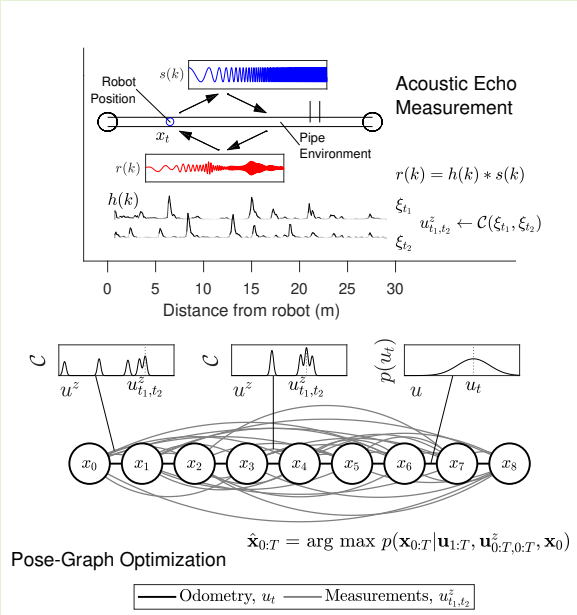
Acoustic Echo Sensing for Robot Localization in Buried Pipe Networks

Rob Worley, Yicheng Yu, Kirill V. Horoshenkov, and Sean R. Anderson

***Abstract*— Robot localization in buried pipes presents many challenges, including the unavailability of sensing methods such as a global positioning system, and the limited perspective of sensors such as vision. This paper addresses these challenges by using acoustic sensing, where sound emitted by robots propagates long distances and around corners in the pipe environment and is used to estimate the distance to reflective features. The reverberant environment produces many echoes that contain useful information for localization, but also produces many *second-order* echoes that have reflected from two features consecutively, which creates a challenge for localization. Therefore, a novel approach for using first and second-order echoes in pose-graph optimization for localization is developed here. Experiments demonstrated that the acoustic echo localization approach gave a relative estimate error of 0.24% (with 0.23% standard deviation). We compared to the results of a Kalman filter localization algorithm (3.4% error) and a pose-graph optimization alternative designed for only first-order echoes (0.80% error), and also compared to results from sensing approaches in the literature including vision sensing (2.17-2.27% error), acoustic sensing with a separate transmitter (3.5% error), and radio beacon sensing (0.047% error). This suggests that acoustic echo sensing surpasses these alternative methods, except those that require external hardware, and gives a 10-fold improvement on other approaches without external hardware. In summary, we have developed a new method for echo-localization in pipes, incorporating second-order echoes into pose-graph optimization, and demonstrated its effectiveness with respect to alternative approaches.**



***Index Terms*— Acoustic Sensing, Pipe Network Inspection, Robot Localization, SLAM, Processing of Acoustic Sensor Data**

I. Introduction

Autonomous robots could be used to improve the inspection and monitoring of buried pipe infrastructure. In the US there is around 2,000,000 kilometres of sewer pipes [1], while in the UK this figure is around 400,000 kilometres [2] along with 400,000 kilometres of water supply pipes. Faults in wastewater pipes can cause pollution and danger to public health, with around 50 serious pollution incidents per year in the UK [3]. Inspection of buried pipes is made difficult due to incomplete knowledge of the location and condition of pipes [1]. Autonomous persistent monitoring of this buried infrastructure could be done using robots to pervasively inspect a network of pipes and report the location of connected pipes and the position of faults.

Robot localization is essential in this application, which motivates this work. Its output is useful for large-scale navigation, small-scale control, for mapping an unexplored environment, and for locating the robot and any detected defects in an environment [4]. An illustration of the fault detection application is given in Figure 1, showing conventional tether sensor based localization estimation, and the alternative sensing approach proposed in this paper. Robust, accurate localization is required for the feasibility of robotic inspection, which will improve the efficiency and effectiveness of buried infrastructure monitoring and modelling, having positive public health and economic impacts.

Robot localization in this application is challenging as the pipe network adds a number of constraints [5]. For autonomous operation over a large area, the robots must be untethered and free to move, which limits communications and available power. The sensing and computing hardware are constrained in size and capability, and there is a constraint on

This work is supported by EPSRC grant EP/S016813/1 Pervasive Sensing for Buried Pipes (Pipebots).

R. Worley and S. R. Anderson are with the Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, UK. e-mail: {r.worley, s.anderson}@sheffield.ac.uk

Y. Yu and K. V. Horoshenkov are with the Department of Mechanical Engineering, University of Sheffield, Sheffield, UK. e-mail: {yicheng.yu, k.horoshenkov}@sheffield.ac.uk

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

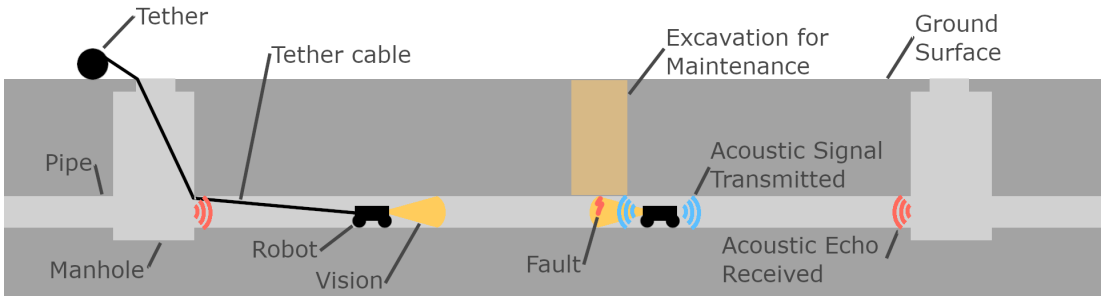


Fig. 1: A sketch of the pipe inspection and maintenance application. A cross section of a buried pipe is shown, connected to the surface by two manholes. There is a fault some distance along the pipe, which can be detected through short-range sensing such as vision. Maintenance will be done through a 1 m wide excavation, giving a target 0.5 m accuracy for fault location estimation. Conventionally, a tether cable is used to measure the position of the robot, and therefore the position of the fault. In this paper, acoustic echo sensing is proposed as an alternative which does not require the external hardware. It estimates the distance to reflective features such as manholes by comparing the transmitted and received acoustic signals.

what kinds of perception can be made at all. Crucially, the conventional use of the tether for localization, where a sensor measures the length of tether cable pulled out by the robot, is prohibited. Furthermore, GPS for measuring position and a magnetometer for measuring heading angle are unavailable or unreliable underground. Extrinsic beacons, as used in tunnel environments [6], require that transmitters are placed around the environment at known locations, which is prohibitive at the large scale and unobservability of a pipe network.

Sensing is an essential input to robot localization, and despite the limitations on sensing in this environment, common robot sensing has been applied and specialised to pipes. Vision has been used for odometry and mapping in gas pipes [7] and in our previous work in sewer pipes [8], as well as for detecting features such as pipe joints [9], and corners and junctions using shadows [10], [11], light [12], projected laser arrays [13], and depth images [14]. The same features can be detected with scanning rangefinders [15], and joints between pipes can be detected with inertial sensing [16]. While effective, cameras and rangefinders have drawbacks in the pipe environment compared to a typical indoor or outdoor application. Cameras require power-costly lights and are sensitive to lighting conditions in the dark environment of the pipe, while scanning rangefinders add to the size which is a problem in the narrow pipe environment, and both methods suffer from a lack of detectable features in a typical pipe. For example, good localization performance is found with scanning rangefinders in train tunnels only with the detection and use of railway features like tracks and power infrastructure, which are not available in the pipe environment [17]. Critically, in the pipe environment, these sensors can only observe nearby features, so *loop-closure* will be infrequent and localization will rely on odometry, which suffers from accumulating uncertainty, causing drift in the position estimate.

Alternative sensing methods specialised for pipes have been deployed on robots. A hydrophone sensor in water-filled pipes can measure a recognisable spatially varying property of the pipe. This measurement has been used in localization [18], developing on an approach designed for using magnetic signals detected by a drilling robot [19], but is limited to observing only nearby features like vision-based sensing. Radio waves can produce a recognisable signal which varies spatially along the length of a pipe [20], and the periodicity of this signal can be used as a reliable method of odometry. The time of travel of an acoustic signal has been used on a tethered robot to estimate the distance from a separate transmitter [21], which measures the robot's position relative to a known location, so the resulting estimate will not drift over time. However, these methods require separate hardware for signal transmission or a tether, which limit the robot's mobility. Radio signals from passive *RFID beacons* also create spatially varying signal. One use of this sensing [22] takes the method developed for hydrophone and magnetic signals [18], [19] and adds an additional calibration process. While effective, the performance would be affected when the spatial signal obtained in calibration is different to that measured in deployment in the heterogeneous and time-varying pipe environment. Furthermore, in practice, this method would require the deployment of beacons at a density of roughly one per metre of pipe, which be a very large amount of additional hardware at the scale of a pipe network where wastewater networks comprise hundreds of pipes with a typical length of 90 metres [1]. The signal strength from radio transmitters varies over space in a pipe in a complex way, which precludes the conventional use of signal strength as a measurement of distance between source and receiver used in other applications [23], [24].

Table I compares different sensing approaches described in the literature. The requirement for hardware external to the robot is noted, as this is a source of additional complexity and cost at the very large scale of buried pipe networks. It is noted whether the sensing approach allows for an absolute measurement of robot position with respect to the end of the pipe, a relative incremental measurement of position, or a semi-absolute measurement of position, meaning that the approach gives a measurement which is absolute with respect to a feature for which the position is unknown, or gives a set of possible absolute measurements. Finally, it is noted whether the approach has been demonstrated in real field conditions.

