whether it can deal with potential systematic errors arising from the use of the used ranging protocol. This is to be tested in the simulation of the entire chain in this work.

4.3.2 Protocol

Traditional MWR, as introduced in [22], consists of nodes that sequentially transmit a single message to all (neighboring) nodes. A node with identifier 1 starts the procedure by transmitting a message that contains his identifier, node n+1 then responds by transmitting a message with his identifier to all (neighboring) nodes. When all messages and their local transmission and reception times are collected (or transmitted to all other nodes), there is enough information collected by all nodes to calculate the distances between all nodes when all of this information is gathered in a central place. This specific method requires that all nodes are within communication range of each other, and no single node or transmission is allowed to fail. In our application, connectivity is sparse (i.e. not all nodes are within communication range of each other) and it is unknown which nodes are within range. Therefore, this simple sequence of events cannot be used.

In our work, the ranging procedure is controlled by pre-assigned *master nodes* that transmit request (REQ) messages to their neighboring nodes. The REQ consists of the sender's identifier, as well as the receiver's identifier. Nodes having received any REQ (even if they are not the addressee) become *slave nodes*. The slave nodes that are called upon in the REQ, respond to the master node with an acknowledgement (ACK) message. This ACK message also consists of the sender's and receiver's identifier. The master node will address all neigh-
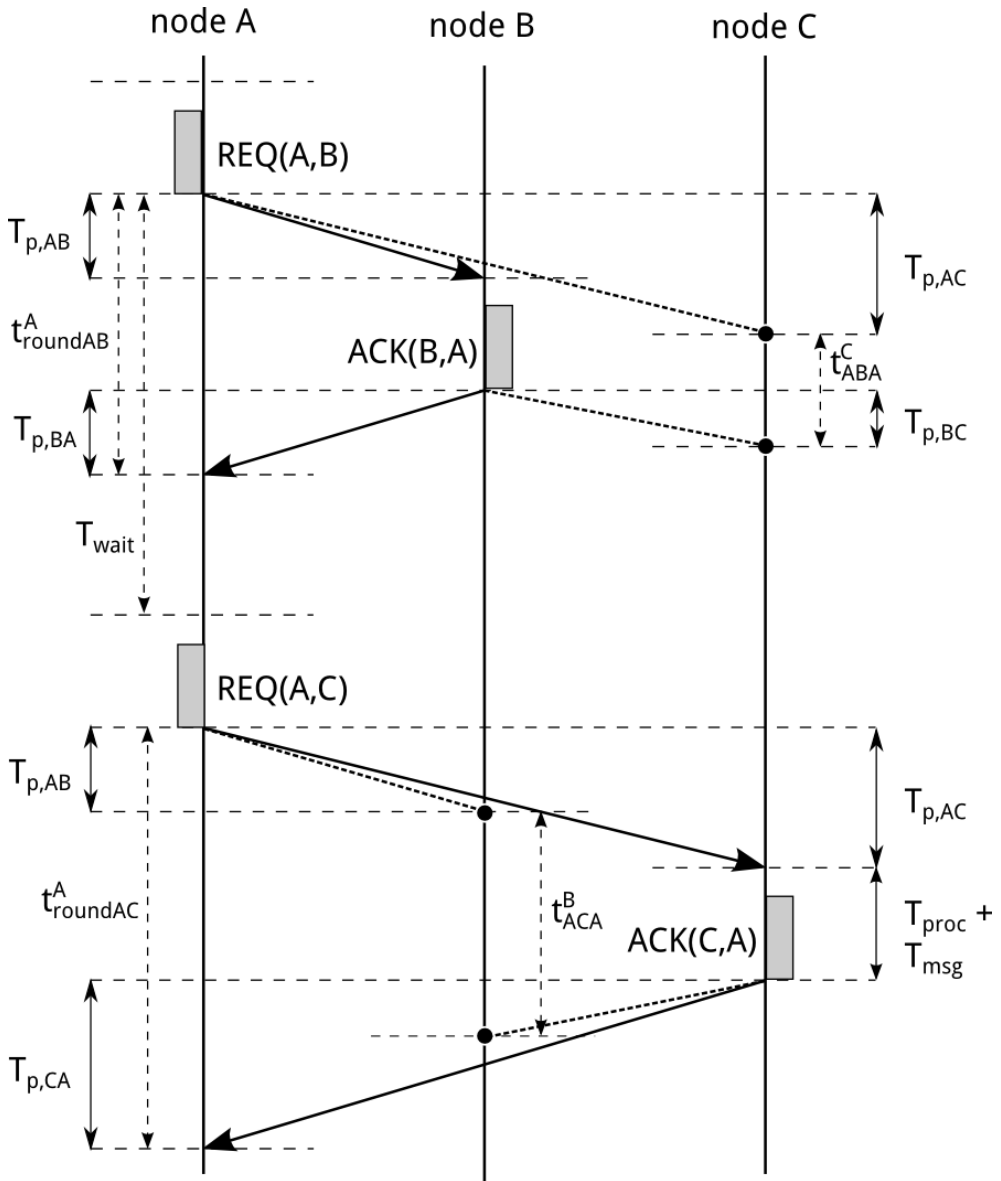
Figure 4.4: One ranging cycle of master node A and slave nodes B and C. The master node emits request (REQ) signals and slave nodes respond with and acknowledgement (ACK) signal if it is addressed to them. Knowledge of the nodes internal timing information $t^A_{roundAB}$, $t^A_{roundAC}$, t^C_{ABA} and t^B_{ACA} and the fixed and known value of the processing time T_{proc} is sufficient to determine the propagation times $T_{p,AB}$, $T_{p,AC}$ and $T_{p,BC}$ between the nodes.
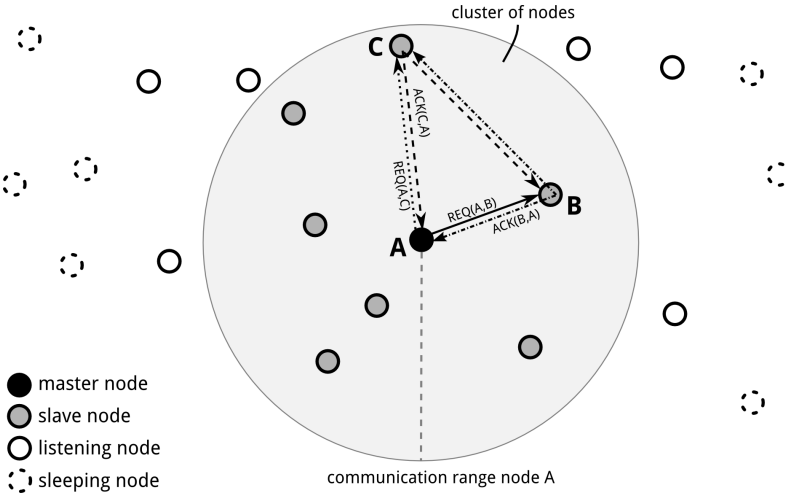
Figure 4.5: Master node A initiates the ranging process to node B and C. The nodes within communication range become slave node and respond to REQ signals with an ACK signal. All nodes within the respective communication range receive the signals and store them: in off-line analysis they can be used to determine round-trip TOF between nodes. The nodes outside the cluster will only receive ACK signals. In this figure, not all signals (arrows) are drawn.

boring nodes one by one. The communication scheme that is used is illustrated in Fig. 4.4. The *cluster of nodes* that is formed by this master node and the slave nodes is illustrated in Fig. 4.5.

In a ranging transaction where master node A requests an answer from node B, the timestamps of transmission of the REQ and reception of the ACK at node A provides knowledge about the round-trip TOF between nodes A and B. Time-parameter $t^A_{roundAB}$ indicates this round-trip TOF between node A and B as measured by node A, plus the processing time and the message duration. The node's internal processing time T_{proc} and signal message time T_{msg} are known beforehand and are fixed (and equal for all nodes), therefore, the round-trip propagation time between A and B can be estimated.

$$T_{p,AB} + T_{p,BA} = t^A_{roundAB} - T_{\text{proc}} - T_{\text{msg}} \qquad (4.1)$$

where $T_{p,AB}$ is the propagation time of a message from node A to B and is $T_{p,BA}$ in opposite direction. After this ranging transactions, master node A performs a similar transaction to node C and to the other nodes within the cluster.

Because nodes B and C are also within each other's communication range, node C receives the ranging signals between nodes A and B; and node B receives the ranging signals between nodes A and C. The time difference t^C_{ABA} between the arrival of REQ(A,B) and ACK(B,A) at node C and the time difference t^B_{ACA} between the arrival of REQ(A,C) and ACK(C,A) at node B, can be used to calculate the propagation time between nodes B and C using:

$$T_{p,BC} + T_{p,CB} = t^C_{ABA} + t^B_{ACA} - 2(T_{\text{proc}} + T_{\text{msg}}) \qquad (4.2)$$

Note that not all nodes within the cluster are within each others communication range.

In the following sections, we will discuss how the master nodes are assigned; how the neighboring nodes are called upon, even though there is a lack of knowledge on which nodes are within communication range; how we attempt to reduce the latency; and how we deal with de-synchronization between nodes. The reader is referred to our earlier work Ref. [14] for more details on the protocol.

4.3.2.1　Picking the master node

The role of master node alternates among all nodes in the network. The advantage of this is that the power consumption is distributed evenly over all nodes (master nodes transmit more signals).

Within the time frame T_{sample} of one sample, in which a complete ranging cycle is completed for all nodes, the role of master node is chosen randomly. This is performed by letting all nodes chose a random delay time T_D at the beginning of a sample. Nodes become master when their sample timer t_s, that is set to zero at the beginning of a sample, exceeds $t_s > T_D$. A node becomes slave node if it receives any REQ signal from a master node before their own sample time reaches its T_D. The master node initiates the ranging transactions as described above, thereby forming a cluster of nodes as in Fig. 4.5.

Throughout the entire network, several of these clusters are formed in which ranging transactions are performed. Every sample, these clusters change based on which nodes have become master node. Overlap between these clusters is needed to reconstruct the entire swarm instead of the individual clusters.

4.3.2.2　Calling the slave nodes

Within one ranging cycle, the master node sends a request to all slave nodes in the cluster, but it is not known beforehand which nodes are within communication range. Regular scanning techniques depend on the availability of sufficient bandwidth, processing power or time to perform broadcasting.

In this work, we propose that the master node initiates the ranging transactions to all possible hardware addresses. But as the total number of nodes in the network can be very large and the connections are sparse, this will be very inefficient as most requests remain unanswered. Instead of requesting to the hardware's *unique identifiers* (UID) the master node requests to highly abbreviated *calling identifiers* (CID). The master node only initiates n_f times a ranging transaction to CID $= \{0, 1, ..., n_f\text{-}1\}$. Slave nodes will respond if and only if their unique hardware identifier satisfies

$$\text{mod}(\text{UID}, n_f) = \text{CID}. \tag{4.3}$$

As multiple nodes will have identical CIDs, the probability arises that multiple nodes will respond to the same request. If the ACKs of the responding slave nodes do not overlap such that the signals cannot be distinguished and decoded anymore at the receiving node, the determination of the round-trip TOF of each of them can still be performed. Parameter n_f can be chosen both off-line as on-line to adjust for the amount of neighboring nodes and the
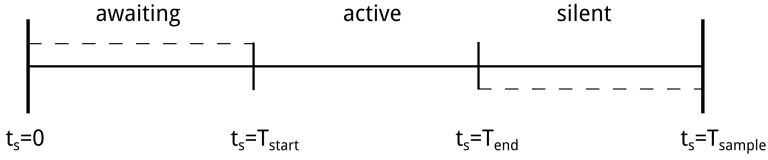
Figure 4.6: Subdivision of sample in three parts: awaiting, active and silent. Each sample, the start delay that determines when to become master is randomly chosen within the active period (drawn from a uniform distribution)..

total signal overlap. In fact, since we do not actively resolve collisions via retransmissions, an optimization for 'first time right' appears relevant .

After transmission of a REQ, the master node waits $T_{\text{wait}} = 2T_{\text{p,max}} + T_{\text{proc}} + T_{\text{msg}}$ before it proceeds with the request for the following CID. Here, $T_{\text{p,max}}$ accounts for the propagation time required to reach the end of the (expected) communication range. This is to allow the reception of all possible ACKs within the (expected) communication radius.

After all ranging transactions have been performed, the nodes will go into a low-energy sleep mode to await the start of the next sample. In this period until awakening, the node will not be active. The node will wake up and start a new sample when the node's internal sample timer reaches $t_s > T_{\text{sample}}$.

4.3.2.3 Synchronization

Absolute synchronization is not required as all distances are obtained using direct or indirect round-trip TOF measurement. It is however beneficial to have nodes synchronized to a level in which samples are aligned such that the avalanche effect introduced in Sec. 4.3.2.4 allows nodes to sleep for the majority of the sample time instead of responding to nodes that are in a previous or next sample.

In order for connected nodes to remain in the same sample, we proposed to subdivide a sample on the node level into time-slots as illustrated in Fig. 4.6. The random delay time to become master node is chosen from a uniform distribution within the range $T_D \in (T_{\text{start}}, T_{\text{end}})$ or the *active time period*. The internal sample timer t_s is reset to $t_s = T_{\text{start}}$ when becoming master; or, at reception of the first signal (any REQ or ACK) in the sample, as illustrated in Fig. 4.7. This will assure that connected nodes remain synchronized to the sample level, as long as (groups of) nodes have not been disconnected from each other. In the *awaiting period* nodes do not become master and can only receive signals. In the *silent period*, nodes have already received their first signal or already became master node. In our work, the three time periods are chosen to be of equal length. It appeared effective to simply choose the three time periods to be of equal length, i.e. $T_{\text{start}} = \frac{1}{3}T_{\text{sample}}$ and $T_{\text{end}} = \frac{2}{3}T_{\text{sample}}$. The total sample time can be chosen based on the requested frequency of a complete ranging cycle and/or on the available energy and storage available during their passages in the environment.
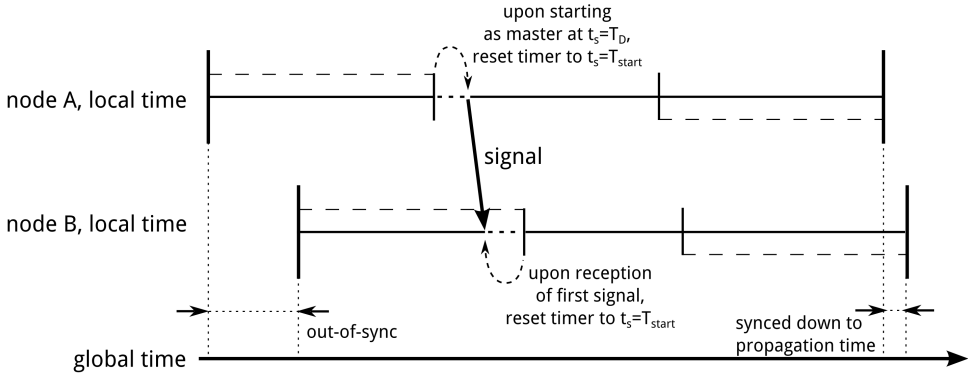
Figure 4.7: Upon first reception of signal in the node's sample, or, upon becoming master node, the sample timer t_s is reset to $t_s = T_{\text{start}}$.

4.3.2.4 Reducing the latency

As seen in Fig. 4.5, a node just outside the communication range of the master node does not become slave node and would have to wait till either another a node within its communication range becomes master or till it becomes a master node itself. To avoid excessive latency in forming master slave combinations, we introduce a dynamic reduction of the delay time T_D for any node just outside existing clusters. In fact, we exploit the insight that if a node does receive an ACK signals without having received the initiating REQ signals, it is very likely to be just outside an already formed cluster. Their remaining delay time before they become master node is reduced by a factor $M_{\text{avalanche}}$ at the reception of any ACK signal. In order to prevent multiple nodes just outside the cluster to become master node around the same time, the reduction of T_D is limited to CID·T_{wait}, that is, we intentionally let it depend on the node's CID, as this reduces the chances of having clusters with multiple master nodes.

4.4 Simulation chain

We simulate the system using the chain illustrated in Fig. 4.3. Three different environments and flow patterns are constructed to represent a broad variety of environments and to obtain typical node trajectories, as we elaborate in Sec. 4.4.1. While traversing the environment, nodes perform distance measurements between each other using a ranging protocol. This is discussed in Sec. 4.4.2, presenting the network simulator. Node positions are then reconstructed based on these distance measurements using the reconstruction algorithm of Sec. 4.4.3.

4.4.1 Trajectory of nodes

After insertion of the nodes in the environment, the nodes will go-with-the-flow of the liquid medium and traverse the environment. In order to obtain the nodes' trajectory, the nodes' movement throughout the environment is simulated. These simulations consider the nodes to
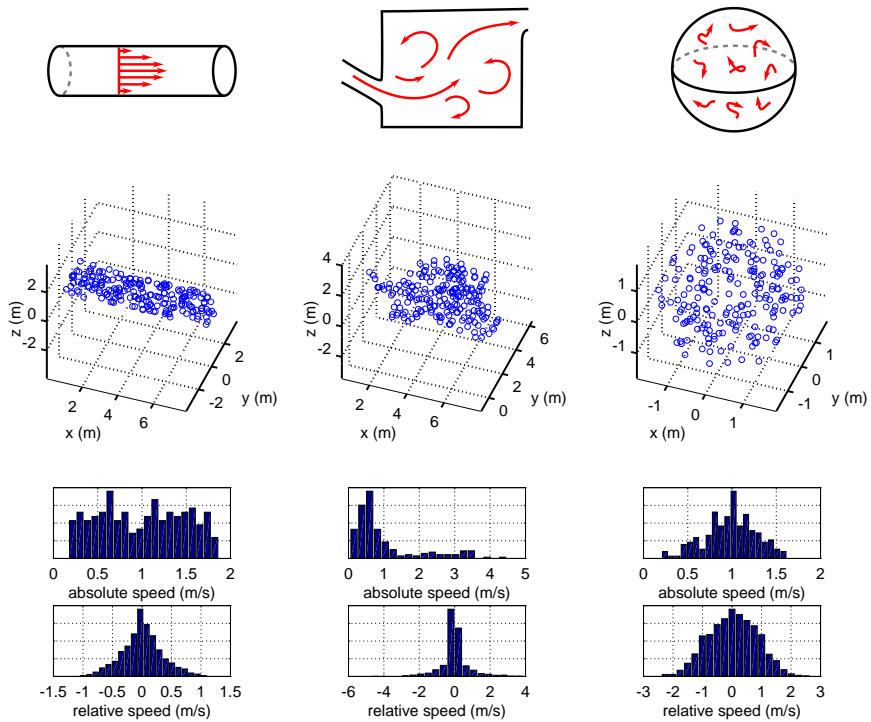
Figure 4.8: Different environments and fluid speed profiles used for the simulations. Each environment is scaled such that each node has on average 20 neighbouring nodes. Relative speed is calculated between nodes within communication radius.

be infinitely small and neutrally buoyant, in fact, they are numerical approximations of tracer particles that move with the flow. Three entirely different scenarios are chosen to account for the wide variety of dynamics that are present in different applications. These are illustrated in Fig. 4.8:

1. an idealized laminar flow profile in a 3-D pipe.

2. 'realistic' turbulent flow in a 2.5-D tank-like environment.