There is a general research gap and need, therefore, to develop sensors for robots that can perceive distant features

Sensing approach	Requires external hardware?	Absolute or relative position measurement?	Demonstrated in the field?
Radio Beacon [22] using Spatial Signal [18], [19]	Yes	Semi-absolute	No
Radio Periodic Signal [20]	Yes	Semi-absolute	Yes
Tethered Acoustic [21]	Yes	Absolute	No
Vision SLAM [8]	No	Relative	Yes
Vision Pipe Joint [9]	No	Semi-absolute	Yes
Hydrophone Spatial Signal [18]	No	Relative	No
Acoustic Echo	No	Absolute	No

TABLE I: Summary of sensing approaches in the literature. Here, an absolute measurement of position is with respect to the ends of the pipe, while a semi-absolute measurement is with respect to a feature for which the position is not known, or is a set of possible absolute measurements.

in pipes, increasing the perspective of the robot and allowing an absolute measurement of robot position, but that do not require any hardware that is not on the robot, such as tethers and beacons. This motivates the use of acoustic waves, emitted by the robot and echoing from the environment, which appear well suited to address this need. In our previous work, acoustic echo localization was performed using Kalman filtering [25], [26], which was demonstrated to be effective in a simple environment comprising a single pipe.

This paper addresses the specific research gap in the development of an acoustic echo sensing approach for localization in more realistic pipe network environments, where *lateral connections* and multiple *second-order* acoustic echoes are present. Where first-order echoes are direct reflections from a feature, second-order echoes occur where the sound wave has undergone two successive reflections. A large number of second-order echoes are produced by lateral connections along a pipe, creating a highly reverberant environment in which processing acoustic sensor measurements is very challenging. Correctly distinguishing second-order echoes from first-order echoes is very important because they can produce considerable error in localization otherwise, due to incorrect data association. This is a challenging problem because second-order echoes cannot be identified from the acoustic wave alone, but it is a crucial problem to address in order to develop robust acoustic echo sensing methods for localization in realistic pipe environments.

Therefore, in this paper we develop a novel localization algorithm based on pose-graph optimization, which is specially designed to incorporate acoustic sensor signals with multiple first-order *and* multiple second-order echoes into the pose-graph. The performance of this new algorithm is investigated and demonstrated on new, more realistic experimental data than previously used [25], [26], where the pipe environment has lateral connections. This performance is compared to a benchmark algorithm based on Kalman filtering, which is designed for first-order echoes and a single second-order echo [25], [26], and also compared to a simpler pose-graph optimization method which is designed for only first-order echoes. We show in the results that the Kalman filtering approach performs poorly on experimental data and the pose-graph optimization method for first-order echoes degrades in accuracy on increasingly realistic experimental data, whereas the new pose-graph optimization method incorporating second-order echoes performs much more accurately and robustly.

In summary, the aim of this paper is to develop a novel localization system using acoustic echo sensing for robots in pipes, and in particular to:

1) develop a novel pose-graph optimization algorithm specialized for acoustic echo sensing in the pipe environment, which is able to process multiple *second-order* echoes and echoes from lateral connections (Section III-C);
2) use new experimental data and benchmarking to show the necessity of a novel localization algorithm for acoustic sensing in this environment (Section IV-D);
3) evaluate the effectiveness of this localization method in the pipe environment over a range of uncertainty in inputs (Section IV-E).

Background on acoustic echo localization is given in section II. The problem definition and proposed solution are given in section III where the new localization algorithm is described. Results are given in section IV including benchmarking the new pose-graph optimization algorithm against a Kalman filter approach using real data. Conclusions are made in section V.

II. Background

While acoustic interference can be used to precisely detect walls close to a robot [27], many approaches using acoustic signals for robot localization compute the acoustic echo impulse response of the robot's surroundings, which is the transfer function between transmitter and receiver. Estimation of this in different environments with variation in position of acoustic transmitter and receiver has been studied for decades [28].

Robot localization can be done using measurements of direction of arrival of sounds from multiple sound sources [29]. In contrast, here time, rather than direction, of arrival of sounds will be considered, which is done using echoes rather than separate sound sources. A co-located acoustic source and receiver has been used for robot localization in a structured room [30], assuming that each echo is clear and unique, there are no multi-path echoes, and each detected echo can be associated with a specific wall in the room. However, in the pipe environment, these assumptions are not applicable.

Estimation of the impulse response can be improved [31] using a *maximum likelihood* (ML) (or *expectation maximization* (EM), or *nonlinear least-squares* (NLS)) optimization

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

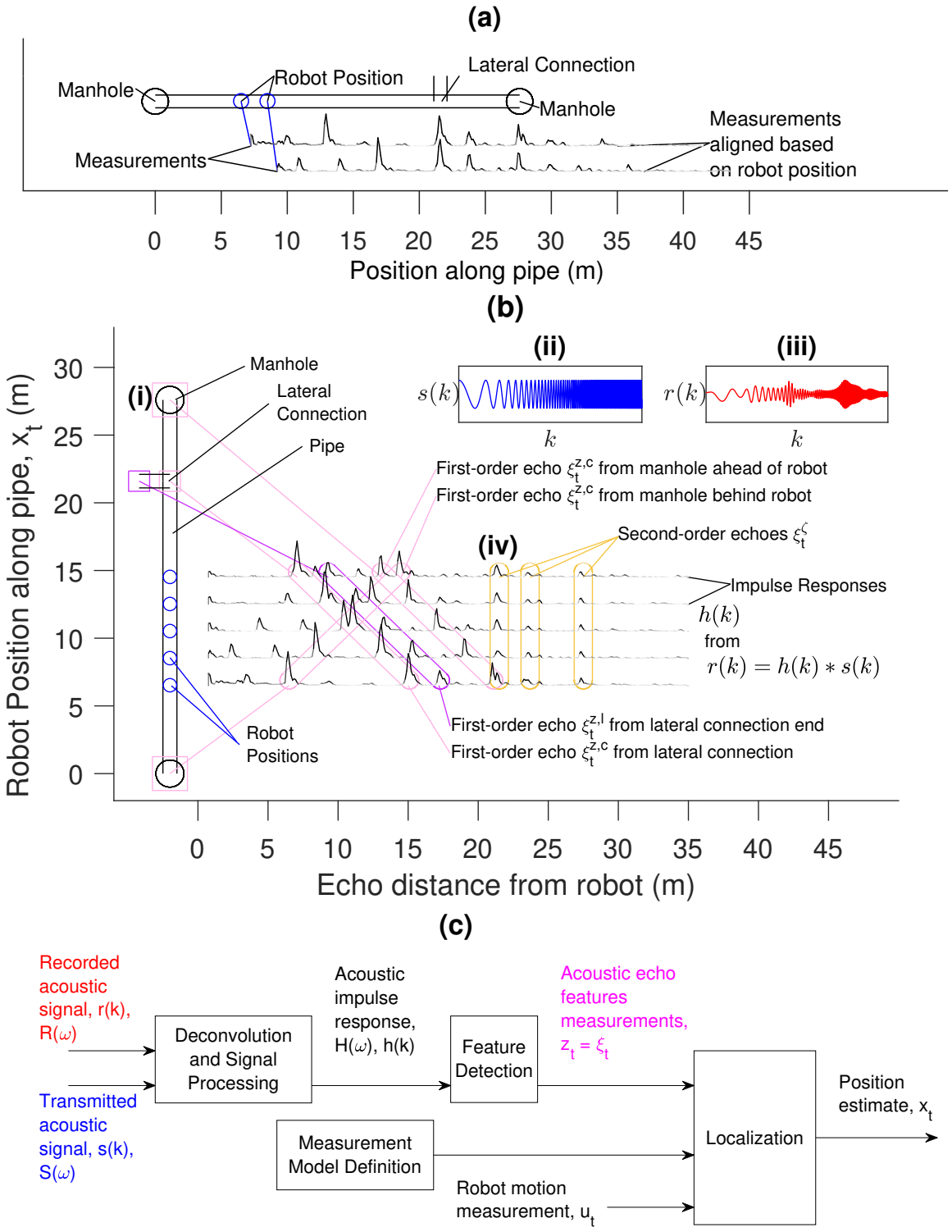


Fig. 2: **(a)** An illustration of acoustic echo sensing in a pipe network **(i)**. For five positions x_t = 6.5, 8.5, 10.5, 12.5, 14.5, the robot emits a signal $s(k)$ **(ii)** and receives a signal $r(k)$ **(iii)**. The computed impulse responses $h(k)$ are shown **(iv)** at the corresponding position x_t, where the first and second-order echoes are labelled. **(b)** An illustration of the system described in this paper, showing the link between the *signal processing*, *feature detection*, *measurement modelling*, and *localization*.

method [32]. The effects of signal to noise ratio, hardware transfer function [33], correlated and coloured noise, and faulty microphones [34] have been considered.

These methods are designed for room environments, and while the results are useful for developing similar methods for pipes, there are differences between the two environments. Pipes are acoustic waveguides, so sound is effectively constrained to one dimension. Acoustic signals can be used to detect features and leaks, and accurately estimate their one-dimensional distance from a sensor [35]. Sound travels long distances in pipes, with research investigating acoustic sensing in pipe lengths from 30 metres [36] to 4000 metres [37]. This is useful in wastewater networks where pipes have a typical length of 90 metres [1]; outside of robotics, acoustic echo sensing can detect features up to a range of 10 kilometres in gas pipes [38]. In room environments, multi-path transmission of sound produces a challenge in signal interpretation as the indirect path signal can be of similar amplitude to the direct path [39]. However in the low-dimensional pipe environment the signal components corresponding to different paths can be distinguished [40], [41], although the path cannot be easily identified from the signal component, and useful information might be found if the signal can be interpreted correctly.

In this work, a single microphone sensor is used. This is suitable for small robots designed for application in pipes, but therefore only the one-dimensional time of arrival of acoustic echoes can be measured. A microphone sensor array can be used to improve the acquisition of the impulse response by increasing the useful sensor frequency range, and improved signal processing can be used to reduce background noise [26]. A microphone sensor array can also allow the estimation of the circumferential direction of each acoustic echo, such as whether the echo has arrived from a blockage on the left or right side of the pipe cross-section, as well as improve the precision of measurement of distance for each echo [42].