3. random node movement in a 3-D spherical environment.

In order to still be able to compare the environments, they are scaled such that in all scenarios, the instantaneous average node velocity is equal to 1 m/s and that if the communication radius is set to 1 meter, the average number of neighboring nodes within communication range is 20 in all scenarios. Each scenario is generated with $N = 200$ nodes and all nodes are initially uniformly distributed within the environment.

In scenario 1, representing a straight pipe-section, the flow of the medium is approximated by a perfect laminar flow profile. The nodes experience a velocity similar to that of the flow. Due to the above-mentioned uniform scaling between scenarios, the pipe has a diameter of close to 2 meter. Even though in practice a pipe with this diameter will not yield a perfect laminar flow with water-like liquids, the scenario of a laminar flow is still relevant, not only as a limiting, extreme case but it also approximates commonly experienced situations.

The tank scenario 2 is generated using a 2-D computational fluid dynamics (CFD) simulator (Java-based) with predefined boundary conditions, similar to what we presented in Ref. [14]. We modeled one inlet and one outlet for the flow. A third dimension is added by uniformly assigning a (vertical) z-position but keeping the z-velocity zero. The tank dimensions (without inlet) are $6 \times 5 \times 0.7$ meter (x,y,z).

The sphere scenario 3 is generated entirely differently, to widen the variety of environments. Nodes move randomly in a sphere of 4 meters in diameter. To express a random walk, the course direction and the node velocity is changed at random from a uniform distribution every 60 ms. A smooth path is obtained by drawing a spline through the course points. These dynamics are inspired by collision on a molecular level.

The different scenarios give rise to different velocity distributions, typical examples are also shown in Fig. 4.8. The relative velocities are taken only from the nodes that are within communication range of each other.

For scenario 1 and 3, a total of 10 different datasets are generated for different iterations of the simulations, each with different randomly drawn values.

4.4.2 Distance measurements

The protocol of distance measurements between the dynamic nodes is simulated in network simulator OMNeT++ [25, 26]. The ranging protocol is translated into a simple sequence of events for each individual node such that they can operate autonomously, using only their internal clock and acting upon input received from neighboring nodes.

The ranging protocol described in Sec. 4.3.2 was coded into the network simulator environment. It takes as input the nodes positions over time that are simulated as described in

Sec. 4.4.1. Furthermore, additional parameters that are used by the ranging protocol can be adjusted like the communication distance, bit rate of transmission, and timing parameters.

For the simulations in this paper, the communication range is set to a fixed 1 meter, resulting in an average node density of 20 neighboring nodes within the communication range. The maximum clock frequency deviation of the nodes is set to 100 ppm. The actual frequency deviation is different for each node and is kept fixed throughout the simulation. At the start of the simulation, the internal clock of the nodes have a random absolute offset from a uniform distribution between 0 and 0.1 seconds.

The sample time is set to $T_{sample} = 1$ sec and we considered a number of calling identifier codes (CIDs) of $n_f = 2^5 = 32$ (5 bits).

The ultrasound transmission rate is set to 40 kbit/sec and nodes have a fixed 1 ms processing time for receiving and decoding signals. The avalanche induction is set to $M_{avalanche} = 4$. The parameters n_f and $M_{avalanche}$ are chosen for a reasonable trade-off in ranging latency and signal overlap. The reader is referred to [14] for a more thorough analysis on these parameters.

Additional noise other than the protocol specific ones are not included in this study as the focus is on the 'noise' introduced using this protocol.

The network simulator creates for each node in the network a datafile in which the transmitted and received signals are stored together with the corresponding timestamps. The signals consist of the identifier of the sender and the identifier of the addressee, as well as the type of signal (REQ or ACK). This data output is similar to what each node would individually store in its own memory while performing the ranging protocol.
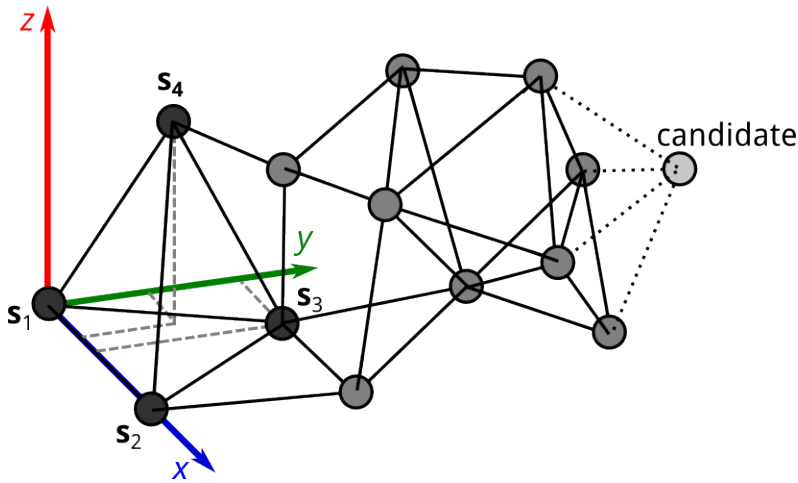
4.4.3 Reconstruction of node positions

Using the stored measurements from the nodes and Eqn. 4.1 and Eqn. 4.2, one can estimate the distances between the nodes within a cluster, as explained in Sec. 4.3.2.

However, the obtained distances are sparse, as only the distances between nodes within a cluster can be obtained. Furthermore, the sparse distances are potentially measured multiple times as each node pair can be part of more than one cluster. The multiple measurements are taken at different times, and therefore are based on a different actual distances. This adds a significant challenge to the reconstruction algorithm as it is not defined which measurement to use and which measured distance is most correct.

In [11–13], we developed a reconstruction algorithm for range-only distance measurements and studied the effect on a large set of possible types of noise, including multiple measured distances between the same node pair. The reconstruction algorithm is schematically visualized in Fig. 4.9. The algorithm uses a guided Random Sampling Consensus (guided-RANSAC) approach to robustly filter out the outlier measurements and based on the geometric consistency of the neighboring nodes, can choose the measurement that is most likely to be the correct one. The nodes positions and the measured distances can be seen as a graph, and can incrementally be constructed as shown in Fig. 4.10. An initial estimate for the nodes positions is constructed which is then optimized using a non-linear graph-optimization algorithm.
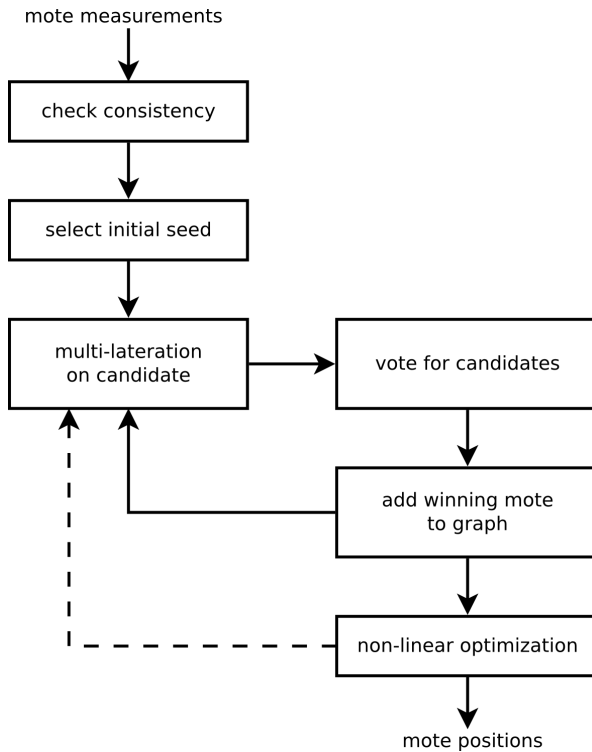
Figure 4.9: Simplified diagram of reconstruction algorithm that takes as input the node measurements and through an iterative RANSAC and non-linear optimization phases it estimates the nodes positions [11–13].

Figure 4.10: Reconstruction of the nodes in the swarm, seen as a graph growing problem. For each node that knows the distances to at least four already reconstructed nodes, a position is estimated using multi-lateration. The first four positions are chosen such that the frame of reference ($\{x, y, z\}$-axes) is defined. A larger amount of distances to neighboring nodes allows to filters out noisy measurements and outliers, therewith improving the estimation.

This reconstruction algorithm considers the distance dataset as a static snapshot of the distances between nodes. From the dynamically generated data, the reconstruction algorithm will attempt to make a static reconstruction of the nodes positions. The difference between the actual (dynamic) distance and the effect the ranging protocol has on the measured distance with respect to this static reconstruction will be considered as noise to the 'ground truth' static snapshot of the nodes positions.

4.5   Performance Evaluation

Performance metrics are calculated for both the ranging between and the reconstruction of the nodes as also used in [12] and [13].

4.5.1   Ranging protocol

The *latency* of the ranging protocol is a measure to study the time differences between different ranging transactions. The maximum latency indicates the time difference between finishing the first and the last ranging transaction in a sample. The average latency is calculated between all individual node pairs that were within communication radius.

In the ranging protocol, the master-slave distance measurements can effectively be seen as round-trip TOF measurements, the error introduced here due to the node movement is only affecting the measurement for the total round-trip TOF time-interval between the master and slave nodes, which in our case is $T_{\text{wait}} < 3ms$. The majority of distances, however, are determined using indirect slave-slave transactions. These indirect round-trip TOF measurements

**Table 4.2:** Performance metrics of reconstruction.

	pipe	tank	sphere
recall $\mathbb{X}$	89 %	78 %	86 %
recall $\mathbb{Y}$	88 %	65 %	85 %
recall $\mathbb{Z}$	88 %	55 %	85 %
MS absolute error (m^2)	8.0e-2	1.5e-1	6.1e-3
MS relative error	1.2e-3	9.2e-3	3.0e-4
global error	6.1e-3	7.6e-2	6.8e-4
local error	2.6e-3	2.6e-2	5.4e-4
average latency (ms)	65±45	58±39	64±46
max latency (ms)	255	205	269
coverage$_{\mathrm{protocol}}$	76 %	73 %	75 %
coverage$_{\mathrm{reconstruction}}$	63 %	65 %	65 %

within a cluster are performed over a much larger time interval, on average $2n_f T_{\mathrm{wait}}/3$, which in these experiments is 46 ms. The effect of the slave-slave movement on the measured distance is therefore much larger than in the master-slave distance measurement. The ranging latency of the entire ranging cycle is much larger. Resulting in an even larger displacement of the nodes on different sides of the swarm.

The *coverage* is defined as the fraction of the node pairs that are within each others communication range that have successfully established either a distance measurement (in the protocol phase) or a reconstructed distance (in the reconstruction phase). Since only distances can be calculated between nodes that are within the same cluster, the coverage will be lower than 100%. Also, the reconstructed coverage cannot exceed the ranging coverage.

4.5.2   Reconstruction metrics

To express and compare the performance of swarm operations like ours, there is not yet a commonly used metric to assess the performance of reconstructing a swarm of nodes. It highly depends on what information is desired from such a swarm. The more conventional metrics like recall and absolute errors do not suffice for our applications. When exploring a yet unknown and difficult-to-access environment, initially one might want to identify the overall structure of the environment: this includes a rough estimation of the local geometry and the overall shape of the total environment. Later, after adding additional sensor information, the fine-grained local geometry typically becomes more relevant as the absolute error over the entire swarm is less relevant. In fact, in the next sections, we address a variety of performance metrics, introduced in [13], that quantify different objectives.
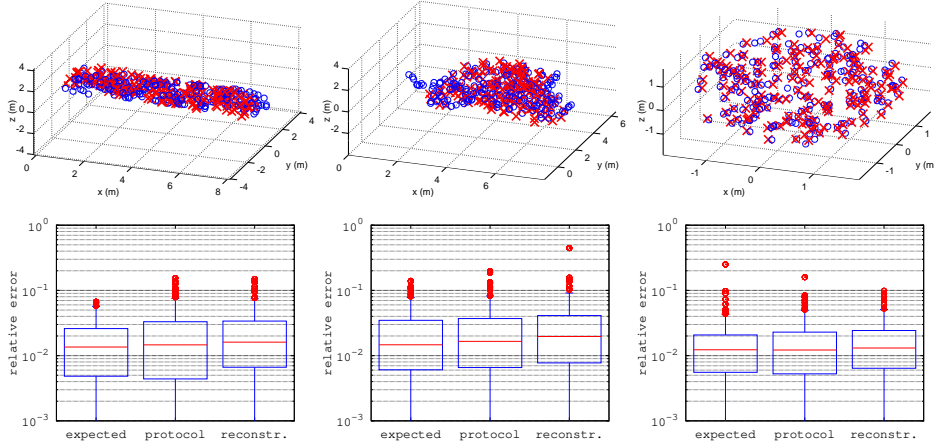
Figure 4.11: Reconstruction of node positions (red crosses) and ground truth positions (blue circles). Below, the relative (distance) errors involved in the simulation chain: the expected error is based on the relative speed distribution and the average latency, the protocol error is from the distance measurements after the ranging protocol, and the reconstructed error is after the node positions are reconstructed.

4.5.2.1 Absolute error

The absolute error is the mean squared absolute distance error of the reconstructed nodes positions relative to their ground truth positions. It is calculated as $\mathbb{E}_{abs} = \sum |\hat{\mathbf{s}} - \bar{\mathbf{s}}|^2 / N$, the sum over all N nodes where $\hat{\mathbf{s}}$ is the reconstructed positions of the node and $\bar{\mathbf{s}}$ its ground truth position. In order to compare ground truth positions with reconstructed positions, a linear fit between the 3-D positions of the initial four nodes and their true positions is performed to solve the general rotation and reflection ambiguity.

4.5.2.2 Relative error

The relative error is the mean squared relative error of the reconstructed distances between nodes, relative to the ground truth distances. It is calculated as $\mathbb{E}_{rel} = \sum |\hat{\mathbf{d}} - \bar{\mathbf{d}}|^2 / M$, the sum over all M reconstructed distances, where $\hat{\mathbf{d}}$ is the reconstructed distance and $\bar{\mathbf{d}}$ the ground truth distance.

4.5.2.3 Global error

The global error $\mathbb{E}_{glob}$ is calculated similar as $\mathbb{E}_{abs}$, but after a rigid transformation of the entire reconstructed swarm (rotation and displacement) such that $\mathbb{E}_{abs}$ is minimized.

4.5.2.4 Local error

The local error $\mathbb{E}_{loc}$ is calculated as mean squared absolute error of subsections of 20 connected nodes after rigid transform of only those 20 nodes with their ground truth. This is

performed over the entire swarm.

4.5.2.5 Recall

The *recall*, $\mathbb{X}$, is the percentage of nodes reconstructed by the reconstruction algorithm. The adjusted recall $\mathbb{Y}$ and $\mathbb{Z}$ are the percentages of nodes that are reconstructed within a specified error condition. For $\mathbb{Y}$, this condition is met if at least 80% of the node's reconstructed distances have a relative error $<10\%$; or when more than 50% of the nodes reconstructed distances have a relative error of $<1\%$. The adjusted recall $\mathbb{Z}$ is similar to $\mathbb{Y}$, but includes that the neighboring nodes to which the relative error suffices this condition, should also fall in the category of $\mathbb{Y}$.

4.5.3   Simulation Results

The reconstruction performance metrics, as well as some of the output metrics of the ranging protocol are summarized in Table 4.2. All indicated metrics are median values of the iterations performed in similar scenarios. From each of the three scenarios, one of the reconstructions is shown in Fig. 4.11, together with the relative distance errors involved.

For all of the three scenarios, the majority of nodes are successfully reconstructed (recall $\mathbb{X}$ between 78% and 89%). This indicates that reconstruction based on the distance measurements from this ranging protocol is feasible.

The errors that are introduced are mainly due to the fact that the node measurements are performed when the nodes are non-static and the ranging procedure is not instantaneous.

The boxplots in Fig. 4.11 indicate the relative distance errors that are introduced throughout the entire process. The expected average distance error is calculated based on only the nodes' relative speed distribution as seen in Fig. 4.8 and the average latency within one cluster of nodes (46 ms). The protocol distance errors are the errors in the distance measurements of the ranging protocol compared to the ground truth. For the ground truth comparison, a snapshot of the actual positions are taken at a time halfway the ranging procedure. The reconstructed error indicate the relative distance errors between the reconstructed positions. Only the distances between the specific node pairs that actually aided in the reconstruction process (the RANSAC agreeing voters) are considered.

As seen in Table 4.2, throughout all simulations, the protocol coverage was between 73%-76% meaning that this fraction of nodes could successfully determine a distance between each other in the protocol phase. And the reconstructed coverage was between 63%-65%, meaning that on average, the reconstructed nodes are reconstructed based on distance measurements to 63%-65% of the neighbors within communication range, the other fraction is either not measured due to the protocol or not agreeing in the RANSAC voting process.

4.6   Conclusion

We have proposed and evaluated a solution for reconstructing the relative and absolute positions of a large quantity of miniaturized nodes in a go-with-the-flow exploration approach in enclosed environments. Our robust reconstruction solution relies only on communication

between nodes in the neighborhood of each other. It appears feasible for a non-static node swarm to estimate the positions and trajectory of the nodes. This is a step forward in creating miniaturized nodes that are to be used in difficult-to-access environments for obtaining a map of the nodes positions and sensing the interior of enclosed environments.

Our simulation chain illustrates that reconstruction algorithms can be used even if these rely on the assumption that the nodes are (quasi) static. The motion of nodes during a measurement cycle introduces an error, say, measurement noise, that with a proper system design can be small compared to dominant other types of noise experienced during the ranging procedure.

Future work includes dealing with the issues related to signal processing and overlap, as well as experiment with real hardware to test ranging accuracy. The verification of the protocol performance in experiments is foreseen as a future step in the EU Phoenix project.

4.7   Acknowledgement

References

[1] Z. Liu and Y. Kleiner,"State of the art review of inspection technologies for condition assessment of water pipes," Measurement, vol. 46, no. 1, pp. 1-15, Jan. 2013.

[2] J. M. M. Tur and W. Garthwaite, "Robotic devices for water main in pipe inspection: A survey," J. Field Robot., vol. 27, no. 4, pp. 491-508, 2010.

[3] T. Hao et al., "Condition assessment of the buried utility service infrastructure," Tunneling Underground Space Technol., vol. 28, pp. 331-344, Mar. 2012.

[4] E. Talnishnikh *et al.*, "Micro Motes: A Highly Penetrating Probe for Inaccessible Environments", in: *Intelligent Environmental Sensing*, edt. H. Leung, Springer Int. Publishing, 2015.

[5] M.N. Andraud *et al.*, "Exploring the unknown through successive generations of low power and low resource versatile agents", 20th Design, Automation and Test in Europe Conference (DATE 2017).

[6] Jules S. Jaffe *et al.*, "A swarm of autonomous miniature underwater robot drifters for exploring submesoscale ocean dynamics", NAT COMMUN 8, Article number: 14189 (2017)

[7] J. Lu, et al, "A small long-life acoustic transmitter for studying the behavior of aquatic animals", Review of Scientific Instruments 87, 114902 (2016); doi: 10.1063/1.4967941

[8] N. Patwari *et al.*, "Cooperative localization in wireless sensor networks", IEEE Signal Processing Magazine (54), July 2005

[9] I.F. Akyildiz, D. Pompili and T. Melodia, "Underwater acoustic sensor networks: research challenges", Ad Hoc Networks 3 (2005) 257-279

[10] J. Bachrach and C. Taylor, "Localization in Sensor Networks", in Handbook of Sensor Networks: Algorithms and Architectures, John Wiley & Sons, Inc., 23 Sept 2005

[11] G. Dubbelman *et al.*, "Robust Sensor Cloud Localization from Range Measurements", in *IEEE Int. Conf. Intelligent Robots and Systems*, Chicago, Illinois, USA, Sept. 2014.

[12] Erik H.A. Duisterwinkel *et al.*, "Mapping Swarms of Resource-Limited Sensor Motes: Using Distance Measurements and Non-Unique Identifiers", 2016 IEEE Symposium

Series on Computational Intelligence (SSCI), 6-9 December 2016, Athens, Greece (pp. 1-8).

[13] Erik H.A. Duisterwinkel *et al.*, "Robust Reconstruction of Sensor Swarms Floating through Enclosed Environments", Wireless Sensor Networks, Scientific Research, vol.10, no.1, January 2018

[14] Erik H.A. Duisterwinkel *et al.*, "Asymmetric Multi-Way Ranging for Resource-Limited Nodes", 8th EAI International Conference on Ad Hoc Networks, September 26-27, 2016, Ottawa, Canada

[15] J. Duderstadt, M.R. William, "Chapter 4:The derivation of continuum description from transport equations". In Wiley-Interscience Publications. Transport theory. New York. p. 218, 1979, ISBN 978-0471044925.

[16] S. Sadeghpour *et al.*, "Single-Element Omnidirectional Piezoelectric Ultrasound Transducer for under Water Communication", Eurosensors 2017, Vol. 1.

[17] G. Berkol *et al.*, "Design of a low-power ultrasound transceiver for underwater Sensor networks", 14th Conference on Ph.D. Research in Microelectronics and Electronics, PRIME 2018.

[18] S. Schlupkothen, A. Hallawa, G. Ascheid, "Evolutionary Algorithm Optimized Centralized Offline Localization and Mapping", 2018 International Conference on Computing, Networking and Communications (ICNC)

[19] Alessandro Pozzebon, "Bringing near field communication under water: short range data exchange in fresh and salt water", 2015 International EURASIP Workshop on RFID Technology (EURFID), IEEE, 2015

[20] Thomas Szabo, "Diagnostic Ultrasound Imaging: Inside Out", second edition, Academic Press, December 2013

[21] E.H.A. Duisterwinkel, Ph.D. dissertation, research in progress.

[22] M. P. Green, "N-way time transfer ('nwtt') method for cooperative ranging", Contribution 802.15-05-0482-00-004a to the IEEE 802.15.4a Ranging Subcomitee, July 2005.

[23] R. Hach, "Symmetric double sided - two way ranging", Contribution 802.15-05-0334-00-004a to the IEEE 802.15.4a Ranging Subcomitee, June 2005.

[24] H. Kim, "Performance Comparison of Asynchronous Ranging Algorithms", IEEE Global Telecommunications Conference, December 2009, Honolulu, Hawaii

[25] N.A.H. Puts, "Analysis and design of an ultrasound positioning system protocol for sensor swarms", MSc. dissertation, Eindhoven University of Technology, The Netherlands, 2016

[26] A. Varga and R. Homig, "An Overview of the OMNeT++ Simulation Environment", SIMUTools, March 03 - 07, 2008, Marseille, France.

5

Asymmetric Multi-Way Ranging for Resource-Limited Nodes

authors:
*Erik H. A. Duisterwinkel, Niels A. H. Puts, Heinrich J. Wörtche*

Cooperative localization in wireless sensor networks (WSN) is used in applications where individual sensor nodes cannot determine their location based on external contact, like e.g. GPS. The applications we focus on are the exploration and mapping of flooded cavities that are otherwise inaccessible or difficult-to-access, e.g. underground (oil-) reservoirs or industrial tanks for e.g. mixing. High levels of miniaturization are required for the nodes to traverse these cavities; nodes will have to be stripped down to a bare minimum. Ultrasound time-of-flight is used as radio communication is infeasible. Network topology is highly unpredictable and fast changing.

We present an asymmetric multi-way ranging protocol for these highly resource-limited, miniaturized, autonomous nodes. The specific set of constraints imposed by these applications, like the use of ultrasound, high latency, low data-rates, and non-static network topology is far-reaching and has not been studied before. Simulations of the protocol show trade-off's that can be made between ranging latency, signal overlap and overall energy budget.
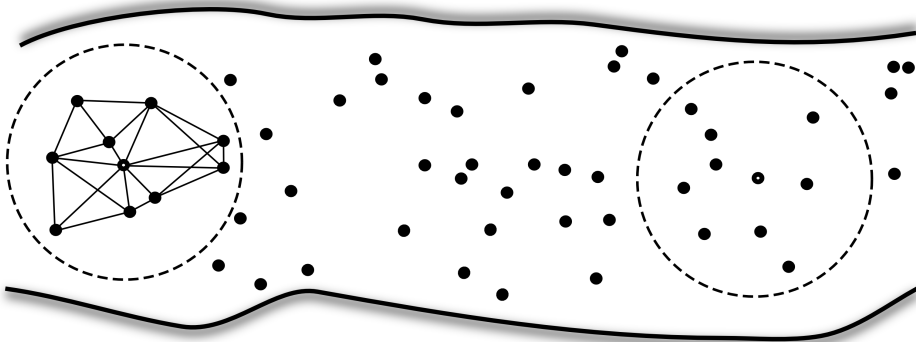
Figure 5.1: Swarm of exploring sensor nodes forming a network within an enclosed environment. Nodes perform ranging transactions to neighbouring nodes within their communication range for localization and further analysis offline.

5.1   Introduction

Underground cavities like (oil-)reservoirs, mines and geothermal sources, and industrial infrastructure like, pipelines, mixing tanks and reactors are systems which have in common that they are hard to access for in situ measurements of system structure, dynamics, conditions and integrity. A straight forward approach which has been proposed and investigated in [1, 2] is based on directly injecting large quantities of miniaturized sensor systems ('sensor motes') into the flooded system[1], let them go-with-the-flow in order to penetrate and to explore the system as visualized in Fig. 5.1.

For these sensor nodes to pass through the environment and explore it without disturbing it or interfering with the dynamics, the nodes need to be scaled down to the centimeter or millimeter scale, depending on the application. This highly limits the resources, like energy, processing and memory, that can be taken on board the nodes. Furthermore, antennas with those dimensions will only efficiently produce radio signals with wavelengths that have an extreme high attenuation in the liquids in these environments, effectively blocking all radio communication. However, ultrasound transducers at these scales do provide larger communication ranges in these environments, but yields other problems for stable and fast communication between dynamic nodes in enclosed environments [3]. Instead of relying on communication of data, measured data is stored in memory and made available for offline analysis after retrieving the nodes from the environment.

A crucial requirement is to obtain knowledge of the positions of the nodes while traversing the environment. Structural information can be extracted from this and sensor measurement of relevant parameters (e.g. temperature, pressure, salinity) can be visualized on a map. However, during operations, neither a distributed system of anchor points nor external beacons will be available. The concept of cooperative localization [4] can be used; nodes perform measurements like time-of-flight (TOF), angle-or-arrival (AOA) or received signal strength (RSS), to gain knowledge about the position of nodes relative to neighbouring nodes within

---

[1]In this paper, *mote* and *node* will be used interchangeably, as well as *system* and *environment*

communication range. This paper introduces a *ranging protocol* to determine distances between nodes using round-trip TOF that can be used for *localization algorithms* like in [5–7], but under the specific constraints that are found in these applications.

Besides the limitation on the nodes resources, localization is further hindered by the fact that network topology – the nodes' positions and their (sparse) connectivity to neighbouring nodes – is non-static and highly unpredictable. As it is not known where neighbouring nodes are positioned, omni-directional ultrasound is used for ranging measurements. This can be achieved using e.g. tube-shaped transducers as in [8].

Ultrasound is often used for ranging applications as the propagation speed is 10^5 times lower than that of radio, therefore allowing for larger timing errors. However, the low propagation speed in combination with the significantly lower data transmission rates (typically 2-40 kbits/sec or even lower in more challenging environments), introduces challenges that are less often seen in radio communication [3]. Latency in the ranging transactions makes that the movement of the nodes becomes significant in the distance determination. The low data-rates, in combination with the enclosed environment and non-static topology makes signal overlap a significant hinder.

In this paper we present a novel asymmetric multi-way ranging protocol, in which trade-offs are made between the energy budget, the ranging latency and the signal overlap to optimally use the on board resources for obtaining nodes positions in offline analysis for the above mentioned applications. Depending on the application or the state that nodes are in, these trade-off's can be adjusted to address the specific situation as good as possible.

The specific challenges in developing the ranging protocol for these applications are addressed in Sec. 5.2. In Sec. 5.3 the design of a ranging protocol is described that attempts to balance between all the parameters involved. In order to assess the suitability of the protocol for these applications, the protocol is simulated in a network simulator as described in Sec. 5.4. Important performance metrics that assess the specific goals are shown in the results section Sec. 5.5. Discussion and future work can be found in Sec. 5.6, the conclusion in Sec. 5.7.

5.2   Protocol design challenges

Traditional ranging protocols consist of three phases: a scanning phase, a ranging phase and a reporting phase [9]. In this paper, an attempt is being made to maximally reduce energy costs of ranging in the specific application cases described above.

Reporting of ranging measurements to neighbouring nodes is not performed as this would require extra node resources and data communication is challenging in this applications. Nodes only store measurements in their own memory.

The ranging-phase can be performed in a variety of methods. We chose for the concept of multi-way ranging (MWR), initially proposed in [10] as N-Way Time Transfer. It exploits the omnidirectional transmission and reception by using all received signals for determining distances between nodes, rather than only the signal between sender and one addressed receiver in e.g. two-way ranging (TWR) methods [11]. Therefore, the total amount of messages needed to complete a full ranging cycle using MWR scales linearly with the number of

nodes, instead of quadratically in TWR methods. It significantly reduces the energy required for performing the ranging procedure.

5.2.0.1   Control of ranging sequence.

As the network topology is non-static and connectivity sparse and fast-changing, it is not known which neighbouring nodes are within communication range. The simple sequence of events in traditional MWR [10], where node $i+1$ transmits a ranging signal after node i, cannot be easily controlled in these applications. An alternative method is proposed in Sec. 5.3 using a master-slave system.

5.2.0.2   Ranging latency.

The ranging latency of a single ranging transaction takes up to 3 ms when nodes are 1 meter apart (twice the propagation time, the message length and the processing time). A full ranging cycle within a large swarm, with all its individual ranging transactions, can easily take 100 ms. Depending on the movement of the nodes, a large latency significantly challenges the localization algorithm as the measured inter-node distances cannot be considered quasi-static. The latency should therefore be kept as low as possible.

Furthermore, from an energy perspective it is beneficial to reduce latency such that nodes are longer in a low-energy sleep state instead of an active listening/decoding state.

5.2.0.3   Signal overlap.

The low data-rates in combination with small inter-node distances in enclosed environments cause a significant amount of potential signal overlap. Signal overlap should be prevented as much as possible as it requires more energy and processing to filter and distinguish signals. The ranging protocol in Sec. 5.3 uses a time-divided communication scheme for determination of the distances to allow for reduction of signal overlap. The amount of bits transmitted should also be kept at a minimum to keep the message length as short as possible to reduce signal overlap.

5.2.0.4   Scanning phase.

As it is not known which nodes are within communication range, often in ranging protocols, a separate scanning phase is initiated before the ranging phase. In this phase, nodes determine which neighbouring nodes are within communication range to determine which nodes to perform ranging measurements to.

Such an additional scanning phase adds to the energy budget. In this paper, the scanning phase is omitted and solved by addressing all nodes by a non-unique calling identifier. It causes a trade-off between ranging latency and signal overlap.

5.3   Protocol

This section introduces a modified version of the regular multi-way ranging protocol to deal with the specific limitations in the usage of ultrasound in a non-static network topology with
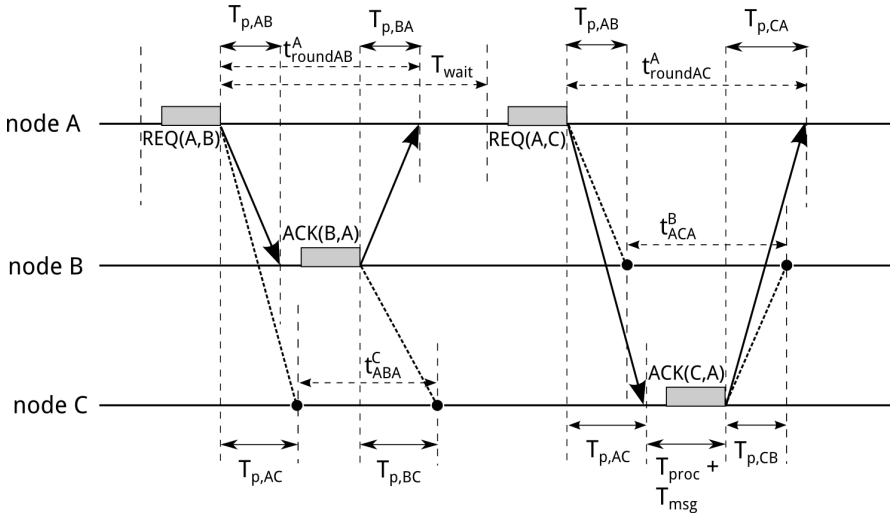
Figure 5.2: One ranging cycle of master node A and slave nodes B and C. The master node transmits request (REQ) signals and slave nodes respond with and acknowledgement (ACK) signal if it is addressed to them. Knowledge of timing information $t^A_{roundAB}$, $t^A_{roundAC}$, t^C_{ABA} and t^B_{ACA} and the fixed value of the processing time T_{proc} is sufficient to determine the propagation times $T_{p,AB}$, $T_{p,AC}$ and $T_{p,BC}$ between the nodes.

highly resource-limited nodes. It will also address the challenge of finding proper trade-off's between e.g. ranging latency and signal overlap. It is important to notice that these trade-off's can be adjusted based on the specific environments or the specific situation that nodes are in.

5.3.1   Asymmetric multi-way ranging

Instead of traditional, 'symmetric', MWR as introduced in [10], here the ranging procedure is controlled by *master nodes* that send request (REQ) signals to its neighbouring nodes that then become *slave nodes* and respond with an acknowledgement (ACK) signal. The communication scheme that is used is illustrated in Fig. 5.2. The *cluster of nodes* that is formed by this master node and the slave nodes is illustrated in Fig. 5.3.

In a ranging transaction between master node A and slave node B, the timestamps of transmission and reception at node A provides knowledge about the round-trip TOF between nodes A and B, denoted as $t^A_{roundAB}$. The node's internal processing time T_{proc} and signal message time T_{msg} are known beforehand and are fixed, therefore, the round-trip propagation time, $T_{p,AB} + T_{p,BA}$, between A and B can be estimated. After this ranging transactions, master node A performs a similar transaction to node C and the other nodes within the cluster.

Since the nodes A, B and C are within each others communication range, also the nodes that are not addressed in the ranging transactions receive the signals. The time difference t^C_{ABA} between the arrival of REQ(A,B) and ACK(B,A) at node C and the time difference t^B_{ACA} between the arrival of REQ(A,C) and ACK(C,A) at node B, can be used to calculate
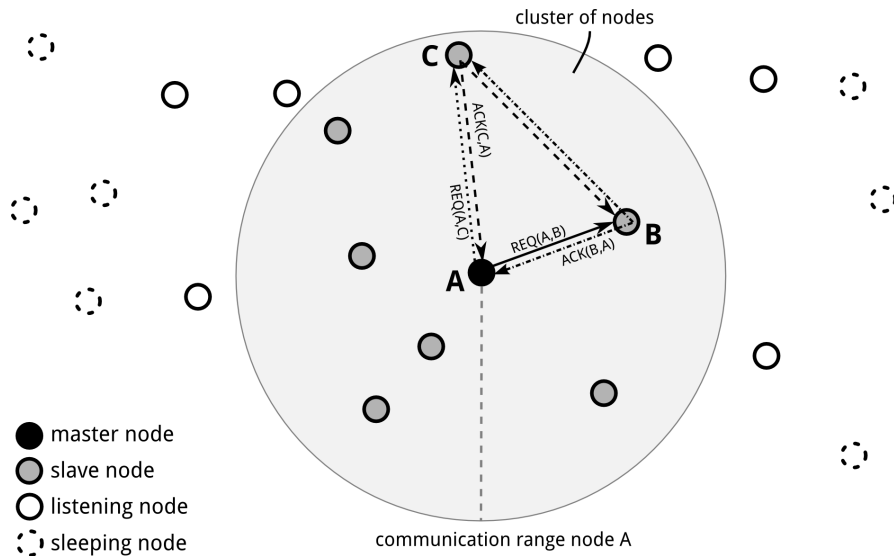
Figure 5.3: Master node A initiates the ranging process to node B and C. The nodes within communication range become slave node and respond to REQ signals with an ACK signal. All nodes within the respective communication range receive the signals and store them: in offline analysis they can be used to determine round-trip TOF between nodes. The nodes outside the cluster will only receive ACK signals. In this figure, not all signals (arrows) are drawn.

the propagation time between nodes B and C using:

$$T_{p,BC} + T_{p,CB} = t_{ABA}^C + t_{ACA}^B - 2(T_{\text{proc}} + T_{\text{msg}}) \tag{5.1}$$

5.3.2 Picking the master node

The role of master node is being alternated between all nodes in the network. The advantage of this is that the power consumption is distributed evenly over all nodes (master nodes transmit more signals) and clusters are more distributed over the swarm.

Within the time frame of one sample T_{sample}, in which a complete ranging cycle is completed for all nodes, the role of master node is chosen randomly. This is performed by having all nodes at the beginning of a sample chose a random delay time T_D. Nodes become master when their sample timer t_s, that is set to zero at the beginning of a sample, trespasses $t_s > T_D$. Nodes become slave node when before t_s reaches T_D, it receives any REQ signal from a master node. The master node initiates the ranging transactions as described above, thereby forming a cluster of nodes as in Fig. 5.3.

As seen in Fig. 5.1, throughout the entire network, several of these clusters are formed in which ranging transactions are performed. Every sample, these clusters change based on which nodes have become master node.

5.3.3 Scanning the slave nodes

Within one ranging cycle, the master node should send a request to all slave nodes in the cluster, but it is beforehand not known which nodes are within communication range. Regular scanning techniques depend on the availability of sufficient bandwidth, processing power or time to perform broadcasting.

In this work we propose for the master node to initiate the ranging transactions to all possible hardware addresses. But as the total amount of nodes in the network can be very large and the connections are sparse, this will be very inefficient as most requests remain unanswered. Instead of requesting to the hardware's *unique identifiers* (UID) the master node requests to highly abbreviated *calling identifiers* (CID). The master node only initiates n_f times a ranging transaction to CID $= \{0, 1, ..., n_f\text{-}1\}$. Slave nodes will respond if and only if their unique hardware identifier suffices

$$\text{mod}(\text{UID}, n_f) = \text{CID} \tag{5.2}$$

As multiple nodes will have an identical CID, the probability arises that multiple nodes will respond to the same request. If the ACKs of the responding slave nodes do not overlap such that the signals cannot be distinguished and decoded anymore at the receiving node, the determination of the round-trip TOF of each of them can still be performed. Parameter n_f can be chosen both offline as online to adjust for the amount of neighbouring nodes and the total signal overlap.