Overall, acoustic echo sensing in a pipe environment is a promising source of information for robot localization. While development of acoustic echo sensing for robot localization in pipe networks is informed by the existing literature for other environments, autonomous robot localization in this environment requires further development of novel algorithmic aspects to make use of the available, but difficult to interpret, acoustic information.

III. Methods

Localization is performed by finding the acoustic impulse response of the environment using deconvolution, detecting features in the pipe, and estimating the robot position using optimization methods, as illustrated in Figure 2.

A. Problem Definition

Robot localization is the problem of estimating a robot's state with respect to the environment, and is typically solved with a probabilistic approach [43]. A continuous-discrete state definition is used, suitable for the pipe network environment [44], given by

$$\mathbf{x}_t = \left(q_t, a_t, x_t\right) \tag{1}$$

where $q_t \in \mathbb{Z}$ is the discrete index of the pipe where the robot is, $a_t \in \{-1, 1\}$ is the direction of the robot along the pipe, and $x_t \in \mathbb{R}$ is the one-dimensional position of the robot along the pipe, which is the variable to be estimated.

The robot moves to state $\mathbf{x}_t$ from state $\mathbf{x}_{t-1}$, modelled by the probability distribution $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t)$, so the odometry measurement is $\mathbf{u}_t = x_t - x_{t-1} + v_t$ where v_t is normally distributed uncertainty with variance Σ_t^u. The robot moves along a single pipe q_t over a series of time steps from $t = 0$ to $t = T$, with constant direction a_t.

At each time step, the robot makes observations $\mathbf{z}_t$ with acoustic echo sensing, given by the probability distribution $p(\mathbf{z}_t|\mathbf{x}_t)$. The observation $\mathbf{z}_t$ contains multiple measurements of distance $\boldsymbol{\xi}_t$ to features in the environment. Using the observations $\mathbf{z}_{t_1}$ and $\mathbf{z}_{t_2}$, an estimate of the relation $\mathbf{u}^z_{t_1,t_2}$ between states $\mathbf{x}_{t_1}$ and $\mathbf{x}_{t_2}$ can be made. The measurements $\mathbf{z}_t$ and estimates $\mathbf{u}^z_{0:T,0:T}$ have uncertainty Σ_t^z.

It is assumed that the pipe is part of a wastewater network. At each end of the pipe there is a *manhole*, and along the main pipe there are an arbitrary number of *lateral connections* to smaller pipes [1], illustrated in Figure 2(a). Pipes vary in length but are an average of 90 metres long [1]. In practice, the manholes may be observable from above ground, but here no prior knowledge of their position, and therefore of the pipe length, is assumed.

The uncertainties Σ_t^z and Σ_t^u motivate the problem and the probabilistic approach. The problem addressed is the estimation of posterior probability distribution $p(\mathbf{x}_{0:T}|\mathbf{u}_{1:T}, \mathbf{z}_{0:T}, \mathbf{x}_0)$ of the trajectory of the robot from measurements from odometry $\mathbf{u}_t$ and acoustic echoes $\mathbf{z}_t$, and knowledge of the initial position $\mathbf{x}_0$.

B. Measurement Model Definition

At a point in time t, the robot stops moving and emits an acoustic signal $s(k)$ over discrete time k, for a period K, before moving and repeating this process at time $t + 1$. Over this period, the robot records the acoustic signal received from the environment $r(k)$. An acoustic echo impulse response $h(k)$ is found using deconvolution, which is computed in the frequency (ω) domain as $H(\omega) = R(\omega)/S(\omega)$, where S, R, and H are the frequency domain representations of s, r, and h respectively. Examples of the transmitted and received acoustic signals are shown in Figures 2(b)(i) and (ii), and acoustic impulse response signal envelopes are shown in Figures 2(a) and 2(b)(iv). Peaks in the impulse response envelope correspond to echoes from reflective features in the environment.

Acoustic echoes will be produced by the manholes at each end of the pipe and by each end of the lateral connections. A map of the position of these reflective features in pipe q is

$$\mathcal{M}^q = \left\{\mathcal{M}^q_c, \mathcal{M}^q_l\right\} \tag{2}$$

where $\mathcal{M}^q_l$ and $\mathcal{M}^q_c$ describe the length and position of each connection or manhole, and where the length of a manhole is defined to be zero. For example, the map is

$$\mathcal{M}^q = \left\{\mathcal{M}^q_c = \left\{0, 21.6, 27.6\right\}, \mathcal{M}^q_l = \left\{0, 2.2, 0\right\}\right\}$$

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

for the 27.6 m long pipe with a manhole at each end, and a 2.2 m long lateral connection 21.6 m along the pipe shown in Figure 2(a)(i).

From the envelope of $h(k)$, the one-dimensional distance ξ_t^ϕ to a feature ϕ in the environment can be estimated by locating peaks at times κ^ϕ in the time (k) domain signal and converting to distance $\xi^\phi = c_a \kappa^\phi$ using the known acoustic wave speed c_a. Feature detection can be done automatically [45]. This is the distance through the pipe network, not the euclidean distance. The measurement model $p(\mathbf{z}_t|\mathbf{x}_t)$ consists of three sets of components, $\boldsymbol{\xi}_t^{z,c}$, $\boldsymbol{\xi}_t^{z,l}$, and $\boldsymbol{\xi}_t^{\zeta}$.

1) First order echoes: The first set describes the direct (*first-order*) echoes from the lateral connections and manholes along the main pipe

$$\boldsymbol{\xi}_t^{z,c} = \left\{ |m^\phi - x_t| : m^\phi \in \mathcal{M}_c^q \right\} \tag{3}$$

The second set describes the direct echoes from the features at the end of each of the lateral connections, and is given by

$$\begin{aligned} \boldsymbol{\xi}_t^{z,l} = \Big\{ & |m_l^\phi - x_t| + m_l^\phi : \\ & \phi \in \Big\{1, \ldots, \Phi\Big\}, m_c^\phi \in \mathcal{M}_c^q, m_l^\phi \in \mathcal{M}_l^q \Big\} \end{aligned} \tag{4}$$

where Φ is the number of features.

2) Second order echoes: The third set describes the second-order echoes from two features, where the sound travels from the robot, to one feature, then another, then back to the robot. When the robot is between any two features in the environment, there will always be a *second-order* echo which travels a total distance equal to twice the distance between those two features. These echoes pose a challenge for localization, as they appear to correspond to a feature which is a fixed distance from the robot regardless of the robot's position. Multiple second-order echoes will be present in a realistic environment with many lateral connections between pipes because the robot is between many pairs of features at any time. The result is a highly reverberant environment. This set of second-order echoes is given by

$$\begin{aligned} \boldsymbol{\xi}_t^{\zeta} = \Big\{ & |m_c^{\phi_1} - m_c^{\phi_2}| + n_l^{\phi_1} m_l^{\phi_1} + n_l^{\phi_2} m_l^{\phi_2} : \\ & \Big\{ \phi_1, \phi_2 \in \left\{1, \ldots, \Phi\right\}, m_c^{\phi_1}, m_c^{\phi_2} \in \mathcal{M}_c^q, \\ & m_l^{\phi_1}, m_l^{\phi_2} \in \mathcal{M}_l^q : m_c^{\phi_1} < x_t, m_c^{\phi_2} > x_t \Big\}, \\ & n_l^{\phi_1}, n_l^{\phi_2} \in \left\{0, 1\right\} \Big\} \end{aligned} \tag{5}$$

While this description of the set of components is complicated mathematically, it is simple to model programatically.

For the example illustrated in Figure 2(a) and (b), for $x_t = 6.5$ these echoes are given by $\boldsymbol{\xi}_t^{z,c} = \{6.5, 15.1, 21.1\}$, $\boldsymbol{\xi}_t^{z,l} = \{17.3\}$, and $\boldsymbol{\xi}_t^{\zeta} = \{21.6, 23.8, 27.6\}$.

C. Pose-Graph Optimization Overview

A practical means of finding the posterior probability distribution $p(\mathbf{x}_{0:T}|\mathbf{u}_{1:T}, \mathbf{z}_{0:T}, \mathbf{x}_0)$ over a sequence of robot poses is needed. The simple approach uses typical pose-graph optimization designed for only first-order acoustic echoes (PGO1$^{\text{st}}$) and the benchmark approach uses a Kalman filter algorithm specialised to a single pipe (KF) [25] to sequentially estimate the posterior $p(\mathbf{x}_t|\mathbf{u}_{1:t}, \mathbf{z}_{0:t}, \mathbf{x}_0)$. Both perform poorly when using measurements from realistic pipe networks, which is shown in Section IV-D. This paper proposes an algorithm which incorporates information from multiple second-order acoustic echoes in the highly reverberant pipe environment into pose-graph optimization (PGO2$^{\text{nd}}$).

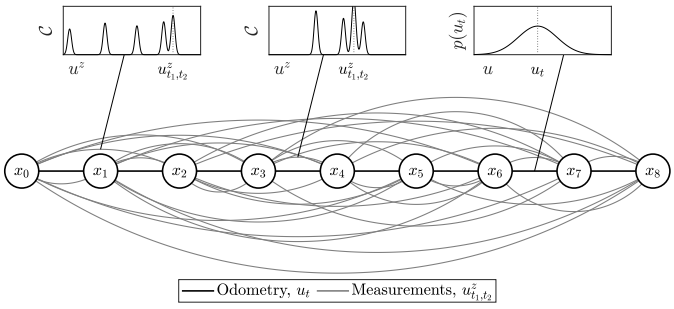


Fig. 3: An example pose-graph $\mathcal{G}$, showing vertices $\mathcal{V} = \mathcal{X} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$ and edges $\mathcal{E} = \mathcal{U} = \{\mathbf{u}_{0:T}, \mathbf{u}^z_{0:T,0:T}\}$.