In order to receive all possible ACK's, the master node will wait $T_{\text{wait}} = 2T_{\text{p,max}} + T_{\text{proc}} + T_{\text{msg}}$ after transmission of a REQ before it sends a request with a next CID. Here, $T_{\text{p,max}}$ accounts for the propagation time required to reach the end of the (expected) communication range.

After all ranging transactions have been performed, the nodes will go into a low-energy sleep mode to await the start of the next sample. When $t_s > T_{\text{sample}}$, nodes will internally initiate a new sample. The sample is initiated in a sleep mode and nodes will wake up upon reception of any signal (using e.g. a threshold detection). It then starts a listening mode in which it can decode incoming signals. Before a master node starts with the first CID, it can transmit a short signal to wake up the neighbouring nodes.

5.3.4 Reducing the latency

As seen in Fig. 5.3, the nodes just outside the communication range of the master node do not become slave node and will have to wait for itself to become master node, or will have to wait for a node within its communication range to become one.

In order to speed up this process and have the network-wide ranging cycle end sooner, an avalanche effect is induced. Nodes that receive ACK signals without having received REQ signals are likely to be just outside an already formed cluster. Their remaining delay time before they become master node is reduced by a factor $M_{\text{avalanche}}$ at the reception of any ACK signal until they become master or slave.

This reduction of the delay time T_D induces an avalanche effect throughout the network such that all nodes become either master or slave within less time after each other, therewith, reducing the ranging latency throughout the network.
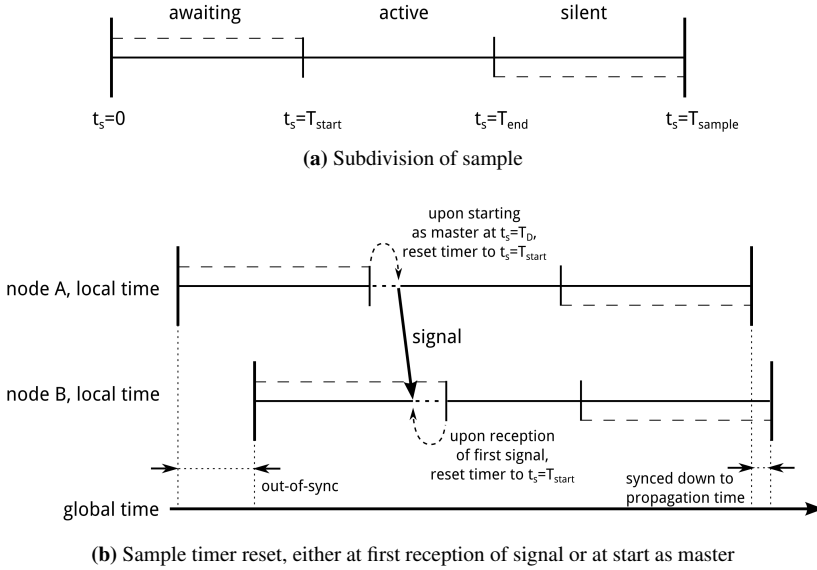
5.3.5 Synchronization

Absolute synchronization is not required for determining distances as all distances are obtained using a round-trip TOF measurement. It is however beneficial to have nodes synchronized to a level in which samples are aligned such that the avalanche effect introduced in Sec. 5.3.4 allows nodes to sleep for the majority of the sample time instead of responding to nodes that are still in the previous or already in the next sample.

For this reason, in order for connected nodes to remain in the same sample, it is proposed to subdivide a sample on the node level into timeslots as illustrated in Fig. 5.4.a. The random delay time is chosen from a uniform distribution within the range $T_D \in (T_{\text{start}}, T_{\text{end}})$ or the *active time period*. The internal sample timer t_s is reset to $t_s = T_{\text{start}}$ when becoming master; or, at reception of the first signal (any REQ or ACK) in the sample, as illustrated in Fig. 5.4.b. This will assure that connected nodes remain synchronized to the sample level, as long as (groups of) nodes have not been disconnected from each other. In the *awaiting period* nodes do not become master and can only receive signals, in the *silent period*, nodes have already received their first signal or already became master node.

In our work, the three time periods are chosen to be of equal length, i.e. $T_{\text{start}} = \frac{1}{3}T_{\text{sample}}$ and $T_{\text{end}} = \frac{2}{3}T_{\text{sample}}$. Note that these periods do not indicate when a node is asleep or in which mode it is in.

5.4 Simulations

The protocol implementation is simulated in OMNeT++ network simulator [12, 13].

(a) Subdivision of sample

(b) Sample timer reset, either at first reception of signal or at start as master

Figure 5.4: Subdivision of sample in three parts: awaiting, active and silent. Each sample, the start delay that determines when to become master is randomly chosen within active period (uniform distribution). Upon first reception of signal in the sample, or, upon becoming master node, the sample timer t_s is reset to $t_s = T_{\text{start}}$.

In order to simulate a dynamic swarm of nodes that passively flow through an enclosed environment, we use a flow simulator to generate the nodes positions over time [14]. The positions are generated based on tracer positions in a fluid flow in a 9 by 8 meter 2-D tank-like environment with an inlet an outlet. The positions of the $N = 200$ nodes at the beginning and end of the simulation time are illustrated in Fig. 5.5.

The average node speed throughout the simulation is 0.20±0.17 m/s with a maximum of 0.80 m/s. The communication range is set to a fixed 1 meter and results in an average node density of 9.8±3.7 neighbouring nodes within the communication range. The clock frequency offset is set to 100 ppm and is fixed throughout the simulation.

The sample time is set to $T_{\text{sample}} = 1$ sec and the amount of CIDs is in this paper is swept between: $n_f = \{8, 16, 32\}$. The ultrasound transmission rate is set to 40 kbit/sec. The avalanche induction is studied by sweeping $M_{\text{avalanche}} = \{1, 2, 4, 8\}$ in which $M_{\text{avalanche}} = 1$ means no induced avalanche.

The output of the simulations consist of the data that are being stored on the nodes internal storage: the messages sent and received, the timestamp at transmission/reception, the timestamp when new samples start and for research purposes also the internal states the nodes are in and at what specific time.

(a) t=0 sec　　　　　　　　　　　　　**(b)** t=16 sec
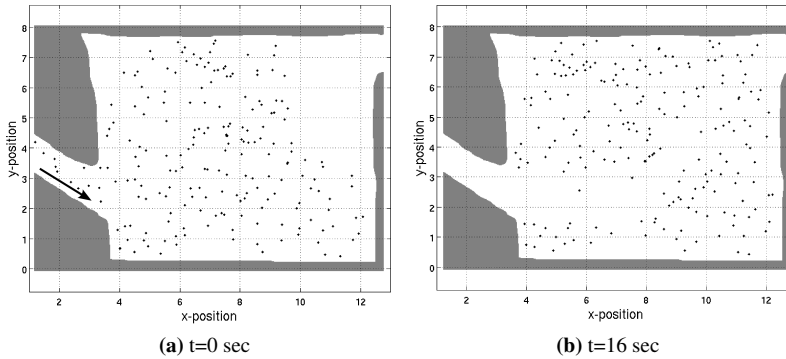
Figure 5.5: Nodes positions throughout simulation in 2D tank-like environment with injection flow from left (indicated by arrow).

5.5　Results

The protocol is analysed based on several performance metrics that assess the design goals for the ranging method. One is the ranging latency within a swarm versus the signal overlap, second is the energy usage of an individual node. And as last, the fraction of the theoretical amount of possible distances that are determined using this protocol: the coverage.

Fig. 5.6 shows for the simulated datasets with $n_f = 16$ an overview of the amount of nodes in a specific state over a single sample. Figure 5.7 shows the main performance metrics of the protocol in a simulated network as described in Sec. 5.4 and discussed next.

5.5.1　Latency versus signal overlap

From Fig. 5.6 it can be seen that the induced avalanche effect assures that the network finished a single ranging cycle sooner. In this sample, the maximum latency goes from 360ms for $M_{\text{avalanche}} = 1$ (no avalanche) down to 150ms, 120ms and 110ms for $M_{\text{avalanche}} = \{2, 4, 8\}$, respectively. As a reference; within a single cluster, the ranging latency is $n_f T_{\text{wait}} = 43$ms.

Fig 5.7a shows the average latency between all ranging transactions in a ranging cycle. Increasing the avalanche effect ($M_{\text{avalanche}}$) yields a smaller latency. The latency of the full ranging cycle is approximately between 4 to 5 times larger as the average latency between the transactions.

Fig 5.7a also shows the average fraction of signals that are received with overlap with another signal. Increasing the avalanche effect and reducing the latency inevitably increases the signal overlap.

At lower values of n_f, the latency drops quicker, but signal overlap is higher; less CIDs are scanned but more nodes will respond to the same REQ signal. There is a clear trade-off between latency and signal overlap.
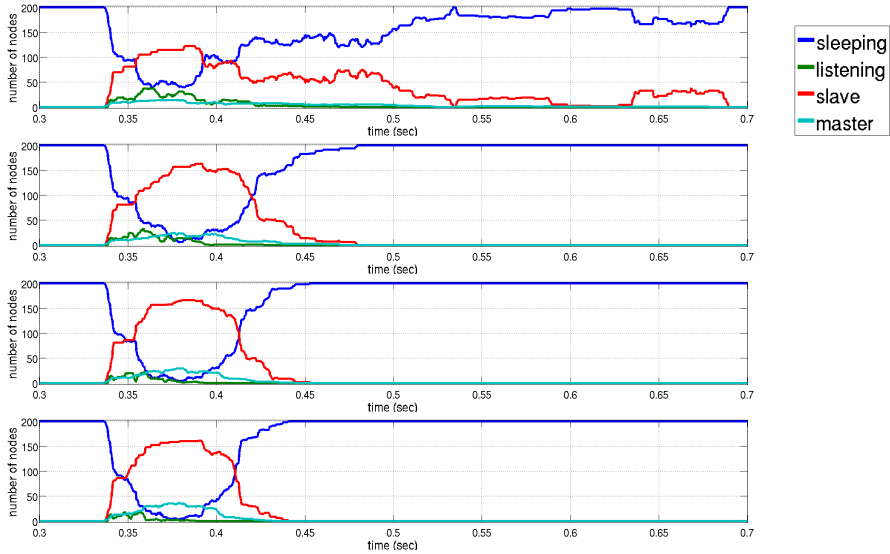
Figure 5.6: Total number of nodes in specific states through first sample. From top to bottom the avalanche parameter: $M_{\text{avalanche}} = 1, 2, 4, 8$. The top graph, $M_{\text{avalanche}} = 1$ means there is no induced avalanche effect. In these simulations, the number of CIDs $n_f = 16$.

5.5.2 Energy efficiency

The energy efficiency in this paper mainly focusses on the nodes' awake time and the amount of signals transmitted (and related to that the amount of signals received and stored). The awake time is defined as the time not spend in the low-energy sleep state, but rather in an active signal transmission or receiving/decoding state.

Fig. 5.7b shows the node's average awake time per sample and the average amount of signals transmitted per node per sample (master and slave nodes together). As the latency decreases with increasing $M_{\text{avalanche}}$, so does the time that nodes need to be awake. With increasing avalanche effect, the number of signals required for transmission increases slightly as more nodes will become master node.

Both the awake time and the number of signals transmitted increase with increasing n_f as more CIDs will have to be transmitted and decoded.

5.5.3 Coverage

The coverage can be defined as the fraction of connections (node pairs that are within each others communication range) for which the ranging procedure yields sufficient information to determine a distance measure. Since only distances can be calculated within a cluster, the coverage will be lower than 100% as not all connections can fall within a cluster. Throughout all simulations, the coverage was between 86%-89%.

Even though for the other 11%-14% no distances can be determined using RT-TOF, the basic connectivity information is available: the received ACK signals that did get received

(a) Ranging latency and signal overlap

(b) Awake time and signals transmitted
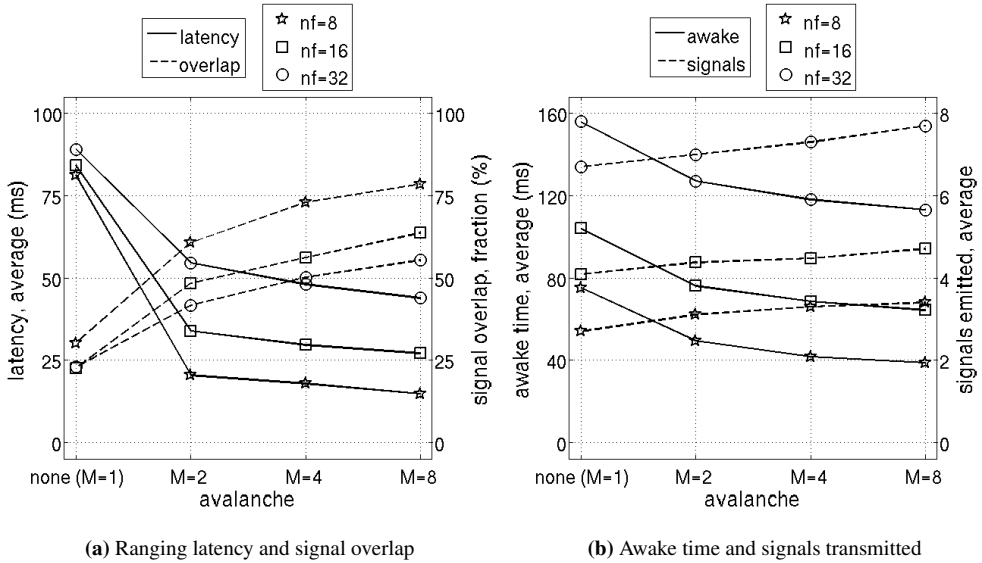
Figure 5.7: Average values of perfomance metrics of simulated ranging protocol using different input parameters n_f and $M_{\text{avalanche}}$. Clear trade-offs are visible.

by the nodes outside the cluster, provides information on which nodes where within their communication range. The localization algorithm can use this information to its advantage.

The coverage does not need to be 100% to localize the entire swarm. In fact, for example in [6], studies are performed where localization is stress-tested on e.g. the loss of large amounts of connections. Also note that each sample, different clusters are formed such that this group of 11%-14% of the connections is different for each sample.

5.6   Discussion and future work

Although the current implementation of the ranging protocol has been simulated over a relative short measurement time. Simulations using extremely large clock deviations of up to 100 000 ppm have been tested and show good alignment of samples over the simulation time of 16 seconds (not shown here). As long as the network is sufficiently connected and not disjoint, the avalanche effect can keep the nodes' internal clock synchronized within a fraction of the sample time T_{sample}.

The simulations in this paper have been performed in a 2-D environment. Although this protocol can be directly used in 3-D, the induced avalanche effect will have quantitatively a slightly different result as the ones presented here. No qualitative differences are to be expected.

Instead of scanning all possible CIDs, all nodes can actively record which nodes it has seen in the past. Upon becoming master node, instead of scanning all available CIDs, the master

node can scan the UIDs of nodes that it has seen in the previous (several) sample periods. This will reduce signal overlap and can reduce the amount of required signals for transmission.

The protocol can fairly easily be adjusted to also account for disjoint networks coming together such that they can become synchronized up to the sample level. This is part of future work.

5.7   Conclusion

This paper illustrates the challenges involved in performing round-trip TOF in a swarm of autonomous nodes without external contact in an unpredictable and dynamic topology with sparse connectivity. The applications require high levels of miniaturization of the nodes and introduce a specific set of constraints and challenges that has not been researched before. A novel asynchronous multi-way ranging protocol has been presented to allow round-trip TOF measurements. Control of the ranging transactions can be performed by master nodes that initiate them to their neighbouring slave nodes. Master nodes are assigned at random each time a new sample starts.

The latency between ranging measurements in the entire swarm can be reduced by inducing an avalanche effect of nodes that become master node. The avalanche effect also reduces the required time for the nodes to be actively listening for signals and allows for synchronization down to a fraction of the sample time.

The trade-off's that are involved in this ranging protocol are a direct consequence of the application: the need for resource-limited nodes, the use of ultrasound and the unpredictable and fast-changing network topology with sparse connectivity. Getting insight in these trade-off's allow for adjusting the ranging protocol based on the specific circumstances that nodes are in. In [2], this exploration method and the ability of nodes to adjust for specific circumstances, is further explored.

5.8   Acknowledgement

References

[1] E. Talnishnikh *et al.*, "Micro Motes: A Highly Penetrating Probe for Inaccessible Environments", in: *Intelligent Environmental Sensing*, edt. H. Leung, Springer Int. Publishing, 2015.

[2] EU Horizon 2020 FET-Open project: PHOENIX. www.phoenix-project.eu

[3] I.F. Akyildiz, D. Pompili and T. Melodia, "Underwater acoustic sensor networks: research challenges", Ad Hoc Networks 3 (2005) 257–279

[4] N. Patwari *et al.*, "Cooperative localization in wireless sensor networks", IEEE Signal Processing Magazine (54), July 2005

[5] J. Bachrach and C. Taylor, "Localization in Sensor Networks", in Handbook of Sensor Networks: Algorithms and Architectures, John Wiley & Sons, Inc., 23 Sept 2005

[6] E.H.A. Duisterwinkel, Ph.D. dissertation, research in progress.

[7] G. Dubbelman *et al.*, "Robust Sensor Cloud Localization from Range Measurements", in *IEEE Int. Conf. Intelligent Robots and Systems*, Chicago, Illinois, USA, Sept. 2014.

[8] H Li, K W Jung and Z D Deng, "Piezoelectric transducer design for a miniaturized injectable acoustic transmitter", Smart Mater. Struct. 24 (2015)

[9] H. Kim, "Performance Comparison of Asynchronous Ranging Algorithms", IEEE Global Telecommunications Conference, December 2009, Honolulu, Hawaii

[10] M. P. Green, "N-way time transfer ('nwtt') method for cooperative ranging", Contribution 802.15-05-0482-00-004a to the IEEE 802.15.4a Ranging Subcomittee, July 2005.

[11] R. Hach, "Symmetric double sided - two way ranging", Contribution 802.15-05-0334-00-004a to the IEEE 802.15.4a Ranging Subcomittee, June 2005.

[12] N.A.H. Puts, "Analysis and design of an ultrasound positioning system protocol for sensor swarms", MSc. dissertation, Eindhoven University of Technology, The Netherlands, 2016

[13] A. Varga and R. Homig, "An Overview of the OMNeT++ Simulation Environment", SIMUTools, March 03 – 07, 2008, Marseille, France.

[14] H.H.W.J. Bosman (2015), for the adaptation and extension of Daniel V. Schroeder's
Java based flow simulator (2013)

6

Sensor Motes for the Exploration and Monitoring of Operational Pipelines

Erik H.A. Duisterwinkel, Elena Talnishnikh, Dirkjan Krijnders and Heinrich J. Wörtche

We report on a first field-test in which miniaturized sensor motes were used to explore and inspect an operational pipeline by performing *in situ* measurements. The spherical sensor motes with a diameter of 39 mm were equipped with an Inertial Measurement Unit (IMU) measuring 3D acceleration, rotation and magnetic field, as well as an ultrasound emitter. The motes were injected into the pipeline and traversed a 260 m section of it with the flow of water. After extraction of the motes from the pipeline, the recorded IMU data were read-out for offline analysis.

Unlike dead-reckoning techniques, we analyse the IMU data to reveal structural information about the pipeline and locate pipe components like hydrants and junctions. The recorded data show different and distinct patterns that are a result of the fluid dynamics and the interaction with the pipeline. Using the magnetic data, pipe sections made from different materials and pipe components are identified and localized. A preliminary analysis on the motes' interaction with the pipeline shows differences in pipe wall roughness and locates structural anomalies.

The results of this field-test show that sensor motes can be used as a versatile and cost-effective tool for exploration and inspection of a wide variety of pipelines.

Figure 6.1: Xploring WiseMote™ (XWM).

6.1   Introduction

Pipeline distribution systems are a backbone technology for operating industries and society. Various industries are looking into optimization and smart maintenance of distribution pipelines worldwide. Pipeline infrastructures represent major investments, e.g. the Dutch freshwater distribution infrastructure comprises 116,000 km of pipelines, representing a value of 20 billion EUR [1]. A majority of these pipelines were installed in the 1950s, 1960s and 1970s and their expected lifetime is 50-60 years. A tremendous effort will be required to sustain a high-quality network in the next decades [1].

Regular inspection of the pipe conditions and integrity is essential to ensure a long service life, reduce risk of failures and help in prioritizing the replacement order of pipe sections and components. A variety of passive and active in-pipe inspection devices have been developed over the past years [2–4]. However, most of the techniques are too expensive to be used in pipelines transporting low-cost commodities like water, let alone in the highly ramified smaller pipelines in domestic areas. Moreover, a significant part of the water distribution network had *ad hoc* repairs in the past that are only partially documented or not documented at all. Physical passage of such in-pipe inspection devices cannot be guaranteed due to the lack of knowledge on the location and the type of pipes and pipe components, bends, flanges and valves. Cost effective and readily applicable exploration tools capable of providing this information are, therefore, in high demand.

In this paper we present results of a field-test in which a modified version of miniaturized sensor motes, Xploring WiseMotes™ (XWMs), are used to explore and to inspect the conditions and integrity of a pressurized water pipeline in operation. The motes were directly injected into a fire-extinguishing water pipeline, where they float with the water through the pipeline and recorded information on the dynamics of the mote using an Inertial Measurement Unit (IMU).

Unlike dead-reckoning techniques that attempt to track positions of an object based on IMU-data [5], we analyse XWMs linear acceleration, rotation, and the sensed magnetic field in order to reveal structural information of the pipeline and to study mote dynamics and their interaction with the environment. The recorded data show a wealth of information, allowing us to explore the pipeline – e.g. to identify and localize pipe-components and pipe materials – and to inspect pipe conditions. We demonstrate this novel technique as cost effective and readily applicable exploration tool to complement or replace existing inspection techniques.

In Sec. 6.2 we describe the conceptual approach underlying the XWM technology, in Sec. 6.3 we present the mote hardware and its modification with the ultrasound extension that has been developed for the field-test. The experimental details are described in Sec. 6.4. Our findings from the data are stated in Sec. 6.5. A discussion provided in Sec. 6.6 and the paper is concluded in Sec. 6.7.

6.2   XWM concept

The XWM technology is based on the concept of injecting miniaturized sensor motes into a flooded environment and let them go-with-the-flow of the medium filling the environment, while performing and storing measurements [6, 7]. XWMs have been designed for operating in environments such as distribution pipelines and industrial mixing tanks. Standard XWMs are equipped with an IMU measuring the 3D linear acceleration and rotation of the mote (dynamics data) and the local magnetic field. The XWM density can be adjusted to the specific density of the carrying medium.

Data stored on an XWM are read out and analysed after retrieval of the mote from the environment. The dynamics data provide insight into processes taking place inside mostly closed and difficult-to-access environments. The magnetic data serve as a reference either to align the dynamics data in respect to the Earth's magnetic field or to identify and locate magnetic anomalies in the environments, such as metallic structures or magnetic beacons.

Experiments performed in an industrial mixing tank revealed the potential to get a better understanding of the small-scale physics governing the mixing process by measuring the mote dynamics [8]. Collisions of the motes with containment walls, mechanical instrumentation or obstacles are indicated by sharp peaks in the accelerometer data. These peaks are typically milliseconds wide with amplitudes exceeding several times g-force. Similar to particle induced scattering processes, the collision statistics and the related momentum transfer reflect the structure of the environment, but also surface and material conditions. If the mote density matches the medium density or the density of a specific phase for multi-phase media, the dynamics data in-between the collisions provide a measure of the local medium (phase) dynamics, reflecting e.g. laminar or turbulent flow patterns.

Correlation analysis of data measured by XWMs with different density and mass distribution, provides the sensitivity to address specific system or process aspects. To gain the required statistics, XWMs are either operated in swarms or a single XWM is operated in repetitive cycles.

The XWMs are, thus, a passive inspection and exploration tool that is significantly different than existing technologies. Additionally, they have a low application threshold. XWMs

Table 6.1: Specifications of the modified XWMs.

IMU sensor	ST LSM9DS0
Functionality	triaxial gyroscope
	triaxial accelerometer
	triaxial magnetometer
Max. sample rate and range	760 Hz; 245-2000 deg/sec
	1600 Hz; 2-16 g
	100 Hz; 2-12 gauss (10^{-4} T)
Data resolution	16 bit
Memory size	2 Gbit
Shell material	Poly-amide 12 (PA 12)
Battery	0.5 Wh rechargeable Lithium
Communication	micro-USB
Measurement time	1 hour (at max. sample rate)
Time synchronisation error	<1 second per hour
Diameter	39 mm
Density	adjustable (0.62-2.2 kg/dm^3)
US emitter	131 kHz

complement available techniques and provide the capabilities of a data collector tool for *in situ* measurements performed during system operation.

6.3   Standard and modified XWM hardware

XWMs are spherical sensor motes with a diameter of 39 mm. The shell consists of two halves that are bolted together with a rubber O-ring in between. The shell is made from 2.7 mm poly-amide 12 (PA 12).

Fig. 6.1 shows a standard XWM with an earlier design of the shell. Fig. 6.2 shows a modified XWM equipped with an ultrasound transducer and a stronger shell, distinguished by a rim framing the transparent half of the shell and providing increased strength to the O-ring groove in the XWM equator plane. The technical specifications are listed in Table 6.1.

The shells have proven to be mechanically robust and chemically inert and tested to be water-tight in excess of 30 bar of water pressure [9]. The gross volume of the XWM is 31 cm^3 and its weight without ballast pieces 20 g, resulting in a minimum density of 0.65 kg/dm^3. The density can be adjusted up to 2.2 kg/dm^3 by adding pre-fabricated brass ballast pieces.

In the centre of the PCB is the IMU chip (ST LSM9DS0) [10] that can be sampled at an adjustable rate and range. Throughout the field-test, sampling rates were set to the maximum values listed in Table 6.1. These settings provided the highest possible resolution to detect short acceleration peaks and subsequent changes in rotation caused by collisions. The magnetic sampling rate allowed to achieve a spatial resolution of about 1 cm at flow speeds of order 1 m/s.
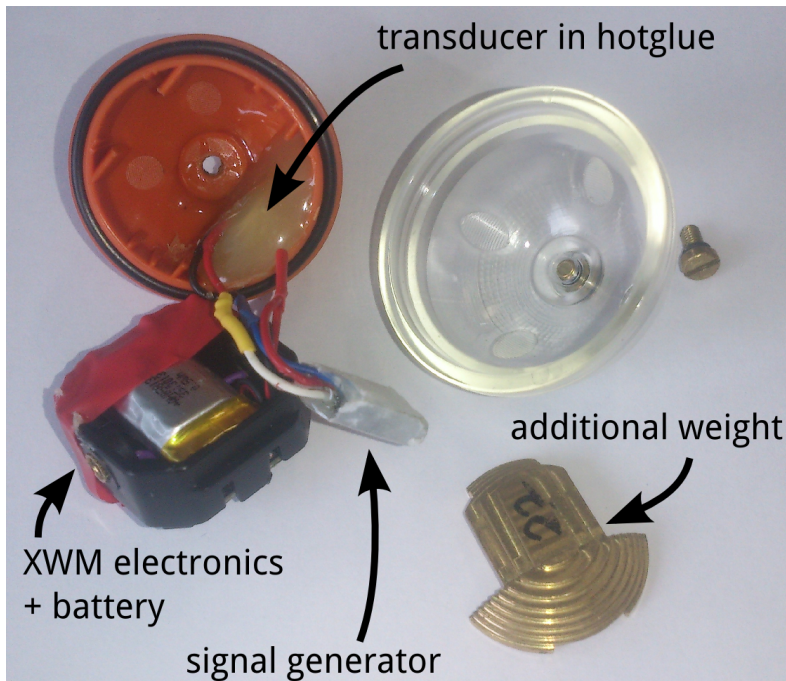
Figure 6.2: Modified XWM equipped with an ultrasound transducer embedded in hot melt glue.

IMU data are stored on a flash memory chip that can be read-out using (micro-)USB interface.

6.3.1 Modified XWM with ultrasound emission

Upon request by the operator of the pipeline, ultrasound emission was implemented to enable mote detection inside of a hydrant in case a mote could not be retrieved by standard procedure. Additionally, it allows to verify arrival of motes at specific locations in the pipeline.

In the modified XWM the ultrasound signal was generated by a $6.35{\times}6.35$ mm^2 tube-shaped piezo-electric transducer driven by a 6 Vp-p square wave signal at 131 kHz, its hoop-mode (or 'breathing'-mode) resonance frequency. As seen in Fig. 6.2, the transducer was embedded in hot glue sticking to the interior of the shell to provide good acoustic coupling with the water. To achieve maximum detectability the XWMs continuously emitted ultrasound. This ensured that the detected signal could be visualized in the frequency domain without trigger or signal processing and could be easily observed using a hand-held scope. For details of the ultrasound emitter design we refer to [11].

The modified XWMs used in the field-test had a density between 1.01-1.03 kg/dm^3, slightly exceeding neutral buoyancy in water (see also Table 6.2). Due to the ballast piece the centre of mass (CoM) was shifted, creating a tendency for the mote to point the ultrasound transducer upwards. For the field-test this was expected to ease the ultrasound detection of
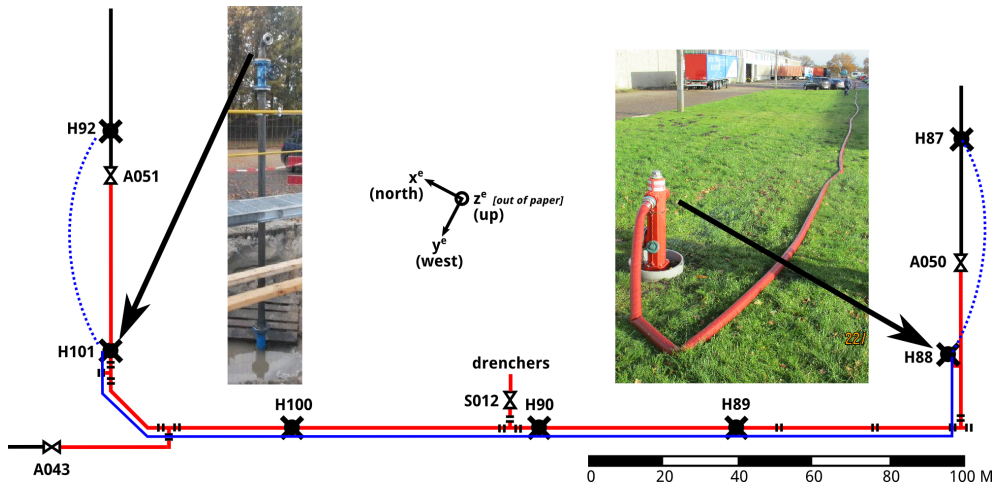
Figure 6.3: Schematic map of the field-test location. The left inlay photograph shows the injection pipe (position H101) without hose, the right inlay shows hydrant H88 with a hose to hydrant H87.

the XWM underneath a hydrant.

Operation of the XWMs IMU was limited by the on-board memory to 1 hour at highest sampling frequencies, operation of the embedded ultrasound transducer was limited to 4 hours by the 0.5 Wh Lithium-battery.

In order to detected the emitted ultrasound an identical transducer was used to build a simple small hydrophone [11]. The transducer was connected to an instrumentation amplifier with an amplification factor of 34 dB. The single-ended output was connected to a laptop controlled hand-held scope (PicoScope 2260) via a coaxial cable of 3 metres, which allowed to lower the hydrophone into a hydrant.

6.4   Field-test

The field-test was performed at Emmtec Industry and Business Park together with Emmtec Services B.V in November 2016. The XWMs were injected in a fire-extinguishing water pipeline, which is schematically shown in Fig. 6.3. The pipeline with an outer diameter of 168.3 mm (DN150) was originally installed underground at a depth of 1.20 m in 1960 (56 years old).

The typical water pressure in the pipeline is 9 bar (0.9 MPa) and no flow is present until hydrants are opened. The pipeline has both visible and hidden structural components such as hydrants (visible), flanges (buried) and t-joints (buried). Throughout the years, various sections of the pipeline have been replaced. A complete documentation of maintenance and repair logs is not available at present. Upon agreement with EmmTec specific objectives of the field-test were established as follows:

Table 6.2: XWM mass configurations with their respective specific densities used in the field-test. The specific density is calculated relative to the density of water at 10°C.

XWM config.	specific density	remarks
A	1.016±0.016	smaller, thinner shell
B	1.005±0.005	minimum density
C	1.013±0.005	
D	1.028±0.009	different mass distribution

- develop and establish injection and retrieval procedures of motes in and from pipelines;

- detect XWMs in hydrants using ultrasound;

- identify and locate pipe sections and components;

- examine conditions and integrity of pipes and pipe junctions;

- detect flow conditions.

In order to inject and retrieve XWMs, hydrant H101 had been replaced by a DN80 vertical injection pipe, as seen in Fig. 6.3. An ultrasonic on-the-line flow meter was installed on the injection pipe. The XWMs were inserted in the injection pipe with no flow in the system.

Valves A051, A043 and A050 (see Fig. 6.3) were closed to isolate the test section from the rest of the pipeline system. Hydrant H88 was kept open for releasing the water. In order to provide a water flow, a fire hose was connected from hydrant H92 to the injection pipe (dashed line in Fig. 6.3). Then, hydrant H92 was opened to provide a water flow, transporting the XWMs from the injection pipe to hydrant H88. The water was released from hydrant H88 into the environment, while the XWMs remained in the system as they could not pass through the open valve of H88.

After a time period estimated to exceed the travel time of XWMs between location H101 and H88, the water flow was stopped by closing hydrant H92. The injection pipe with the attached fire hose was disconnected from the hydrant H92. Hydrants H87 and H88 were connected by a different fire hose (a second dashed line on right in Fig. 6.3). Then hydrant H87 was opened providing the reverse flow. The reverse water flow was also freely released from the fire hose attached to the injection pipe into the street. The water flow was cut off once XWMs were retrieved by closing the hydrant H87. XWMs data were uploaded to a computer and motes were configured for another run.

6.4.1 Execution

Four experimental runs were performed using different combinations of XWM configurations and flow speeds resulting in 6 recorded datasets. In the initial two runs a single XWM was injected, in the last two runs two XWMs were injected simultaneously. Mote configurations and flow speeds as measured in the injection pipe are listed in Table 6.2 and Table 6.3. The flow speed was adjusted by the opening of hydrant H92 for the forward flow and hydrant H87 for the reverse flow. The flow speed could only be controlled with limited accuracy within the range of 1-5 m/s in the DN80 pipe. Due to the larger cross-section of the underground

Table 6.3: Overview of experimental datasets including XWM configurations (taken from Table 6.2), the nominal flow velocity representing the longest sustained velocity measured in the injection pipe (DN80); the travel duration that indicates the period between flow start and mote arrival at the last hydrant[1]. The colours refer to the colour coding in the figures, if not indicated differently in the figure caption.

run	dataset	XWM configuration	nominal flow	duration[1]
1	1 (red)	A	3-5 m/s	272 s
2	2 (green)	B	2 m/s	419 s
3	3 (blue)	C	1 m/s	1052 s
3	4 (cyan)	B	1 m/s	970 s
4	5 (magenta)	C	2 m/s	465 s
4	6 (black)	D	2 m/s	432 s

[1] For dataset 1, the duration indicates the time after which the flow was stopped.

pipe, the flow in the DN150 pipe was estimated to be 0.25-1.40 m/s, which yields Reynolds numbers ranging from $0.3 \cdot 10^5$ to $1.6 \cdot 10^5$, indicating turbulent flow in the DN150 pipe.

In all runs, the XWMs were successfully injected into the pipeline and transported with the flow of the water through the DN150 pipeline and extracted with the reverse flow. Minimum damage was done to the XWMs and no noticeable damage was done to the pipeline infrastructure or to the environment.

Measurements of all motes are extracted and stored in their respective datasets as given in the Table 6.3. Datasets 2-5 represent data resulting from flawless operation and retrieval of motes. We experienced irregularities in mote operation or recording of datasets 1 and 6.

Dataset 1 was recorded by the mote in configuration A with an older, less robust shell for comparison reasons. In run 1, a noticeable sound of a mote colliding with the 90 degree bend of the injection pipe was heard seconds before the mote extraction. The visual inspection of the retrieved mote revealed cracks in the shell. Although the mote experienced leakage and electric short circuit, all data until the collision were recovered.

Dataset 6 was recorded by the mote in configuration D, the heaviest one. Its center of mass was shifted significantly by the ballast piece, which was found loose after recovery, presumably, causing a short circuit seconds before extraction. All data until the short circuit were recovered.

6.5   Data analysis & results

In the data analysis, two coordinate reference systems are used:

- The IMU internal frame of reference $\{x^i, y^i, z^i\}$, with x^i and z^i in the equator plane.

- The Earth frame of reference $\{x^e, y^e, z^e\}$, with z^e aligned along the gravitational force.
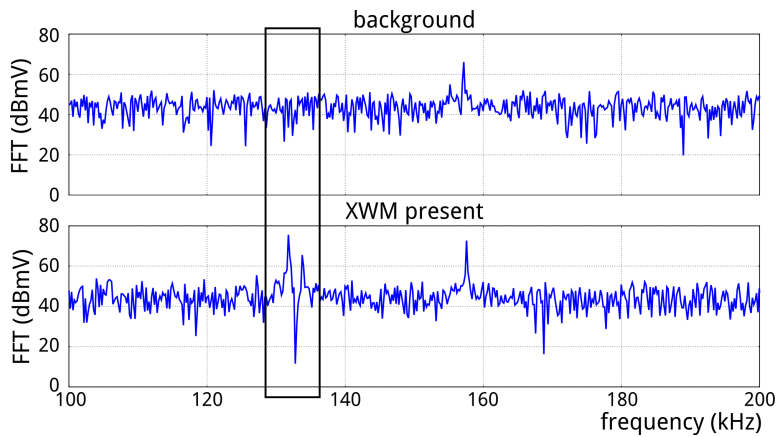
Figure 6.4: Detection of 131 kHz XWM ultrasound signal against background.