Pose-graph optimization effectively aims to find an estimated trajectory $\hat{\mathbf{x}}_{0:T}$ which produces estimated measurements $\hat{\mathbf{u}}_t$ and $\hat{\mathbf{u}}^z_{t_1,t_2}$ which give the minimum error, in a sense, when compared to the observed measurements $\mathbf{u}_t$ and $\mathbf{u}^z_{t_1,t_2}$. Specifically, this the sum of weighted squared error, so is akin to weighted least squares estimation. This is defined as follows.

A pose-graph $\mathcal{G} = (\mathcal{V} = \mathcal{X}, \mathcal{E} = \mathcal{U})$ has vertices $\mathcal{V} = \mathcal{X} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$, and edges $\mathcal{E} = \mathcal{U}$ comprising the measurements $\mathbf{u}_t$ and $\mathbf{u}^z_{t_1,t_2}$ between each pair of poses. An example is illustrated in Figure 3.

Error terms are defined between each edge in $\mathcal{E}$

$$e_t = \hat{\mathbf{u}}_t - \mathbf{u}_t = \hat{\mathbf{x}}_t - \hat{\mathbf{x}}_{t-1} - \mathbf{u}_t \tag{6}$$

$$e^z_{t_1,t_2} = \hat{\mathbf{u}}^z_{t_1,t_2} - \mathbf{u}^z_{t_1,t_2} = \hat{\mathbf{x}}_{t_2} - \hat{\mathbf{x}}_{t_1} - \mathbf{u}^z_{t_1,t_2} \tag{7}$$

which is the difference between the estimated $(\hat{\mathbf{u}}_t, \hat{\mathbf{u}}_{t_1,t_2})$ and measured $(\mathbf{u}_t, \mathbf{u}^z_{t_1,t_2})$ relative pose. A cost function is defined

$$\mathbf{F}(\hat{\mathbf{x}}_{0:T}) = \sum_t e_t^{\mathrm{T}} \Sigma_{e_t}^{-1} e_t + \sum_{t_1,t_2} e^{\mathrm{T}}_{t_1,t_2} \Sigma^{-1}_{e_{t_1,t_2}} e_{t_1,t_2} \tag{8}$$

with uncertainties Σ_{e_t} and $\Sigma_{e_{t_1,t_2}}$, and the objective for robot localization is to obtain an optimal estimate of the set of poses $\hat{\mathbf{x}}_{0:T}$ that minimise $\mathbf{F}(\hat{\mathbf{x}}_{0:T})$.

The challenge is in correctly constructing the pose-graph $\mathcal{G}$, from both first order and second order acoustic echoes, and therefore $\mathbf{F}(\hat{\mathbf{x}}_{0:T})$, so that the optimization gives the desired estimate. Therefore, the challenge is in accurately finding the measurements $\mathbf{u}^z_{0:T,0:T}$ from the observations $\mathbf{z}_{0:T}$.

D. Pose-Graph Optimization using First-Order Echoes Only (PGO1st)

For comparison to the proposed PGO2$^{\text{nd}}$ algorithm, a simpler pose-graph optimization algorithm, PGO1$^{\text{st}}$, is implemented which is designed with the assumption that *second-order* echoes have been removed from the measurement sets

leaving only *first-order* echoes, and that the robot does not pass lateral connections along the pipe axis.

A *scan-matching* [46] approach is used to estimate the measurements $\mathbf{u}^z_{0:T,0:T}$ between poses using the acoustic measurements, illustrated in Figure 3. This effectively estimates the correspondence, or data association, between acoustic echoes observed at different times. These are given by

$$u^z_{t_1,t_2} = \arg\max_{u^z} \left\{ \mathcal{C}_1 \left(\mathbf{x}^\xi_{t_1} + u^z, \mathbf{x}^\xi_{t_2} \right) \right\} \tag{9}$$

where $\mathbf{x}^\xi_{t_1} = \left\{ -\boldsymbol{\xi}_{t_1}, \boldsymbol{\xi}_{t_1} \right\}$ and $\mathbf{x}^\xi_{t_2} = \left\{ -\boldsymbol{\xi}_{t_2}, \boldsymbol{\xi}_{t_2} \right\}$, and where $\mathcal{C}_1$ is a correlation function. This has been specialised to the acoustic echo measurements here, with some simplification as the space has one dimension rather than two, but some additional complexity as the measurements are directionless.

The scan-matching estimates are typically accurate when there is low measurement uncertainty, but outlier inaccurate values are possible. A robust pose-graph optimization approach [47], based on *graduated nonconvexity*, is used where estimates $u^z_{t_1,t_2} \in \mathcal{U}$ which cause a large amount of residual error in the optimization are removed from the pose-graph $\mathcal{G}$. Typically there are sufficiently many accurate measurements so the estimation is not compromised if some are incorrectly removed.

Due to the simplifying assumptions made in the design of this algorithm, it is seen to perform poorly when using measurements from realistic pipe networks, as shown in Section IV-D.

E. Pose-Graph Optimization using First- and Second-Order Echoes (PGO2nd)

An algorithm is needed to address the challenges of acoustic measurements in realistic pipe environments, which obtains accurate estimates $\mathbf{u}^z_{0:T,0:T}$ from the acoustic measurements. Estimates $u^z_{t_1,t_2}$ are found as in equation 9 from a correlation function $\mathcal{C}$. This function is generally multimodal, as illustrated in Figure 3, and the uncertainties in measurements can mean that a *wrong* value can be assigned to $u^z_{t_1,t_2}$. The presence of *second-order* acoustic echoes $\boldsymbol{\xi}^\zeta_t$ affects the correlation function between measurements at all time indices, as they appear to correspond to features which are a fixed distance from the robot at any time. The presence of *first-order* acoustic echoes $\boldsymbol{\xi}^z_t = \{\boldsymbol{\xi}^{z,c}_t, \boldsymbol{\xi}^{z,l}_t\}$ from both from features along the main pipe and from the far end of lateral connections affects the correlation function between measurements corresponding to poses in different sections of the pipe. Therefore, these echoes need to be classified and incorporated differently.

The algorithm is outlined as follows:

1) $\left[\tau_{0:J}\right] = \texttt{estimate-measurements-location}\left(\boldsymbol{\xi}_{0:T}, \Sigma^z_t\right)$
2) $\left[\hat{\boldsymbol{\xi}}^\zeta_{0:T}, \hat{\boldsymbol{\xi}}^z_{0:T}\right] = \texttt{estimate-measurement-class}\left(\boldsymbol{\xi}_{0:T}, \tau_{0:T}, \Sigma^z_t\right)$
3) $\left[(\mathbf{x}^\xi_{0:T}, \mathbf{l}^\xi_{0:T})\right] = \texttt{estimate-feature-positions}\left(\hat{\boldsymbol{\xi}}^z_{0:T}, \mathbf{u}_{0:T}, \tau_{0:T}, \Sigma^u_t, \Sigma^z_t\right)$
4) $\left[\mathbf{u}^z_{0:T,0:T}\right] = \texttt{estimate-feature-relations}\left((\mathbf{x}^\xi_{0:T}, \mathbf{l}^\xi_{0:T}), \mathbf{u}_{1:T}, \tau_{0:T}, \Sigma^u_t, \Sigma^z_t\right)$

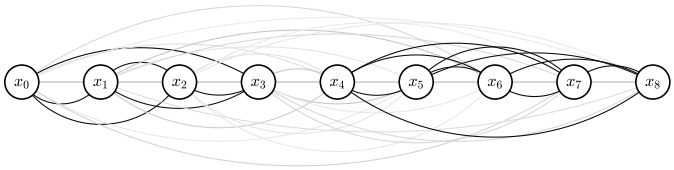


Fig. 4: An auxiliary pose-graph $\mathcal{G}^\mathcal{P}$, with vertices $\mathcal{V} = \mathcal{X} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$ and edges $\mathcal{E} = \mathcal{P}$, with darkness corresponding to probability $\hat{p}^\zeta_{t_1,t_2} \in \mathcal{P}$ that the edge connects vertices with poses in the same section of pipe.

5) $\left[\hat{\mathbf{x}}_{0:T}, (\hat{\mathbf{x}}^\xi_{0:T}, \hat{\mathbf{l}}^\xi_{0:T})\right] = \texttt{pose-graph-optimization}\left(\mathbf{u}^z_{0:T,0:T}, \mathbf{u}_{1:T}, x_0, \Sigma^u_t, \Sigma^z_t\right)$

These steps are detailed in the following subsections.

1) Estimate Measurement Location: Poses in the same section of pipe between two lateral connections are grouped together as this facilitates the classification of echoes as first-order echoes $\boldsymbol{\xi}^z_t$ and second-order echoes $\boldsymbol{\xi}^\zeta_t$ and the estimation of $u^z_{t_1,t_2}$ between poses in different sections of the pipe. This is done by constructing and partitioning an auxiliary pose-graph $\mathcal{G}^\mathcal{P}$. This process is described here and summarised in Algorithm 1.

A weighted pose-graph $\mathcal{G}^\mathcal{P} = (\mathcal{V} = \mathcal{X}, \mathcal{E} = \mathcal{P})$ is constructed, where the edges $\mathcal{E} = \mathcal{P}$ between each vertex correspond to the probability that the two vertices, or poses, $\mathcal{V} = \mathcal{X}$, are within the same section of pipe. This is illustrated in Figure 4. This probability can be estimated using the fact that measurements at poses in the same section of pipe will have the same *second-order* acoustic echoes, so the sets of measurements will have some similarity independent of their corresponding pose. A probability distribution over the *second-order* measurement space ξ^ζ_t for each time index is created from the set of measurements $\boldsymbol{\xi}_t$, as a sum of normal distributions $\mathcal{N}$ with mean μ and variance σ^2 for each element $\xi_n \in \boldsymbol{\xi}_t$

$$p(\xi^\zeta_t | \boldsymbol{\xi}_t) = \sum_n \mathcal{N}(\mu = \xi_n, \sigma^2 = \Sigma^z_t) \tag{10}$$

The similarity between sets of measurements at different time indices is estimated using the *Kullback-Liebler Divergence* $D_{t_1,t_2} = D(p(\xi^\zeta_{t_1}|\boldsymbol{\xi}_{t_1}) || p(\xi^\zeta_{t_2}|\boldsymbol{\xi}_{t_2}))$. The directionless values $D^\zeta_{t_1,t_2} = \frac{1}{2}\left(D_{t_1,t_2} + D_{t_2,t_1}\right)$ for each pair of time indices give a set of divergences $D^\zeta_{0:T,0:T}$. The divergence values in $D^\zeta_{0:T,0:T}$ tend to fall into two groups d^ζ_1 and d^ζ_2, which are identified using *k-means* clustering. The set d^ζ_κ with lower values is used to find a threshold ϵ^d as

$$\epsilon^d = \texttt{mean}\left(d^\zeta_\kappa\right) + \texttt{standard-deviation}\left(d^\zeta_\kappa\right) \tag{11}$$

using the mean and standard deviation of the values in d^ζ_κ.