## 6.5.1 Ultrasound detection

Detection of mote arrival at the last hydrant H88 was verified for run 3, with two motes injected (see Table 6.3). The hydrophone was lowered into hydrant H88 before reversing the flow, i.e. during a phase the motes were expected to be located directly underneath or at least nearby H88. The measured ultrasound spectrum is compared to the ultrasound spectrum measured in the absence of motes, as seen in Fig. 6.4. A signal exceeding the noise level with 30 dB can be observed around 131 kHz in the frequency spectrum which clearly indicates the presence of the motes. This signal was used to verify arrival of at least one mote at hydrant H88 and served as a reference in the analysis. The detection of stuck motes was not tested because the applied retrieval procedure worked flawlessly.

6.5.2 Field-test data

Fig. 6.5 shows the IMU data recorded during the full period of run 2. Six run phases can be identified in the data based on the experimental procedure:

- XWM insertion and rest in injection pipe;
- start of forward water flow;
- arrival at hydrant H88;
- water flow stop;
- start of reverse water flow;
- extraction of XWM.

Before insertion into the injection pipe, the XWM was moving significantly while being held in hands. After insertion, the XWM hit the water and sank to the bottom of the pipe where it came to a complete standstill. The XWM started moving and rotating at the moment the flow was gradually increased. Upon arrival at hydrant H88, the XWM remained stuck, but kept rotating and moving underneath the hydrant until the mote finally sank to the bottom
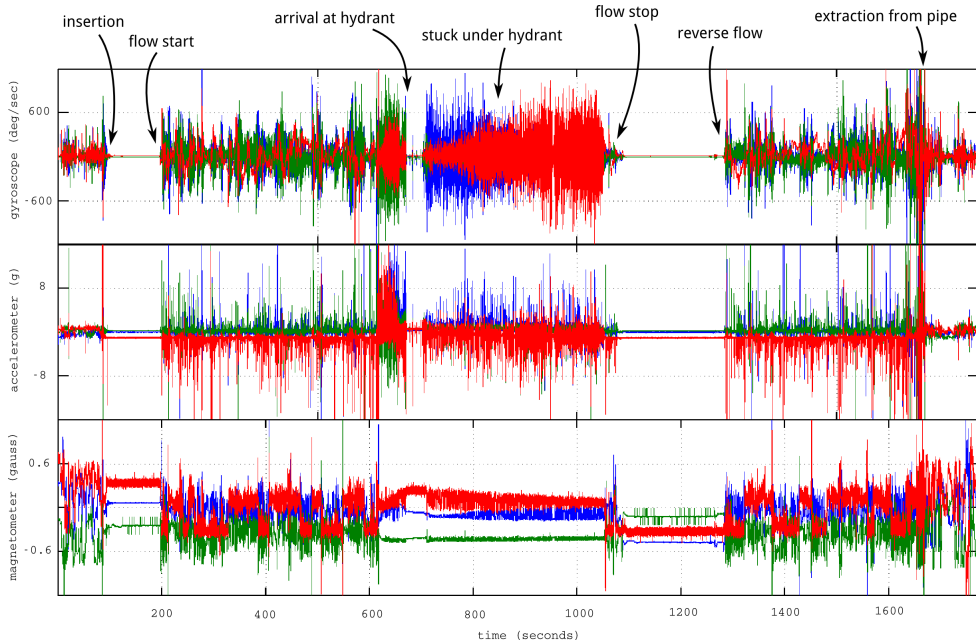
Figure 6.5: Complete dataset 2 (see Table 6.3), x^i-components in blue, y^i-components in green and z^i-components in red.

Table 6.4: Upper limits of the standard deviations σ of the IMU measurements in datasets 1-6. The deviations are calculated for a 1 minute recording interval in the period after insertion and before flow start, i.e. when the motes are at rest (see Fig. 6.5).

σ	x	y	z
gyroscope	<0.88 deg/s	<1.1 deg/s	<0.66 deg/s
accelerometer	<0.053 g	<0.027 g	<0.10 g
magnetometer	<0.0043 gauss	<0.0041 gauss	<0.027 gauss

of the pipe after the flow was stopped.

Table 6.3 shows the duration of the travel between start of the flow and arrival at hydrant H88. From run 3 and 4, with datasets $\{3, 4\}$ and datasets $\{5, 6\}$, respectively, it can be seen that the different motes traverse the same path in a different time. Depending on the weight and the mass distribution, the motes experience a different flow profile. This will be elaborated on in the discussion section.

Table 6.4 provides the upper limits of the standard deviations (σ) of the IMU measurements when the motes were at rest. The deviations are calculated for a 1 minute recording interval in the period after insertion and before the start of the flow (see Fig. 6.5), when the motes approximately were at rest and the local magnetic field was static. The σ values reflect the combined effect of noise of the IMU sensor channels and minor oscillations relative to the position of rest. The σ values indicate the sensitivity limits for detecting changes in the mote dynamics and the local magnetic field. These values provide a benchmark to compare the data recorded in periods when the flow was on, as shown for dataset 2 in Fig. 6.5, revealing the richness and complexity of the data caused by the interaction of the mote with the flow and the pipeline.

6.5.3 XWM orientation

The XWM dynamics are governed by the density and the mass distribution. In the field-test, due to the added ballast pieces, the XWM centre of mass was shifted in the direction slightly off the z^i-axes, causing, similar to a tumbler toy, an alignment of the z^i-axes along the z^e-axes due to gravity, as illustrated in Fig. 6.6.

Distortions in the flow and collisions with the pipe wall, cause a change in the mote orientation, followed by a precession rotation in respect to the z^e-axis. This can be observed in the data, e.g. in Fig. 6.7 where a small part of the data is visualized in detail. The gyroscope data show high frequency rotations around the x^i- and y^i-axis and smoothly varying rotations around the z^i-axis. This dynamics data suggests that the XWMs do not roll against the pipe wall; rather they float almost freely and align with the gravitational field as observed in Fig. 6.7 by the almost constant value of the accelerometer z^i-component of -1 g.

6.5.4 Magnetic data and pipe sections

The z^i-component of the measured magnetic field between flow start and arrival at final hydrant is plotted for all datasets in Fig. 6.8. In all datasets, a specific magnetic pattern is clearly visible throughout the pipeline. In order to perform a better comparison of datasets, a *first-order position estimate* is made, indicated using symbol $\hat{d}^*$, by multiplying the timestamps with an average mote velocity. The average mote velocity for each of the datasets is calculated as the pipeline length, 262 m, divided by the travel duration between flow start and arrival at the last hydrant as given in Table 6.3.

Even though the different runs are performed at different flow velocities and velocity gradients, this pattern is present in all six datasets. This pattern can be seen as a structural signature of the pipeline and is used in the analysis to distinguish pipe sections and to synchronize the different datasets.
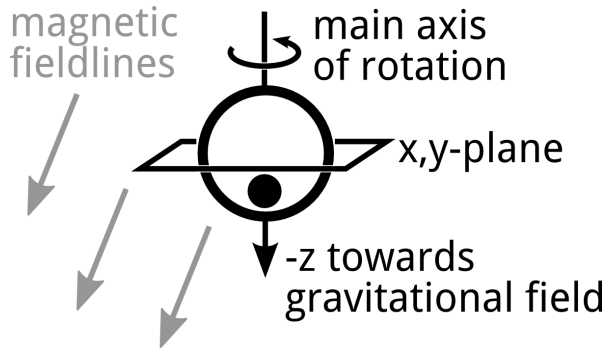
Figure 6.6: Sketch showing the orientation of the XWM in the Earth magnetic and gravitational field. The shift of the centre of mass in negative direction along the z^i-axis causes a stable orientation in the unperturbed case. Distortions will cause precession-nutations about the z^i-axis.

It should be noted that the actual mote velocity varies over time and the position inside of the pipeline. Thus, the calculated position estimate $\hat{d}^*$ is not a representation of the *actual position* d, but rather a first-order estimate.

Part of the raw data from dataset 3 is visualized in detail in Fig. 6.7. It shows a pattern that is representative for the entire dataset. For the section before $\hat{d}^* = 72.5$ m, the magnetic z^i-component exhibits a 'low' value which in a step-like transition changes into a 'high' value for positions beyond. In sections with 'low' magnetic values the gyroscope data show that the mote rotated according to precession and nutation behaviour about the z^i-axis similar to the tumbler toy principle. A clear correlation between the strong oscillations in the x^i- and y^i-component of the gyroscope and the magnetic field is visible. This suggests the presence of a fixed and static magnetic field throughout this section of the pipe, similar to e.g. the Earth magnetic field.

In areas with 'high' magnetic values the above described correlation between rotation and measured magnetic field is far less visible and the measured field is weaker. It indicates that the sensed magnetic field is governed by a weaker field that is dependent on the radial position of the XWM within the pipe. The 'high' values of the measured magnetic field are the result of strong shielding due to a non-zero magnetic susceptibility of the pipe material. The dependency of the position of the mote on the measured magnetic field suggests the pipe is made of a magnetized material, as explained in [12].

The Earth's magnetic field at the experimental site points downwards at an angle of 68 degrees from the surface[1] [13]. In this experiment an absolute comparison with the Earth's magnetic field cannot be made because of magnetic distortions introduced by steel ballast pieces and the Lithium battery inside of the XWM. However, we can estimate the step size in measured 'low' and 'high' values of the magnetic z^i-component. In datasets 1-5 the observed step size for all transitions is 0.41±0.01 gauss, which is 15 times the σ sensitivity limit. In dataset 6 the step size is 0.34 gauss and is significantly smaller due to the fact that the IMU

---

[1]M = {18654, 538.6, 45661.6} nT (x^e, y^e, z^e) in 'Earth' frame.
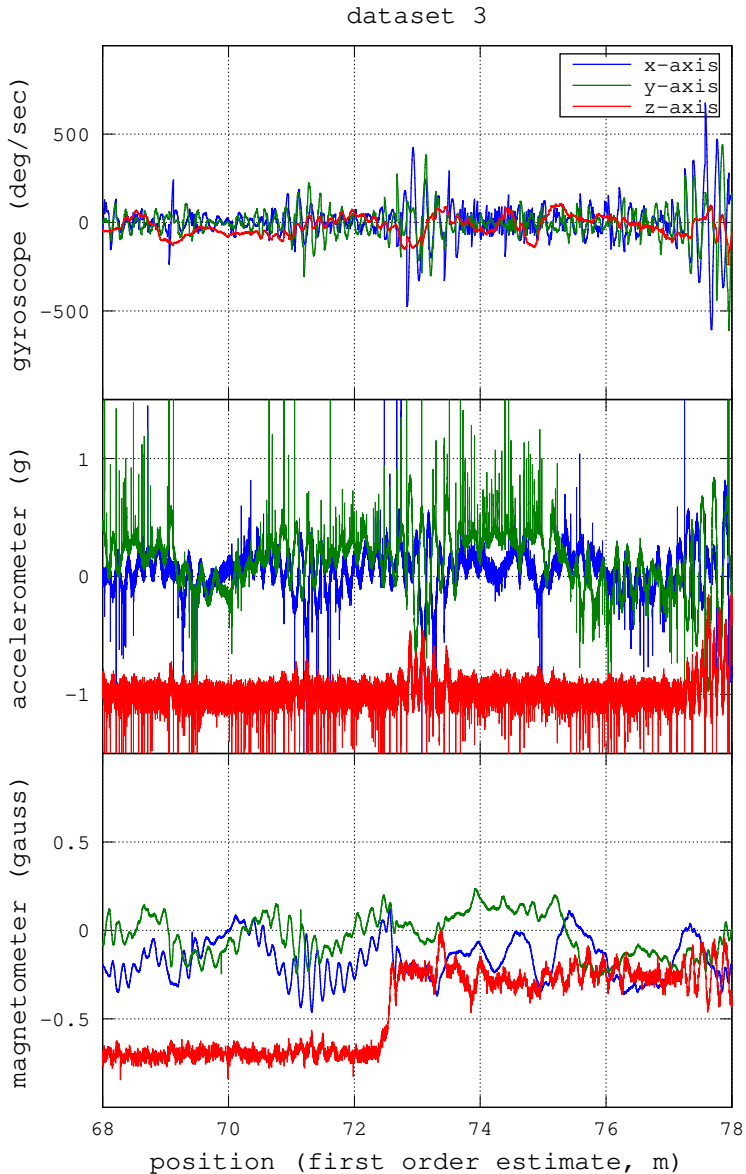
Figure 6.7: A fraction of dataset 3 (see Table 6.3) with gyroscope data (top), accelerometer data (middle), and magnetic data (bottom). The x, y, z-components refer to the IMU internal reference frame.
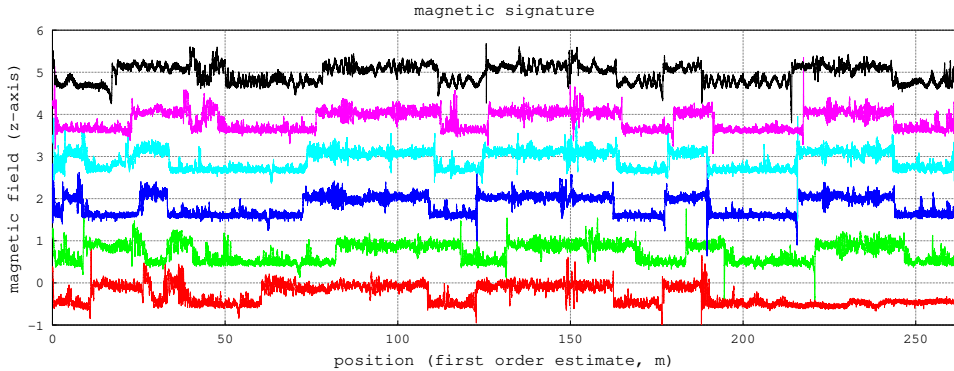
Figure 6.8: The magnetometer z^i-component of dataset 1 (red, bottom) through 6 (black, top), for run parameters see Table 6.3. The position is based on a linear conversion as described in the text. For the sake of better visualization, the amplitude values are shifted by a constant (vertical) offset value. .

z^i-axis of this particular mote was less aligned with the z^e-axis of the Earth frame.

We, therefore, conclude that two different materials are used for the pipe: a (ferro-)magnetic material and a material with a magnetic susceptibility that is close to zero. In one of the provided maps, it is suggested that the pipe is made of glass-reinforced epoxy (GRE), hereafter referred to as 'plastic pipe', and has a magnetic susceptibility close to zero. The material in the other pipe sections is likely to be cast-iron, hereafter referred to as 'metal pipe', as it is the only (ferro-)magnetic material commonly used in piping systems in the 1960s when the pipe system was constructed [1].

6.5.5   Identifying mote location and pipe 'anomalies'

Besides the observed step-like pattern in the magnetic data as discussed in the previous section, the magnetic data exhibit anomalies that are caused by e.g. pipeline components like junctions and hydrants. Identifying pipeline components from the data that have a known position will aid in improving the position estimate of the motes.

A clear anomaly is visible in the magnetic z^i-axis at $\hat{d}^* = 62.5$ m in dataset 3, as seen in Fig. 6.9. This anomaly is also observed around the same position in all other datasets as seen in Fig. 6.8 and is located inside a plastic pipe section. The magnetic signature indicates the presence of a metal object or pipe in the positive z^i-axis direction above the main pipe. The object compresses the magnetic field lines, increasing the magnetic field and resulting in a stronger (negative) measured magnetic field. The first-order position estimate of this anomaly in datasets 2-6 is in the range 60 m $\leq \hat{d}^* \leq$ 69 m as can be seen in Fig. 6.8. Actual positions of pipeline components can be estimated on the map in Fig. 6.3 and are indicated by the symbol d. In particular, the hydrant H100 is expected around $d = 62$ m. Therefore, we associate the anomaly with the hydrant H100.

A similar magnetic signature is measured at $\hat{d}^* = 114$ m but now in the $x^i y^i$-plane, as
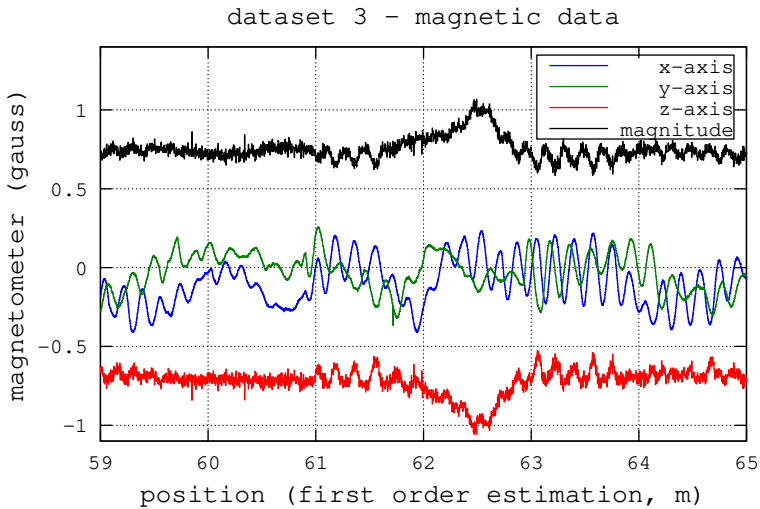
Figure 6.9: Magnetic data of dataset 3 (see Table 6.3), the x, y, z-components refer to the IMU internal reference frame, the position is analogue to that in Fig. 6.8. The magnetic signature fits with hydrant H100 expected at this position.
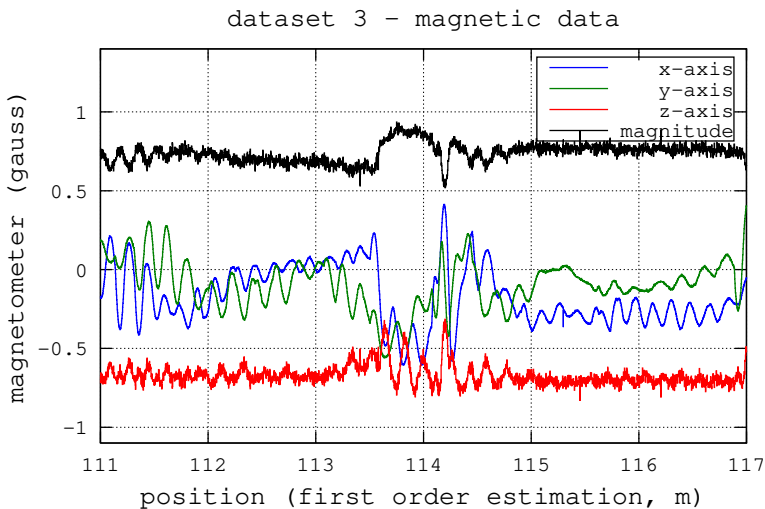
Figure 6.10: Magnetic data of dataset 3 (see Table 6.3), the x, y, z-components refer to the IMU internal reference frame, the position is analogue to that in Fig. 6.8. The magnetic signature fits with the drencher expected at this location.
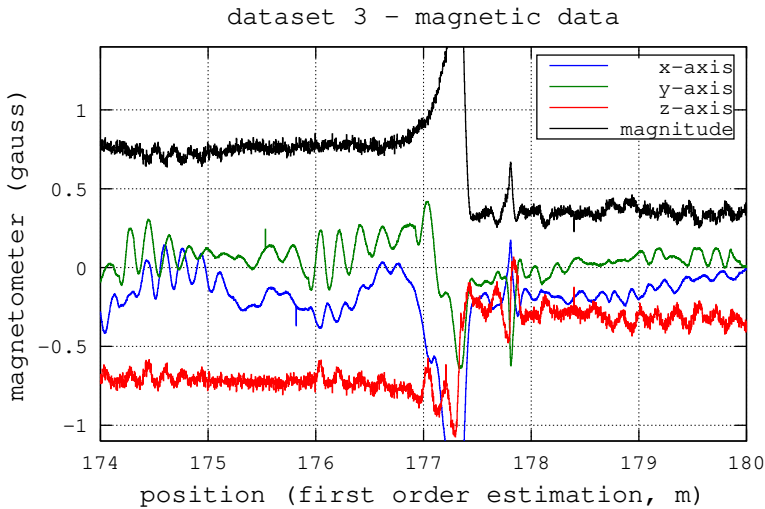
Figure 6.11: Magnetic data of dataset 3 (see Table 6.3), the x, y, z-components refer to the IMU internal reference frame, the position is analogue to that in Fig. 6.8. The position is approximately matching the expected position of hydrant H89.

shown in Fig. 6.10. This indicates a metallic object parallel to the (Earth) surface. This anomaly is seen in the range 109 m $\leq \hat{d}^* \leq$ 118 m in the different datasets as can be seen in Fig. 6.8. On the map, the drencher is expected at $d = 120$ m and is drawn as pipe perpendicular to the main pipe and parallel to the surface. It is therefore hypothesized that this anomaly is due to the metal of the drencher pipe.

A third magnetic anomaly can be identified as hydrant H89 and can be seen in Fig. 6.11. A sharp peak in the magnetic field measurements can be detected just after a transition from plastic pipe section to metallic pipe section just before $\hat{d}^* = 178$ m. In all other datasets this peak appeared between 171 m $\leq \hat{d}^* \leq$ 178 m, as seen in Fig. 6.8. As this happens inside a magnetized pipe, the peak in magnetic field is less obvious in indicating the direction of a possible pipe. The estimated position of the observed anomaly is close to the position where hydrant H89 is expected (on the map at $d = 180$ m). Therefore, this third point is identified as the position of hydrant H89.

In order to obtain a better position estimate, we can now use the pipe transitions observed in Sec. 6.5.4 as synchronization points between the datasets, and the three identified anomalies in this section as landmarks with known actual positions. These points, together with the known location of injection pipe and the hydrant H88 are used in a spline interpolation of the estimated position. This *adjusted position estimate* is indicated with the symbol $\hat{d}$. In Fig. 6.12 the magnetic signature is then plotted against $\hat{d}$. The pipe sections and other annotated anomalies are visualized on the embedded schematic of the pipe system. In this figure, the pipe transitions and the landmarks as observed in the data are synchronized between the different datasets, and aligned with the three identified pipe-components on the schematic of the pipe system.
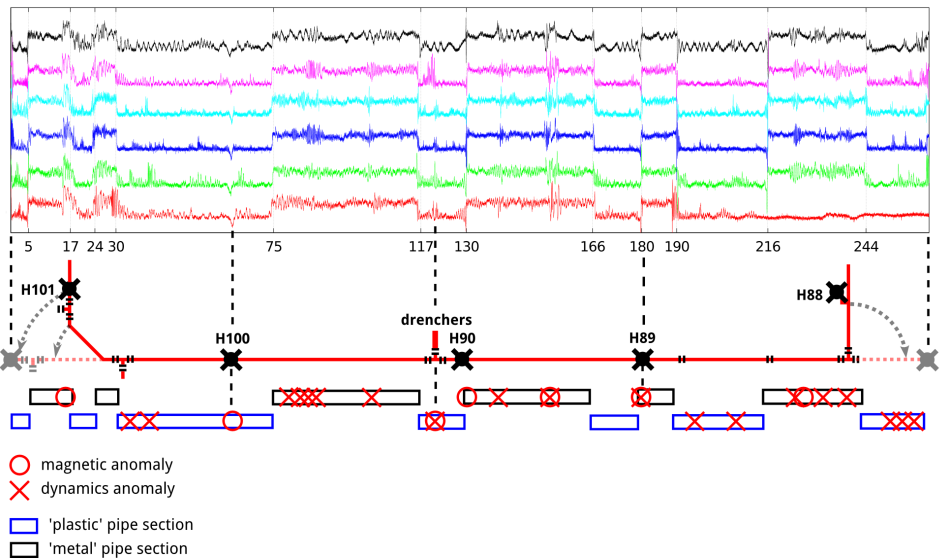
Figure 6.12: Dataset identical to the one shown in Fig. 6.8. The position has been adjusted based on the synchronization points between the datasets and the identified landmarks. For details we refer to section 6.5.5. For comparison, a schematic map of the Emmtec location with the identified landmarks indicated and a scheme of the pipeline sections has been included. The position (in unit m) of pipe section transitions are indicated as well.
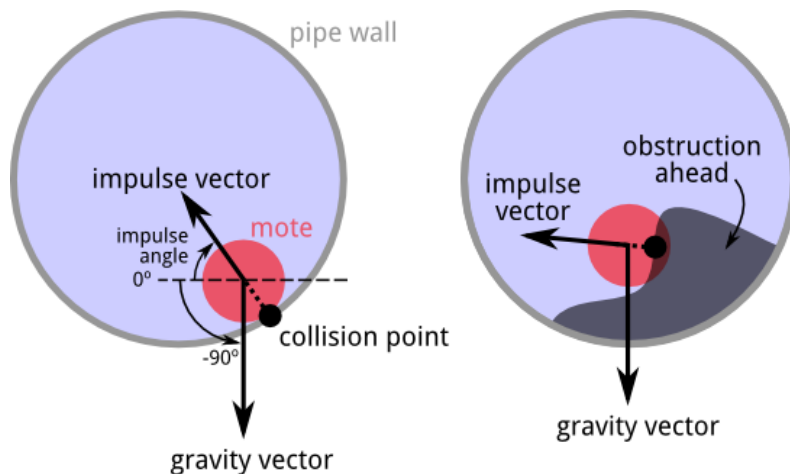
Figure 6.13: During a collision a mote receives an impulse that is estimated by the acceleration it experiences. The impulse direction is estimated relative to the gravitational force. On the left a collision with the pipe wall, on the right a collision with an obstruction that is ahead of the mote.
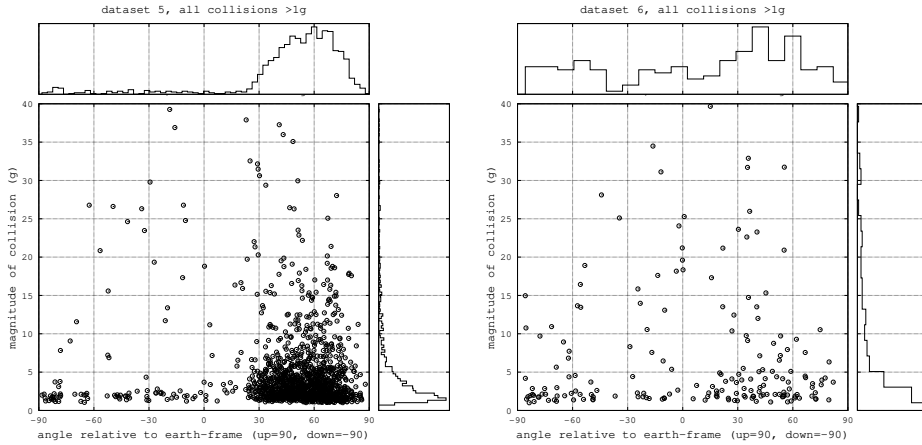
Figure 6.14: Magnitude of the collision versus impulse angle for collisions exceeding a magnitude of 1g detected in dataset 5 (left) and dataset 6 (right). The distribution of dataset 5 is representative also for datasets 1-4.
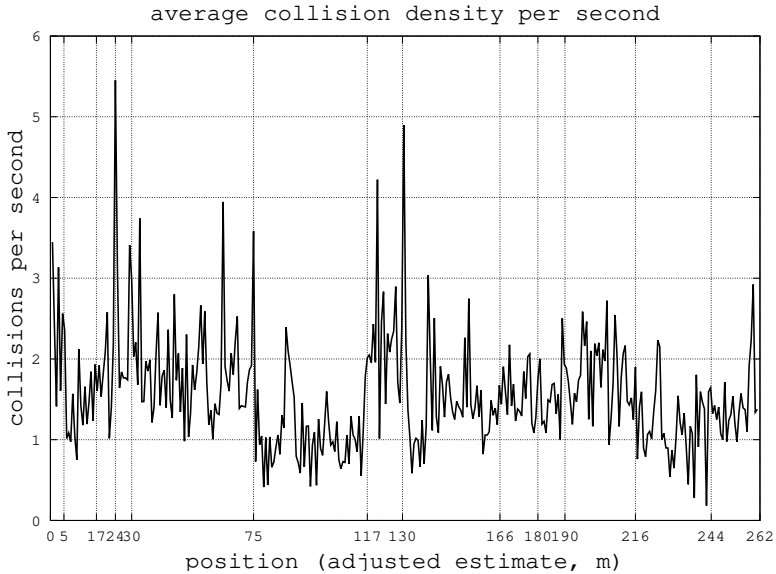
6.5.6   Collisions

The measured data show that the motes experienced a large number of collisions (see Fig. 6.5 and Fig. 6.7). Collisions are marked as sharp peaks in the accelerometer data with a duration ranging from 0.5-4.0 ms. The majority of these collisions are accompanied by a sudden change in mote rotation as can be seen in Fig. 6.7.

During a collision, for a short period of time, a mote experiences a strong acceleration due to the sudden change in direction. The total impulse given to the mote is proportional to the integral of the acceleration during this collision. In this work, we estimate the impulse magnitude and direction based on the acceleration measured in the collision peak.

The measured acceleration has two contributions: the gravitational acceleration and the acceleration due to the collision, as depicted in Fig. 6.13. During the short period of the collision, the acceleration due to the dynamics of the flow and also the rotation of the mote can be neglected.

Just before and right after a collision the only measured acceleration of a free floating mote is due to gravity represented in the IMU system by a vector $\{x_g^i, y_g^i, z_g^i\}$ which is aligned with the gravitational force and equal to 1 g. The measured impulse vector in the IMU-frame $\{x_i^i, y_i^i, z_i^i\}$, can be transformed into the Earth-frame $\{x_i^e, y_i^e, z_i^e\}$ by applying a rotational-operation. This rotational operation is equivalent to that of rotating the gravity vector as measured in the IMU frame $\{x_g^i, y_g^i, z_g^i\}$ towards the gravity vector defined in the Earth-frame $\{x_g^e, y_g^e, z_g^e\} = \{0, 0, \text{-1}\}$. We can then calculate the direction of the impulse vector in the Earth-frame as an angle relative to the gravitational field, as shown in Fig 6.13. We take an angle of 0 degrees to be parallel to the surface of the Earth, -90 degrees points downwards, and +90 degrees points upwards.
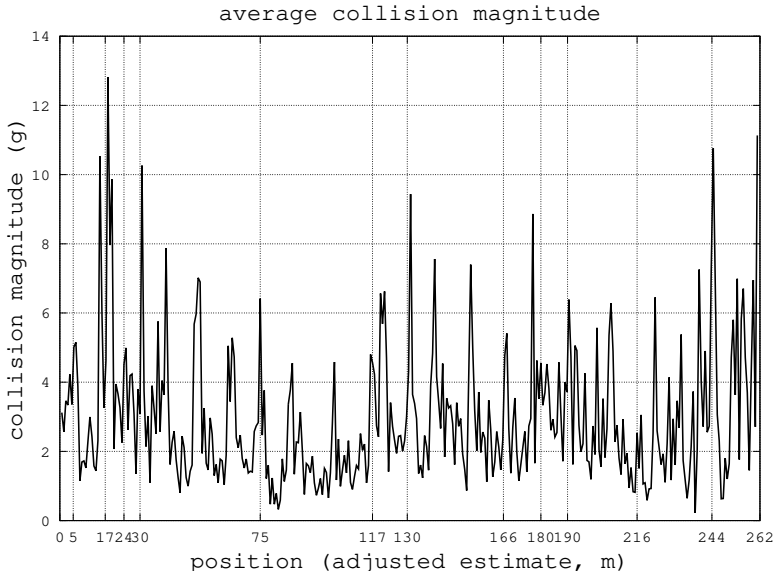
Figure 6.15: Collision count per second of greater than one g collisions plotted as a function of the adjusted position similar to Fig. 6.12. Average taken from all datasets.

Fig 6.14 shows the detected collisions from datasets 5 and 6. The collision magnitude is indicated on the y-axis of the graph and the impulse angle in the Earth-frame on the x-axis of the graph. It is observed that in dataset 5, the vast majority of the collisions are pointed upwards, similarly to datasets 1-4. This indicates that the collisions mainly happen on the bottom of the pipe. In dataset 6, however, the amount of collisions is significantly less and the impulse directions are almost uniformly distributed.

Our analysis shows that the motes from datasets 1-5 collide on average between 1-10 times per metre while the heaviest mote in dataset 6 collides only 0.1-1 times per metre. This difference can be explained by an offset in their center of mass. The motes used for datasets 1-5 have their center of mass offset in the z^i direction in the IMU-frame. It is located in the same plane in which the rim of the shell is oriented: the $x^i z^i$-plane. The mote used for dataset 6 is the same as the one from dataset 2 and 4, but with an extra added ballast piece in the y^i direction. As a consequence, the center of mass shifted to outside the $x^i z^i$-plane, breaking the symmetry of the mote between the gravitational alignment and the rim. This seems to have a stabilizing effect on the motes' rotation and its trajectory in the pipe, as the mote is floating more freely through the pipe, colliding on all sides almost evenly.

In Fig. 6.15, the collision densities (in collisions per second) are plotted as a function of the adjusted position estimate similar to Fig. 6.12. The density is taken as average over all datasets. The density significantly varies as a function of position and clusters around specific positions. For example, at $\hat{d} = 75$ m and $\hat{d} = 130$ m, clusters of sharp density peaks are observed that coincide with positions of transitions from a plastic pipe to a metal pipe. It is also observed that the collision density of entire sections differ from each other, e.g. the

Figure 6.16: Average collision magnitude of greater than one g collisions plotted as a function of the adjusted position similar to Fig. 6.12. Average taken from all datasets.
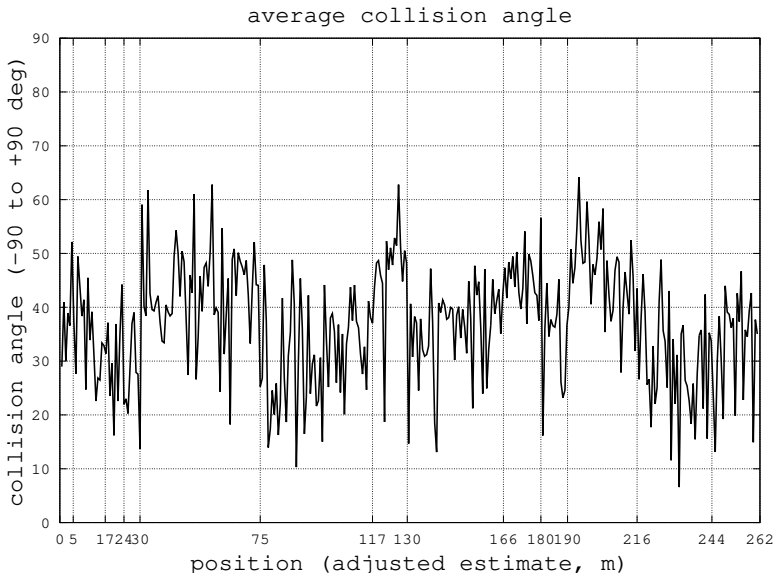
Figure 6.17: Average impulse angle of greater than one g collisions plotted as a function of the adjusted position similar to Fig. 6.12. Average taken from all datasets.

collision density in the metal pipe between 75 m$\leq \hat{d} \leq$ 117 m is significantly lower than that of the plastic pipes before and after.

In a similar fashion, the average collision magnitude and impulse angle are plotted in Fig. 6.16 and Fig. 6.17, respectively. The impulse angle plot shows a clear differentiation between pipe sections. A tendency is visible that high magnitude collisions correspond to lower absolute impulse angles, i.e. impulse vectors almost parallel to the surface of the Earth. This suggests that these collisions happen on larger obstacles, causing larger losses in forward velocity, e.g. head-on collisions on the drencher pipe at around $\hat{d} = 120$ m.

6.6 Discussion

6.6.1 Analysis approach

It should be noted that the data-driven analysis presented in this paper is just a first pragmatic approach for analysing the data. Our objective was to investigate and eventually establish the feasibility to extract information on the structure and the condition of the pipeline during operation. If in forthcoming laboratory- and field-tests the required database on dynamic and magnetic patterns can be accumulated, a much more profound data analysis will become available.

The authors also foresee a model-driven approach based on modelling the flow and the mote dynamics depending on a specific pipeline structure. Mote measurements can be simulated using this model and compared with the actual measured data to identify what type of structure the motes have encountered. However, the required simulations with high granularity including e.g. mote-obstacle collisions inside fluids are challenging.

6.6.2 Pipe sections and transitions

Pipe sections have been identified and localized. The findings are based on a magnetic step-like pattern which can persistently be detected in all datasets. The 'high' to 'low' step size for datasets 1-5 is 0.41±0.01 gauss, demonstrating that this structural feature is measured by all motes in different runs with high consistency and repeatability. The difference in correlation between the motes' rotations and the measured magnetic field suggests that the pipe sections are made of two different materials, one with high magnetic susceptibility and one with low magnetic susceptibility. These observations are also partially confirmed by the statistics of the mote collisions with the pipe wall (Fig. 6.15 through Fig. 6.17). The latter findings suggest a correlation between motes dynamics and magnetic measurements in different pipe sections, and which is consistent throughout the various datasets.

6.6.3 Collision dynamics

The statistical analysis of the collision dynamics shows that, except for dataset 6, the motes are mainly hovering slightly above the bottom of the pipeline and are colliding with it at a rate between 1 to 10 times per metre.

Localized peaks in the collision statistics (higher collision rate, larger magnitude, lower impulse angle) suggest the presence of 'obstacles', pipe-components or other anomalies that

are different from unobstructed straight pipe sections. The locations of some anomalies in the mote dynamics coincide with identified pipe-transitions, hydrants and the drencher. Other peaks have not been identified yet and need a better look into the dynamics to relate to e.g. junctions or obstructions.

6.6.4   Pipe conditions and integrity

We did not yet achieve a quantitative assessment of the integrity of the pipe and the junctions. Aside from the dynamics anomalies, the motes interact differently with pipe walls in different sections. The distinction between 'metal' and 'plastic' sections alone is not sufficient to uniquely categorize the observed collision dynamics. This might suggest a difference in wall structure, reflecting the ageing, wear or roughness of the pipe sections.

Since no additional information on the used pipe-components was available, a more thorough study into the mote dynamics for known pipe structures and conditions under laboratory conditions needs to be performed to provide the required information for an advanced data analysis.

6.6.5   Position estimate

The actual positions of the motes are not known except for the start, the end and arguably the three identified landmarks. The actual motes' velocities are not only dependent on the overall average flow velocity, but also on the local flow characteristics and the interaction with the pipe and obstacles in it. Head-on collisions, for example, are accompanied by a loss in forward velocity.