The probability $\hat{p}^\zeta_{t_1,t_2}$ that two measurement sets were taken in the same pipe section is estimated using the divergence values $D^\zeta_{t_1,t_2}$ and the computed threshold ϵ^d, as well as the

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

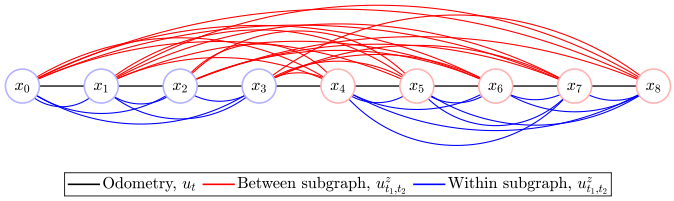


Fig. 5: An example pose-graph $\mathcal{G}$, showing vertices $\mathcal{V} = \mathcal{X} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$, coloured according to the estimated subgraphs $\mathcal{X}_{\boldsymbol{\tau}_j \in \boldsymbol{\tau}}$, and edges $\mathcal{E} = \mathcal{U} = \{\mathbf{u}_{0:T}, \mathbf{u}^z_{0:T,0:T}\}$, coloured according to the between subgraph and within subgraph measurements.

relative pose predicted by odometry $\mathbf{u}_{1:T}$, with

$$\hat{p}^{\zeta}_{t_1,t_2} = -\mathcal{S}(D^{\zeta}_{t_1,t_2}, d_0 = \epsilon^d, \delta = \Sigma^{z\,-1}_t) \\ \mathcal{N}(\sum_{t=t_1}^{t_2} \mathbf{u}_t, \mu = 0, \sigma^2 = \Sigma^u_t(t_2 - t_1)) \tag{12}$$

where $\mathcal{S}$ is a logistic sigmoid function with a midpoint d_0 and steepness δ, and $\mathcal{N}$ is a normal distribution with mean μ and variance σ^2. The weighted pose-graph $\mathcal{G}^{\mathcal{P}} = (\mathcal{V} = \mathcal{X}, \mathcal{E} = \mathcal{P})$ with edges $\mathcal{P}$ made of $\hat{p}^{\zeta}_{t_1,t_2}$ is constructed.

Spectral graph partitioning is used to separate the weighted pose-graph $\mathcal{G}^{\mathcal{P}}$ into discrete subgraphs $\mathcal{X}_{\boldsymbol{\tau}_j \in \boldsymbol{\tau}}$ with dissimilar echo measurements. This is done by finding the eigenvector $\mathbf{v}_{0:T}$ which corresponds to the second smallest eigenvalue of the *graph Laplacian*. The difference between the value of $\mathbf{v}_{0:T}$ corresponding to each time index is found, creating the vector $\Delta\mathbf{v}_{0:T}$ with elements

$$\Delta v_t = v_t - v_{t-1}, t \in \{1, 2, \ldots, T\}, v_t, v_{t-1} \in \mathbf{v}_{0:T} \tag{13}$$

Large changes in $\mathbf{v}_{0:T}$ should occur at boundaries between subgraphs corresponding to different pipe sections. The boundary time indices are

$$\mathbf{t}^{\zeta}_{j \in 0,1,\ldots,J} = \{\texttt{find-peaks}\,(|\Delta\mathbf{v}_{0:T}|)\} + \{0, T\} \tag{14}$$

and therefore the set $\boldsymbol{\tau}$ of sets of time indices corresponding to the subgraphs for each pipe section is

$$\boldsymbol{\tau} = \{\boldsymbol{\tau}_j = \{\mathbf{t}^{\zeta}_j + 1, \mathbf{t}^{\zeta}_j + 2, \ldots, \mathbf{t}^{\zeta}_{j+1}\} : \\ j \in \{0, 1, \ldots, J-1\}\} \tag{15}$$

which gives the partition of the graph vertices into $\mathcal{X}_{\boldsymbol{\tau}_j \in \boldsymbol{\tau}}$. The results of this is illustrated in Figure 5, where the vertices corresponding to each of two subgraphs are coloured accordingly.

2) Estimate Measurement Class: With this estimation of relevant subgraphs with time indices $\boldsymbol{\tau}_j \in \boldsymbol{\tau}$, the function `estimate-measurement-class` estimates which measurements are the second-order acoustic measurements by comparing measurement sets corresponding to poses in each section of pipe. The probability distributions across the one-dimensional pipe axis space are summed, and the largest peaks in probability are identified as those which likely correspond

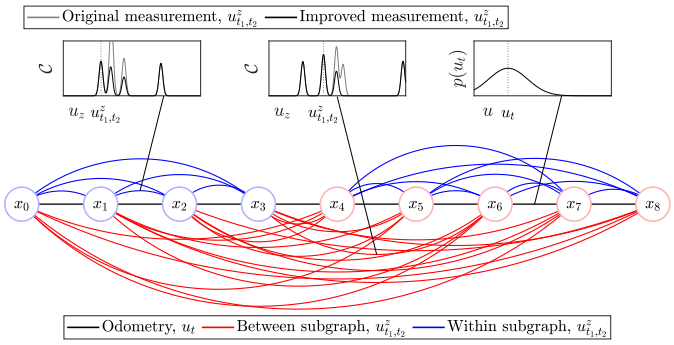


Fig. 6: An example pose-graph $\mathcal{G}$, showing vertices $\mathcal{V} = \mathcal{X} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$ and edges $\mathcal{E} = \mathcal{U} = \{\mathbf{u}_{0:T}, \mathbf{u}^z_{0:T,0:T}\}$ with improved measurements $\mathbf{u}^z_{t_1,t_2}$. Subgraphs $\mathcal{X}_{\boldsymbol{\tau}_j \in \boldsymbol{\tau}}$ corresponding to sections of pipe are indicated with coloured vertices. Coloured edges indicate the measurements between and within these subgraphs.

to second-order measurements as they have been observed multiple times. This is described by

$$\boldsymbol{\xi}^{\zeta}_{\boldsymbol{\tau}_j} = \texttt{find-peaks}\left\{\sum_{t \in \boldsymbol{\tau}_j} p(\xi^{\zeta}_t | \boldsymbol{\xi}_t)\right\} \tag{16}$$

The second-order measurements can be assigned as $\boldsymbol{\xi}^{\zeta}_t = \boldsymbol{\xi}^{\zeta}_{\boldsymbol{\tau}_j}$ for $t \in \boldsymbol{\tau}_j$. The first-order measurements are therefore given by set subtraction $\{\boldsymbol{\xi}^z_t\} = \{\boldsymbol{\xi}_t\} - \{\boldsymbol{\xi}^{\zeta}_t\}$.

This classification of measurements improves the estimation of $\mathbf{u}^z_{0:T,0:T}$ for all time indices, as the correlation function $\mathcal{C}$ from which $\mathbf{u}^z_{0:T,0:T}$ is estimated will have fewer confounding modes.

3) Estimate Feature Positions: The possible positions of the features corresponding to each first-order measurement $\boldsymbol{\xi}^z_t$ must be estimated. The measurements contain echoes from each end of the lateral connections, $\boldsymbol{\xi}^z_t = \{\boldsymbol{\xi}^{z,c}_t, \boldsymbol{\xi}^{z,l}_t\}$, where there are Ψ measurements in total at time t. From each measurement it cannot be determined whether a feature is the near ($\boldsymbol{\xi}^{z,c}_t$) or the far ($\boldsymbol{\xi}^{z,l}_t$) end of a lateral connection. A two-dimensional space (x, l) is defined, where x is the position along the pipe axis, and l is the distance from the pipe axis. Each individual measurement $\xi^{\psi_1}_t$, for $\psi_1 \in \{1, 2, \ldots, \Psi\}$, is projected into this space, and occupies a point in the linear subspace

$$\pm\xi^{\psi_1}_t = x^{\psi_1}_t + l^{\psi_1}_t \tag{17}$$

A lateral connection will exist as a pair of features $(x^{\psi_1}_t, l^{\psi_1}_t)$, $(x^{\psi_2}_t, l^{\psi_2}_t)$ at the same position along the pipe axis x, so hypothetically a second measurement $\xi^{\psi_2}_t$ could form this pair, for $\psi_2 \in \{1, 2, \ldots, \Psi\}$ where $\psi_2 \neq \psi_1$. If $\xi^{\psi_2}_t < \xi^{\psi_1}_t$, then the features can be computed as

$$(x^{\psi_2}_t, l^{\psi_2}_t) = (\xi^{\psi_2}_t, 0) \tag{18}$$

$$(x^{\psi_1}_t, l^{\psi_1}_t) = (x^{\psi_2}_t, \xi^{\psi_1}_t - x^{\psi_1}_t) \tag{19}$$

These hypothetical features are collected in $(\mathbf{x}^{\xi}_{0:T}, \mathbf{l}^{\xi}_{0:T})$.

This classification of measurements improves the estimation of $\mathbf{u}^z_{0:T,0:T}$ for pairs of time indices corresponding to poses in different sections of pipe.