The deviations between the observed range $\hat{d}^*$ in the first order estimated positions of anomalies and the actual positions of the landmarks vary significantly as shown in Sec. 6.5.5. While H100 (d=62 m) is still inside the corresponding $\hat{d}^*$ range, the positions of the drencher (d=120 m) and H89 (d=180 m) are clearly outside the estimated range. An explanation for these deviations might be found in collision statistics. In Figs. 6.15 and 6.16 the frequency and the magnitude of mote collisions are shown as a function of the adjusted positions $\hat{d}$. Upstream the drencher, motes were passing the pipe section 75 m $\leq \hat{d} \leq$ 117 m with reduced collision frequency and intensity. In consequence, the actual mote velocity had the tendency to exceed the average velocity and the motes passed the drencher early, which translates into smaller position values compared to positions calculated using the average velocity. Between the drencher and H89, the motes passed section 117 m $\leq \hat{d} \leq$ 130 m with increased collision frequency and sections 130 m $\leq \hat{d} \leq$ 166 m and 166 m $\leq \hat{d} \leq$ 180, respectively, with reduced collision frequency, where the first sections partly reduced the velocity excess, which fits the observed deviation pattern.

Overall, the estimation of the position seems to be reasonable. Besides the identified landmarks, the schematic map contains more details that can be used in further verification of positions. For example, two flanges are drawn between hydrant H89 and the 90 degree corner in the schematic. This could coincide with the pipe transitions that have been measured. These might be repairs that are documented in this map. Also, just before arriving at the final hydrant H88, three 90 degree corners are drawn in the schematic. These could coincide with the dynamic anomalies found in the last pipe section.

Spline fitting was used for an improved conversion between travel time and mote position, synchronizing motes positions on a variety of points and adjusting the positions to the known locations of five landmarks.

6.6.6 Mote mass distribution

The field-test demonstrated the impact and the importance of the motes' mass distribution. The center of mass offset caused the motes to align with the gravitational field, which provided the base for transforming data measured in the IMU frame of reference to the Earth frame of reference. Additionally, we suppose that the offset in the centre of mass relative to the plane of the rim causes motes to follow a more stable trajectory and reduces the number of collisions with the pipe wall as seen in dataset 6. We can use this fact to probe different aspects of the pipeline by using motes with different mass distributions.

6.7 Conclusion

We demonstrated for the first time that a water pipeline can be explored using small centimetre sized sensor motes with an inertial measurement unit and ultrasound emitter.

The measured data are extremely rich and complex, exhibiting different and distinct patterns. Additional information about the pipeline, i.e. the presence and position of different pipe sections and pipe components and the used materials, was deduced using the data-driven approach presented in this paper. Furthermore, an initial statistical analysis of the interactions of the motes with the pipe wall revealed structural differences between the pipe sections. These findings show that XWMs can be used to monitor the condition of pipelines as such statistical information can be an indicator for wear or ageing of the pipeline at different positions and can be measured at different times.

In this paper we showed the importance of understanding the influence of different mote weights and mass distributions on the mote interaction with the pipeline. It can be used to probe different aspects of the pipeline. Future works includes performing measurements with a larger number of motes to increase the statistics and to reveal structural information on a smaller scale (i.e. larger sensitivity). Furthermore, data-driven analysis can be complemented with a model-driven approach to combine results and reveal more structural information of the pipeline from the data.

6.8 Acknowledgement

References

[1] H. Kater, de, *et al.*, "Inspectietechnieken voor rationeel saneringsbeleid van leidingnetten," KWR Nieuwegein, BTO 2010.013, 2010.

[2] Z. Liu and Y. Kleiner, "State of the art review of inspection technologies for condition assessment of water pipes," *Measurements*, vol. 46, no. 1, pp. 1-15, Jan. 2013.

[3] J.M. Mirats Tur and W. Garthwaite, "Robotic Devices for Water Main In-Pipe Inspection: A Survey," J. Field Robotics, vol. 27, no. 4, pp. 491-508, 2010.

[4] T. Hao, *et al.*, "Condition assessment of the buried utility service infrastructure," *Tunneling and Underground Space Technology*, vol. 28, pp. 331-344, 2012.

[5] R. Harle, "A Survey of Indoor Inertial Positioning Systems for Pedestrians," *IEEE Commun. Surveys & Tutorials*, vol. 15, no. 3, pp. 1281-1293, 2013.

[6] E. Talnishnikh, *et al.*, "Motes for environment mapping," U.S. patent 9397732, July, 19, 2016.

[7] E. Talnishnikh, *et al.*, "Micro Motes: A Highly Penetrating Probe for Inaccessible Environments," in *Intelligent Environmental Sensing*, edt. H. Leung, Springer Int. Publishing, 2015.

[8] H.J. Wörtche, "Advanced Analysis: Motes Dynamics in a Mixing Tank," INGU Solutions, 2014.

[9] A. Bogerman and E. Talnishnikh, "Test: Pressure test with Xploring WiseMotes (XWMs) at Wavin", INCAS[3], Dedemsvaart 2016.

[10] "LSM9DS0 - iNEMO inertial module: 3D accelerometer, 3D gyroscope, 3D magnetometer," STMicroelectronics, 2017.

[11] V.N. Sekhar, "Ultrasound Implementation in XWM," Eindhoven University of Technology, Department of Electrical Engineering, The Netherlands, 2016

[12] Huang Xinjing, *et al.*, "Magnetic Charge and Magnetic Field Distributions in Ferromagnetic Pipe," *Appl. Computational Electromagnetics Soc. J.*, vol. 28, no.8, pp. 737-746, August 2013.

[13] NOAA: National Centers for Environmental Information, geomagnetic database [Online]. Available: https://ngdc.noaa.gov/geomag-web/#igrfwmm

7
Conclusion

This thesis focused on a new approach for exploration of enclosed environments that contain a (semi-)fluid medium. Examples of these are man-made environments such as pipelines, industrial mixing tanks and reactors, and natural environments like underground channels, reservoirs and geothermal sources. Structural and in-situ information of these environments is requested by industry for e.g. guaranteeing operations, for safety purposes, or for optimizing processes. Current state-of-the-art technology does not provide this information at all, or not in a cost-effective manner. We proposed to explore these environments by means of a "go-with-the-flow" approach that uses small buoyant autonomous sensor nodes in large quantities. While traversing the enclosed environment, the sensor nodes perform measurements for parameters of interest, and perform ranging, i.e. they measure their positions relative to neighboring nodes. Measurements are stored in memory and read out when the nodes are extracted from the environment. After extraction, the stored data is analyzed in order to gain the desired knowledge.

The overarching research question addressed in this thesis was: "Can we design an efficient and effective 'go-with-the-flow' system for characterization of enclosed fluidic environments?"

Position determination of the nodes is of pivotal importance in order to design an effective go-with-the-flow approach. To minimize power consumption and size of the nodes, in this thesis we focused on the approach to obtain positions in a cooperative manner. Nodes perform ranging measurements using ultrasound and robust reconstruction of the positions of all nodes happens offline. We also explored alternative localization approaches. Specifically, in Chapter 6 chapter we demonstrated the possibilities of simultaneously localizing node positions and mapping a pipe surface and its features by studying nodes' movement and the local magnetic field. Accordingly we narrowed our overarching research question down to the following three specific research questions:

1. Can we develop a robust algorithm for the reconstruction of node positions using rang-

ing measurements that are obtained in the go-with-the-flow approach?

2. Can we develop a ranging protocol for nodes to determine distances while traversing the environment, such that nodes need to perform as little effort as possible during the go-with-the-flow phase?

3. Can we apply other methods to perform localization?

7.1   Robust reconstruction

Our first research focus was on developing a robust reconstruction algorithm to explore the maximum abilities of the offline analysis. With robustness we mean the ability to deal with severely corrupted distance measurements. For the reconstruction we used a RANSAC-based approach as this is designed to be robust for both inlier and outlier measurements. Our approach resolves ambiguities due to errors in the distance measurement by enforcing geometric consistency of the node positions. We stress-tested the algorithm with increasingly corrupted distance measurements in order to find the minimum required specifications of the nodes' distance measurements.

The quality of the reconstruction algorithm was assessed based on specific performance metrics. Common performance metrics, like recall (the number of nodes reconstructed) and absolute position error alone were not sufficient to assess the performance in our applications. We therefore introduced performance metrics to more adequately characterize the performance that is required in go-with-the-flow applications.

For example, due to the incremental nature of reconstructing an initial estimate, error build-up occurs. Especially in elongated environments this results in larger absolute errors while the local shape can still be properly reconstructed. In go-with-the-flow applications it is expected that nodes will disperse throughout the environment and that not all nodes will stay within the same connected swarm; different partial swarms can exist. Similarly, in the reconstruction phase, reconstruction of the nodes may halt when the algorithm cannot deal well enough with the weakest link in the swarm in terms of measurement quality, lowering the recall metric. Up to the moment the reconstruction would halt because it encountered a weak link in the swarm, successful reconstruction for a smaller subset of nodes in the swarm was achieved. In developing the reconstruction algorithm we kept our focus on attempting to reconstruct all nodes in the swarm as this put a further demand on the robustness of the algorithm to deal with measurements that are significantly more corrupted than assumed in idealized cases.

We further challenged the offline reconstruction by treating each measurement cycle as a one-shot measurement, meaning that the measurements of only one measurement cycle were available for reconstruction. This fits in the design philosophy that minimum effort should be performed online.

We accounted for the difficulties when actually developing and deploying actual sensor nodes in experiments by introducing large amounts of different types of inlier and outlier measurement noise. This is to account for different imperfections in the system and develop robustness against them.

Besides revealing how robust the algorithm is, the simulations showed interesting node design trade-offs. For example, success of the reconstruction is mainly driven by the number of inlier distance measurements available for reconstruction, i.e. distance measurements that are close enough to the actual distance. Connectivity is therefore favoured over identifiability of the nodes. Furthermore, there seems to be a minimum number of unique identifiers needed, depending on the amount of measurement noise. Increasing this number (making the identification more unique), does not improve the reconstruction result, it only reduces the required time to complete the computations. This shows again that more effort offline can reduce effort online.

Many related research papers focus on optimizing the reconstruction algorithm such that positions are estimated more accurately and with less computational effort. However, often these highly optimized algorithms only function as intended when an initial estimate is given that is already close to the ground truth. In go-with-the-flow applications, and especially the ones without a priori information on the environment structure, there is no knowledge that can yield this initial estimate without building it from the ground up. Our focus was on the reconstruction of this initial estimate, yielding a robust basis for future go-with-the-flow experiments.

7.2   Ranging protocol

In order to address the challenges related to size and communication constraints, in Chapters 4 and 5 we successfully developed a novel ranging protocol with the philosophy that nodes should do as little as possible. The multi-way ranging approach is chosen as a basis because the total number of messages needed for sending scales linearly with the number of nodes, whereas in more traditional methods like two-way ranging it scales quadratically. In order to estimate the distances between all node pairs every node stores all received signals, even if the signal is not addressed to this node. In offline processing, this information is used to estimate distances between nodes indirectly. This reduces the required communication, and hence energy, to estimate distances. However, no existing multi-way ranging protocol was directly suitable for the go-with-the-flow approach as existing protocols are all meant for a situation where all nodes are connected to each other rather than for a highly sparse swarm that can emerge in go-with-the-flow applications in elongated environments, like pipelines.

In Chapter 4 we designed a simulation chain in which the ranging protocol was coupled with the reconstruction algorithm. A variety of node trajectories through different environments was simulated. The ranging protocol was simulated for the nodes in these trajectories and the resulting range measurements served as input for the reconstruction algorithm. This integrated simulation chain showed that movement of nodes can be seen as another type of measurement noise when using a reconstruction algorithm that assumes a quasi-static scenario, like ours.

7.3   Go-with-the-flow pipe characterization using other localization techniques

We also studied localization using different methods than the cooperative localization method. In Chapter 6 we performed a go-with-the-flow experiment in underground pipelines using newly developed sensor nodes with inertial and magnetic measurement sensors and ultrasound emitter. Nodes were inserted into the pipeline, and with the flow of the medium they traversed the 300 meter pipeline. Upon extraction, the sensor data was retrieved and used to simultaneously estimate the position of the nodes as well as to reveal structural information of the pipeline. The method validated the time-based approach for position determination and to simultaneously localize the nodes and map the environment using the environment signature. The experiment also validated the go-with-the-flow approach in general as the centimeter-sized sensor nodes successfully managed to traverse the full pipeline length and useful new insights where obtained for the exploration goal.

The inertial and magnetic measurements from the nodes were analyzed in a data-driven manner. Knowledge on the node positions and pipeline characteristics was generated in incremental steps. The initial position estimate was made using time-based position estimation. Specific characteristics of the pipeline were identified in the measured magnetic signature of the environment and used as landmarks to align the position estimates of different nodes. This allowed for further analysis of other pipeline characteristics like obstacles and pipe roughness. Besides these new findings, this also suggests that the go-with-the-flow approach is a promising technique to identify changes over time, like aging, wear, anomalies, when used in a regular monitoring program.

7.4   Overall conclusion

Overall, these findings suggest that the go-with-the-flow approach seems promising and potentially useful for a wide variety of applications.

7.5   Recommendations for future research

7.5.1   Reconstruction algorithm

We have shown that in the offline reconstruction algorithm, one can compensate significantly for highly corrupted range measurements. However, this is only the case when sufficient distance measurements can be considered inlier measurements. And as with many wireless networks, often the network either functions properly and most messages are transacted correctly, or it does not function well and most messages are not transacted correctly. It is therefore recommended to focus most of the future research on making sure that enough inlier measurements are performed, instead of working on making the reconstruction algorithm more robust against different kinds of noise.

Exploitation of the temporal dependency of consecutive measurements could alleviate some of the required robustness by benefiting from the temporal geometric consistency of the

environment and the node locations. By exploiting this consistency, measurement ambiguities can be filtered out.

7.5.2   Ranging protocol

Key factors remaining for the go-with-the-flow approach to work with the proposed ultrasound distance measurements and the reconstruction algorithm are how the signal encoding, signal propagation, signal reflections and reverberation, and signal overlap influence the ability to perform proper distance measurements. These factors have the potential to significantly influence the feasibility of performing distance estimation in enclosed environments under the strict hardware constraints. Addressing them should be a primary step in further development of this method of localization. This step can involve simulations, but more importantly also actual experiments.

7.5.3   Using other localization techniques

The work in this thesis shows feasibility of localizing and simultaneously mapping the environment using methods other than the established localization methods.

In our work we focused on the techniques for which we recently developed the hardware (inertial measurements and magnetic field). Other methods that are not directly related to 'localization' in existing literature might also yield effective localization methods for the difficult-to-access environments considered in this thesis. Examples of these are temperature or pressure gradients and for example proximity sensing as described in Chapter 1. The author challenges researchers to also get creative with other methods to estimate location rather than directly choosing for established methods.

7.5.4   Overall recommendations

Much of the work in this thesis (except for Chapter 6) has been based on theory and simulations. We started from the assumption that performing ranging measurements in these applications is possible. In this thesis we pushed the boundaries of this ideal assumption by introducing many types of noise and e.g. stress-testing the reconstruction algorithm until it would fail. However, this research did not include experimental validation of the go-with-the-flow localization method using ultrasound distance measurements. It is therefore recommended to validate and refine the approaches in real-world experiments.

In Chapter 6 we studied possibilities to perform measurements in a go-with-the-flow approach that are already feasible in practice and by doing significant offline effort to localize nodes and explore the environment. Simultaneously localizing the nodes and exploring the environment proved to be possible in real experiments using basic principles of time-based localization and by measuring physical parameters related to the environment signature. The combined approach in this thesis laid the foundations for interesting future research that can include both the more complex cooperative localization using distance measurements between nodes, as well as focusing on more basic measurement principles for detecting e.g. environment signatures and proximity of nodes.

Acknowledgement

First and foremost I would like to thank my promotors Jan Bergmans and Heinrich Wörtche. Jan offered me a tremendous amount of structure in my PhD research to keep track of the bigger picture and to not get lost in wherever my curiosity would take me. Thank you for guiding me, not only in my academic progress, but also in my personal development and future career. Heinrich offered me the opportunity to propose my own PhD research and arranged the funding for it. Thank you Heinrich for believing in me, helping me shape my thoughts and ideas into this project, and giving me a helping hand when I needed it the most.

Of course this PhD project would have not been possible without the help of my daily supervisors Gijs Dubbelman, Libertario Demi and Elena Talnishnikh; and also Jean-Paul Linnartz. Thank you for being great sparring partners to shape my ideas, to teach me new things, and help me finishing papers and this thesis.

Furthermore there are many people that I would like to thank for making my PhD period more enjoyable, bearable and interesting. From INCAS[3] there are Corina Vogt and Dirkjan Krijnders for supervision and senior advisory things; the engineers, especially Arjen Bogerman, Erik Kallen, Jan Stegenga, and Rajender Baddam for helping me with the engineering part of my experiments, most of which unfortunately did not make it into this thesis, I had a lot of fun and a couple of 400 Volt shocks more than I wished for while working on the setups; and about a dozen fellow pre-doc's, especially Mike, Zee (Shaojie), Froukje, Hedde and Edda, all of which were in a similar position as I was, sharing most of the time together. From TU/e there is my room-mate Mojtaba Farmani; and the students which I supervised, Niels Puts and Vivek Narayanan Sekhar, that helped me with my research; and all the people I did not mention specifically by name, thank you for everything.

From within the Phoenix project I would especially like to thank Peter Baltus for exchanging ideas and always being critical on research issues and applicability of technological solutions.

Before finishing my thesis, I started working at Antea Group; I would like to thank Léon Verhoeven and Chris van der Veen for taking a leap of faith and giving me an amazing position within the company. Furthermore, thank you Léon Verhoeven and César Blaauwgeers for your almost personal crusade to make me finish my PhD thesis amidst all the nice and new projects within the company.

Dankjewel pap en mam voor alles wat jullie hebben gedaan voor mij, dankzij jullie geweldige opvoeding heb ik mij kunnen ontwikkelen tot de persoon die ik nu ben. Dankjewel opa en oma voor het bijwonen van mijn master-diploma uitreiking en de trots die jullie uitstraalden op de treden van het Academie gebouw in Groningen, sorry dat de komende ceremonie helemaal in Eindhoven is. Dankjewel vrienden, broer en zussen voor het er zijn.

And last, but definitely not least, I would like to thank my wife Candela. You make me see the bigger things in life in a different manner, you pick me up when I'm feeling down, you make me push through when I don't see things happening anymore. Thanks for your support and infinite love.

Curriculum Vitae

Hendrik Albert (Erik) Duisterwinkel was born on 31-08-1986 in Groningen, The Netherlands. After finishing VWO degree in 2004 at Wessel Gansfort College in Groningen, The Netherlands, he studied Applied Physics at University of Groningen in The Netherlands, both bachelor and masters degree. For his masters degree he did a research internship at the University of Buenos Aires in Argentina at the Image Processing Laboratory on ghost imaging. In 2011 he graduated cum laude within the Medical imaging and Proton therapy group at the Nuclear Acceleration Institute (KVI) in Groningen on developing a Compton spectrometer for high-intensity medical X-ray. He then proposed his own PhD research and received funding to study a novel swarm-based sensor inspection system for enclosed environments. From 2013 he started this PhD project at University of Eindhoven in The Netherlands of which the results are presented in this dissertation. Since 2017 he is employed at Antea Group.