Algorithm 1 Estimate Measurement Location

input $\boldsymbol{\xi}_{1:T} = \{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_T\}$, $\mathbf{u}_{1:T}$, Σ_t^z, Σ_t^u

Get probability distribution of second-order measurement ξ_t^ζ over space for each measurement set $\boldsymbol{\xi}_t$
for $t = \{1, 2, \ldots, T\}$ **do**
$p(\xi_t^\zeta | \boldsymbol{\xi}_t) = \sum_n \mathcal{N}(\mu = \xi_n, \sigma^2 = \Sigma_t^z)$
end for

Measure the difference between each pair of distributions using the Kullback-Liebler Divergence, creating matrix $D_{0:T,0:T}$
for $(t_1, t_2) = \{(1,1), (1,2), \ldots, (2,1), \ldots, (T,T)\}$ **do**
$D_{t_1,t_2} = D(p(\xi_{t_1}^\zeta | \boldsymbol{\xi}_{t_1}) || p(\xi_{t_2}^\zeta | \boldsymbol{\xi}_{t_2}))$
$= \sum_{\xi^\zeta} p(\xi_{t_1}^\zeta | \boldsymbol{\xi}_{t_1}) \log \left(\frac{p(\xi_{t_1}^\zeta | \boldsymbol{\xi}_{t_1})}{p(\xi_{t_2}^\zeta | \boldsymbol{\xi}_{t_2})} \right)$
end for
$D_{0:T,0:T}^\zeta = \frac{1}{2} \left(D_{0:T,0:T} + D_{0:T,0:T}{}^{\mathrm{T}} \right)$

Determine threshold for similar measurement sets
$\{d_1^\Lambda, d_2^\Lambda\} = \texttt{k-means} \left(D_{0:T,0:T}^\zeta, k = 2 \right)$
$\kappa = \arg\min_{k \in \{1,2\}} \left\{ \texttt{mean} \left(d_k^\zeta \right) \right\}$
$\epsilon^d = \texttt{mean} \left(d_\kappa^\zeta \right) + \texttt{standard-deviation} \left(d_\kappa^\zeta \right)$

Estimate probability that two measurement sets were taken in the same pipe section using sigmoid $\mathcal{S}$ and Gaussian $\mathcal{N}$
$\hat{p}_{t_1,t_2}^\zeta = -\mathcal{S}(D_{t_1,t_2}^\zeta, d_0 = \epsilon^d, \delta = \Sigma_t^{z\,-1})$
$\mathcal{N}(\sum_{t=t_1}^{t_2} \mathbf{u}_t, \mu = 0, \sigma^2 = \Sigma_t^u (t_2 - t_1))$

Construct pose-graph with adjacency matrix $\hat{P}_{0:T,0:T}^\zeta$ with elements $\hat{p}_{t_1,t_2}^\zeta$

Do spectral graph partitioning
$L_{0:T,0:T}^\zeta = \texttt{graph-laplacian} \left(\hat{P}_{0:T,0:T}^\zeta \right)$
find $\mathbf{v}_{0:T}$ *using* $\lambda_2 = \mathbf{v}_{0:T}{}^{\mathrm{T}} L_{0:T,0:T}^\zeta \mathbf{v}_{0:T}$ *where* λ_2 *is the second smallest eigenvalue of* $L_{0:T,0:T}^\zeta$

Define subgraphs with time indices $\boldsymbol{\tau}$
$\Delta v_t = v_t - v_{t-1}, t \in \{1, 2, \ldots, T\}, v_t, v_{t-1} \in \mathbf{v}_{0:T}$
$\mathbf{t}_{j \in 0,1,\ldots,J}^\zeta = \{\texttt{find-peaks}(|\Delta \mathbf{v}_{0:T}|)\} + \{0, T\}$
$\boldsymbol{\tau} = \{\boldsymbol{\tau}_j = \{\mathbf{t}_j^\zeta + 1, \mathbf{t}_j^\zeta + 2, \ldots, \mathbf{t}_{j+1}^\zeta\} : j \in \{0, 1, \ldots, J-1\}\}$

4) Estimate Feature Relations: Each set of features $\left(\mathbf{x}_{t_1}^\xi, \mathbf{l}_{t_1}^\xi\right)$, $\left(\mathbf{x}_{t_2}^\xi, \mathbf{l}_{t_2}^\xi\right)$ at all time indices $t_1, t_2 \in \{1, 2, \ldots, T\}$ are compared and the value of u_{t_1,t_2}^z, an estimate of the relation between poses at time indices t_1 and t_2, is found using

$$u_{t_1,t_2}^z = \arg\max_{u^z} \left\{ \mathcal{C}_2 \left(\left(\mathbf{x}_{t_1}^\xi + u^z, \mathbf{l}_{t_1}^\xi \right), \left(\mathbf{x}_{t_2}^\xi, \mathbf{l}_{t_2}^\xi \right) \right) \right\} \quad (20)$$

where $\mathcal{C}_2$ is a two-dimensional correlation function. As in equation 9, this effectively estimates the correspondence, or data association, between acoustic echoes observed at different times.

5) Pose-Graph Optimization: Using the output from the above algorithm, the pose-graph $\mathcal{G} = (\mathcal{V} = \mathcal{X}, \mathcal{E} = \mathcal{U})$ is constructed with improved measurements $u_{t_1,t_2}^z \in \mathcal{U}$ between poses $\mathbf{x}_{t_1}, \mathbf{x}_{t_2} \in \mathcal{X}$. This is illustrated in Figure 6, where the indicated measurements u_{t_1,t_2}^z between subgraphs are improved by the estimation of feature positions, and measurements u_{t_1,t_2}^z between all time indices are improved by the classification of *second-order* measurements. Subsequently, pose-graph optimization is done to estimate the sequence of robot poses using $\mathbf{u}_{0:T,0:T}^z$ and $\mathbf{u}_{1:T}$. Robust pose-graph optimization as described in Section III-D for the *simple* algorithm is used here.

F. Benchmark Algorithm

This algorithm estimates the trajectory $\mathbf{x}_{0:T}$ by iteratively finding $p(\mathbf{x}_t | \mathbf{u}_{1:t}, \mathbf{z}_{0:t}, \mathbf{x}_0)$ using Kalman filtering (KF). The algorithm is specialised to simple single pipe environments as it detects a single *second-order* echo and incorporates it into the estimation differently from first-order echoes [25].

IV. Results

In this section, experimental data is used to demonstrate, evaluate and benchmark the acoustic echo sensing system and localization algorithm. Experiments investigate the uncertainty in the acoustic echo sensing by measuring the echo detection precision and recall with a variation in three sensor parameters. Acoustic echo experimental data is extended by adding simulated uncertainty, which is used to compare the PGO1st and PGO2nd pose-graph localization algorithms and to evaluate the performance of the proposed localization algorithm over a range of values of uncertainty in robot motion and perception.

A. Sensing with Acoustic Echoes

The sensitivity of the acoustic echo sensing to three parameters is investigated here: the sensing duration, the position of the sensor in the pipe cross-section, and the background noise in the environment. A PYLE PDS122 loudspeaker and GRAS 46AE microphones were used to investigate the sensing across a 300 mm diameter pipe cross-section, and a Visaton 2242 loudspeaker and a MSM321A3729H9CP MEMSensing Microsystems Co., Ltd. microphone were used to investigate the other two parameters in a 150 mm diameter pipe.

The effect of these parameters on the echo measurement is measured by the *precision* and *recall* of the feature detection system. In terms of feature detection, precision is defined as the ratio of true positives to the sum of true positives and false positives, and recall is defined as the ratio of true positives to the sum of true positives and false negatives.

The echo detection system which produces a set of detected features $\boldsymbol{\xi}_t$ is described in Section III-B. The echo detection threshold is 0.02, which is relative to the amplitude of the signal. The detected features can be compared to a set of expected features for a given sensor position, as described in Section III-B, to calculate the precision and recall of the detection.

Figure 7(a) illustrates the acoustic echo measurements made for a range of lengths of sensing duration. This length impacts the robot's operation and energy consumption as it should be stationary while outputting an acoustic signal for the duration.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

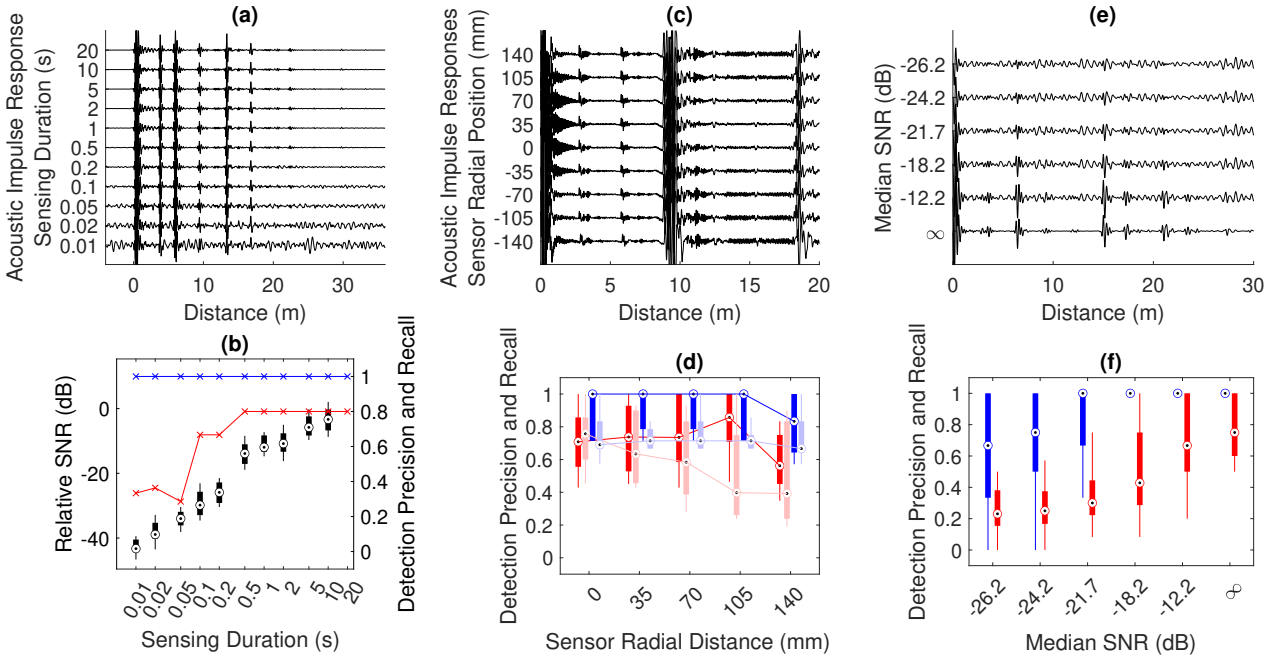


Fig. 7: The effect on acoustic echo sensing of three parameters: (a, b) the sensing duration, (c, d) the sensor position in the pipe cross-section, (e, f) the background noise. (a, c, e) show the measured acoustic impulse response. (b, d, f) show the precision and recall of the feature detection in red and blue respectively using bar charts. In (b), the black graph shows the relative signal-to-noise ratio (SNR). In (d), the light colour graphs relate to the data when higher-frequency acoustic echoes are included in the impulse response, while the dark colour graphs relate to the data when only the low-frequency acoustic echoes are included. In all cases, the SNR is defined as $10\log_{10}(P_{signal}/P_{noise})$, where P is the signal power.

Figure 7(b) shows a measurement of relative *signal to noise ratio* (SNR) between each duration and the longest duration. The SNR decreases with decreasing sensing duration, and the precision of the feature detection drops considerably at a sensing duration below 0.5 s, while the recall remains at 1.

The total sensing duration is made up of two components: the transmission duration varied here and the recording duration. The relatively slow speed of sound means that the recording duration must be sufficiently long to allow perception of distant features. For example, sound propagating at 343 metres per second to a feature 17 metres away will take around 0.1 seconds to return, which puts a lower limit on the recording duration which allows this perception. This physical limitation on total sensing duration therefore imposes diminishing gains on reducing the transmission duration below the recording duration required for a desired perception range.

There is a measurable sensitivity of this sensing method to sensing duration. However, a robot in this application requires perception over the range of tens of metres common in buried pipe environments, and takes hundreds of seconds to traverse a pipe, and there is therefore low incentive to use a short sensing duration where the sensing method is less effective.

Figure 7(c) illustrates the acoustic echo measurements made with the sensor at different positions across the diameter of the pipe. Complex acoustic propagation modes exist in a pipe above a threshold frequency [26], which are to be avoided when a simple impulse response is desired. The excitement of these modes depends on the position of the acoustic sensor relative to the pipe axis, which may be difficult to control in practical application of the sensing method due to the size of the robot relative to the pipe. The sensitivity of the feature detection is measured here.

Figure 7(d) shows a reduction in feature detection precision when the sensor is positioned farther from the pipe axis, but shows that precision remains constant for a position up to 70% of the pipe radius when low-pass filtering is applied. This filtering enables the sensor to be placed anywhere in the cross-section, allowing flexible application to a robot.

Figure 7(e) illustrates acoustic echo measurements made in the presence of different magnitudes of background noise. The echo impulses observable with no background noise diminish relative to the increasing background noise.

Figure 7(f) shows the feature detection performance diminish for increasing background noise amplitude, both in terms of precision and recall. In practice, the magnitude of the background noise cannot be controlled, illustrating the need for a robust localization algorithm. Improvements to sensing might increase the amplitude or duration of the transmitted acoustic signal in the presence of higher magnitude background noise or avoid using acoustic sensing when there is substantial background noise.

B. Localization with Acoustic Echoes

This section demonstrates robot localization using acoustic echo sensing measurements. A robotic platform shown in

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2024.3421236

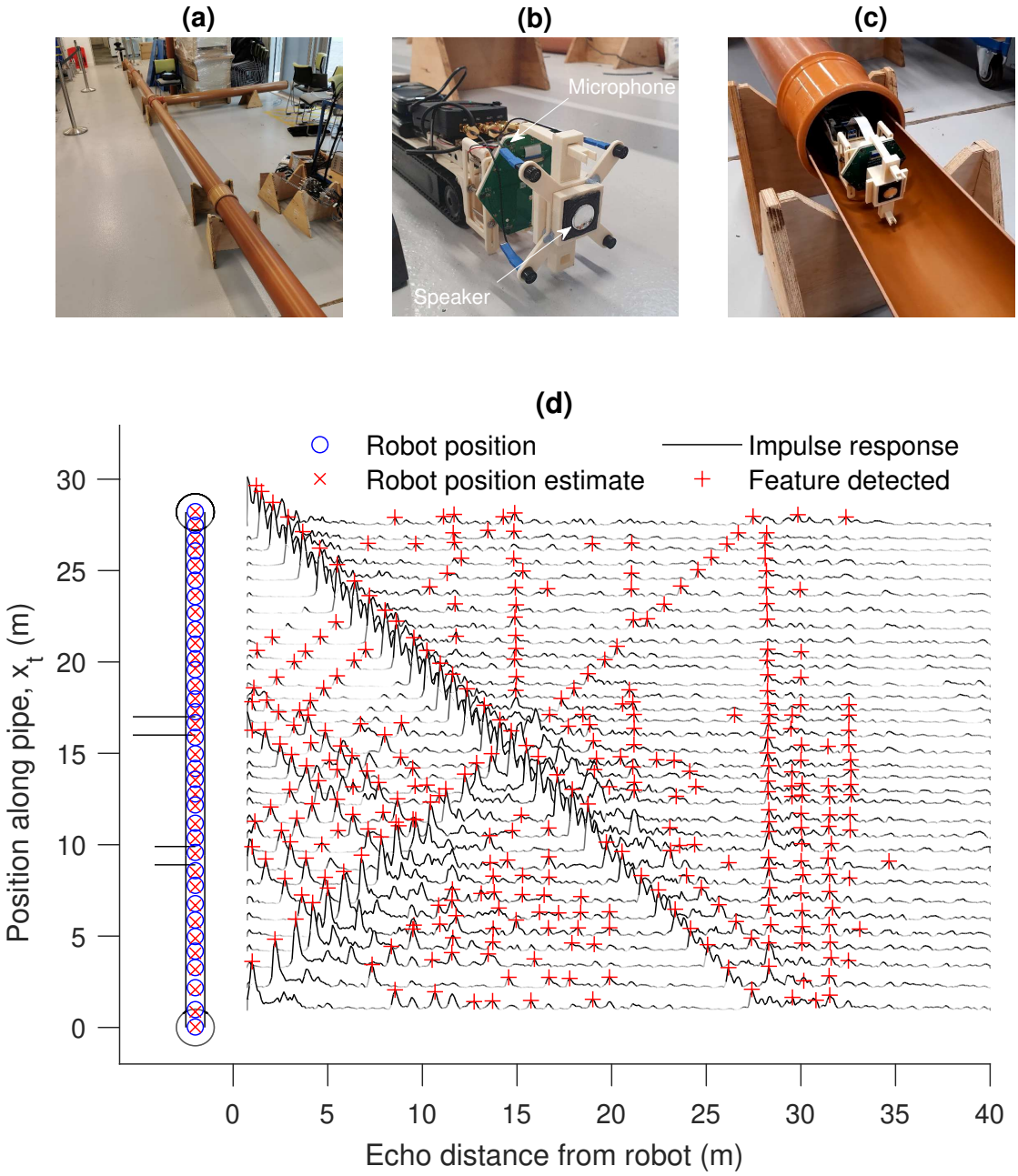


Fig. 8: Illustrative results from experimental data. (a) A photograph of example experimental apparatus, showing a main pipe with a lateral connection. (b) and (c) Photographs of the robot platform used to record the experimental data. (d) Experimental data recorded by the robot in a pipe. Right: The impulse responses and detected features are shown, offset in the y-axis to the position that they were recorded at. Left: A sketch of the pipe with manholes and lateral connections, and the robot trajectory, similar to the labelled sketch in Figure 2.

Figures 8(b) and (c) was moved along pipes like that shown in Figure 8(a). Starting at one end of the pipe, the following procedure was enacted. The robot was moved forwards approximately 1 metre, after which an odometry measurement was made relative to the previous position based on control input, and the ground truth position was measured using an attached tether. The robot then made an acoustic echo measurement using a Visaton 2242 loudspeaker and a MSM321A3729H9CP MEMSensing Microsystems Co., Ltd. microphone. This process was repeated while the robot traversed the length of the pipe. The two sets of experimental data presented in this paper are: Set 1, where the pipe was 27.6 m long, with a 5.5 m long lateral connection 21.6 m from one end; Set 2, where the pipe was 28.2 m long, with 2.15 m and 3.3 m long lateral connections 9.4 m and 16.5 m from one end.

Figure 2(a) and (b) illustrate data similar to that in Set 1, although the lateral connection is a different length. Figure 8(c) shows the measured acoustic impulse responses in Set 2, along with the detected features which are input to the localization algorithm. The localization estimate made using the echoes is

shown on a sketch of the pipe. There is a close match between the estimated and true robot positions.

C. Localization Algorithm Measurement

To quantify the robustness of the localization algorithm to the uncertainty in measurements and motion expected in practice, the experimental data can be extended by adjusting variables in simulation.

Experimental data Set 1 and Set 2 contain odometry, acoustic measurements, and the ground truth measurement of position for a trajectory through two different pipe environments. Based on these data, for a given trajectory $\mathbf{x}_{0:T}$ and environment, sequences of measurements $\boldsymbol{\xi}_t$ and $\mathbf{u}_t$ can be simulated. The uncertainty in the simulated robot motion Σ_u and uncertainty in the measured distance for each measurement Σ_z can be varied. To model error in feature detection, a random number N_{fp} of false positive measurements can be added at each time step, and a random number N_{fn} of true measurements can be removed, creating false negative measurements. These numbers are drawn randomly from 0 to θ_{fp} and θ_{fn} for false positives and false negatives respectively, so $N_{fp} \sim \mathcal{U}(0, \theta_{fp})$ and $N_{fn} \sim \mathcal{U}(0, \theta_{fn})$. From these parameters, estimates of mean precision, $\tilde{\theta}_p$, and mean recall, $\tilde{\theta}_r$, are

$$\tilde{\theta}_p = \frac{N_{tp}}{N_{tp} + \frac{1}{2}\theta_{fp}} \tag{21}$$

$$\tilde{\theta}_r = \frac{N_{tp} - \frac{1}{2}\theta_{fn}}{N_{tp}} \tag{22}$$

where N_{tp} is the number of true positive measurements.

The algorithm *error rate* was measured, which is the proportion of time for which the trajectory estimate error was above 0.5 m, a threshold consistent with the precision expected in industry [48].

D. Localization Algorithm Comparison

1) Comparison of acoustic echo localization methods: The performance of the benchmark KF, simple PGO1st, and advanced PGO2nd algorithms measured as described in Section IV-C is compared here. The algorithms incorporating acoustic echo sensing are compared to an estimate made using only odometry. The results are shown over a variation in the value of Σ_u, with constant values of $\Sigma_z = 0.06$ m, $\theta_{fp} = 1$ ($\tilde{\theta}_p \approx 0.97$), and $\theta_{fn} = 1$ ($\tilde{\theta}_r \approx 0.97$). The acoustic measurements are based on experimental data Set 1, described in Section IV-B.

As described in Section III-D, the PGO1st algorithm is designed with the assumptions that *second-order* measurements are removed and that lateral connections are not passed by the robot, and the KF algorithm is specialised for echoes in a single pipe. Figure 9 shows the comparison where: (a) the assumptions are valid, (b) second-order measurements are not removed from the measurement sets, (c) the robot passes lateral connections along the pipe, and (d) both assumptions are not valid, which represents the most challenging and realistic case. Tables II and III show the median error rate and median estimate error respectively for each algorithm in case (d).

Uncertainty Σ_u (m)	0.50	1.0	1.5	2.0
Method	Median error rate			
KF	0.11	0.29	0.44	0.61
PGO1st	0.0	0.07	0.82	0.84
PGO2nd	0.0	0.0	0.0	0.036

TABLE II: Comparison between localization algorithms using the measured estimate error rate.

Uncertainty Σ_u (m)	0.50	1.0	1.5	2.0
Method	Median estimate error (m)			
KF	0.25	0.93	1.7	3.5
PGO1st	0.087	0.22	1.8	3.0
PGO2nd	0.061	0.065	0.086	0.22

TABLE III: Comparison between localization algorithms using the measured estimate error.

Except in the case where only odometry is used, a similar trend is seen for both the error and the error rate. The estimate error and error rate of the KF algorithm is generally largest, except compared to the use of only odometry. The error rate of the PGO1st algorithm becomes more substantial with increasingly realistic experimental data for which the design assumptions are not valid, while the error rate of the PGO2nd algorithm remains low in the most realistic case. These results show the need for the additional complexity when using the novel approach proposed here.

2) Comparison to alternative sensing in the literature: The average localization estimate error using different sensing approaches in the literature is compared to the estimate error found here using acoustic echoes in Table IV.

As shown in Figure 9, the localization estimate error depends on the input odometry measurement uncertainty. This is reported at various values in the literature; measurements of 1.6% [22] and 5.3% [20] are found for the estimate error per unit of distance travelled. These would correspond to error of 0.46 m and 1.5 m in the 27.6 m long pipe used in this experiment. From Figure 9, these would correspond to a odometry measurement uncertainty of around 0.25 m and 0.75 m respectively, when using the odometry model used in this experiment. For a fair comparison between the acoustic echo localization approach and the other sensing approaches in the literature, the acoustic echo localization accuracy will be compared at a larger value of odometry uncertainty, of 1.0 m.

It is challenging to make a direct comparison between the dissimilar sensing approaches. Vision-based SLAM [8] and pipe joint detection [9] and hydrophone-based spatial signal sensing [18] are expected to increase in estimate error with increasing distance along a pipe. The same is true to an extent for calibrated radio beacon sensing [22] which is incorporated into the localization estimate using the previously developed spatial signal approach [18], [19]. In contrast, localization based on sensing from periodic radio signals [20], tethered acoustic sensing [21], and the acoustic echo sensing proposed in this work are not expected to increase in estimate error with

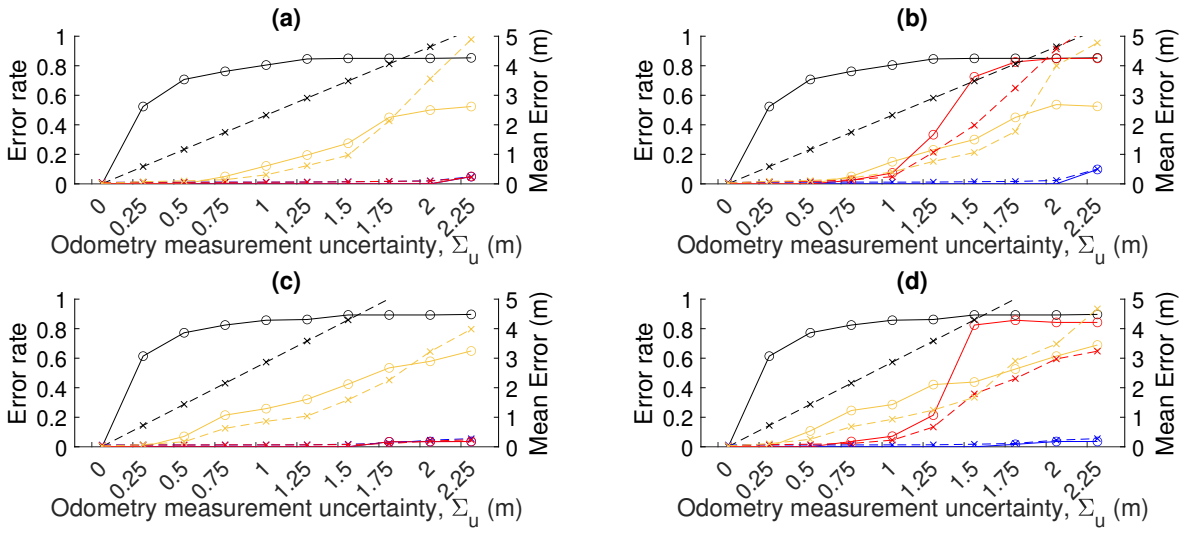


Fig. 9: A comparison of the use of only odometry (black) and the benchmark KF (yellow), PGO1st (red), and PGO2nd (blue) algorithms, in the case: (a) without second-order measurements or lateral connections; (b) with second-order measurements; (c) with lateral connections; (d) with both second-order measurements and lateral connections. In all cases, the median estimate error and median error rate over 50 trajectories are given, with variation in the odometry uncertainty.

increasing distance. Therefore, the best comparison measure could be either the average estimate error per unit pipe length or the absolute average estimate error. Both measures are given in Table IV.

At the chosen level of odometry uncertainty of 1.0 m, the PGO2nd algorithm gives a median estimate error of 0.065 m, which is 0.24% of the distance travelled, with a standard deviation of 0.061 m, or 0.23% of the distance travelled. Equivalently, the PGO1st algorithm gives a median error of 0.22 m, or 0.80% of distance travelled, and the KF algorithm gives a median error of 0.93 m, or 3.4% of distance travelled. This level of odometry uncertainty is similar to that found in the literature in dry pipes with large, stable robots. In application to more challenging environments, with uneven, wet surfaces, and with the flow of fluid, this odometry uncertainty is expected to be larger. Figure 9 suggests that in this case, the error from the PGO1st and KF algorithms would be substantially larger, while the error of the PGO2nd would only slightly increase.

The accuracy of the acoustic echo sensing approach developed in this paper is surpassed by the use of the spatial signal [18], [19] from calibrated radio beacons [22], which achieves a relative estimate error of 0.047%, However, the acoustic echo localization approach gives almost ten times less error than those sensing approaches which do not require additional external hardware to be placed around the environment. This accuracy is sufficient for the fault localization application with a target precision of 0.5 metres.

While other sensing approaches in the literature produce less accurate localization estimates in this case, they have characteristics which may have utility in the application. For example, vision based SLAM can be used to create a detailed map of the environment, and the use of vision sensing may be required for pipe condition assessment in this application. Periodic radio signal sensing may be used up to a range of several kilometres [20] which may be necessary in transportation pipeline applications. Future work should consider the use of a combination of sensing approaches, which may improve localization results and provide an extended sensing utility.

Sensing approach	Test pipe length (m)	Average estimate error (m) (per unit pipe length (%))
Sensing with external hardware		
Radio Beacon [22] using Spatial Signal [18], [19]	50	0.023 (0.047)
Radio Periodic Signal [20]	65	0.3 (0.6)
Tethered Acoustic [21]	6	0.21 (3.5)
Sensing without external hardware		
Vision SLAM [8]	20.3	0.44 (2.17)
Vision Pipe Joint [9]	90	2.04 (2.27)
Hydrophone Spatial Signal [18]	5	0.175 (3.5)
Acoustic Echo	27.6	0.065 (0.24)

TABLE IV: Comparison to sensing approaches in the literature.

E. Localization Algorithm Evaluation

The PGO2nd algorithm described in Section III-E is evaluated over varying uncertainty in the robot's motion and sensing, as described in Section IV-C. The values of Σ_u, Σ_z, $\tilde{\theta}_p$, and $\tilde{\theta}_r$ are varied, with default values of $\Sigma_u = 0.5$ m, $\Sigma_z = 0.05$ m, $\theta_{fp} = 4$ ($\tilde{\theta}_p \approx 0.9$), and $\theta_{fn} = 1$ ($\tilde{\theta}_r \approx 0.97$). The acoustic measurements are based on experimental data Set 2 described in Section IV-B.

For comparison, two odometry-based methods were used. The first uses only odometry, and the second also uses prior knowledge of the position of each end of the pipe. The

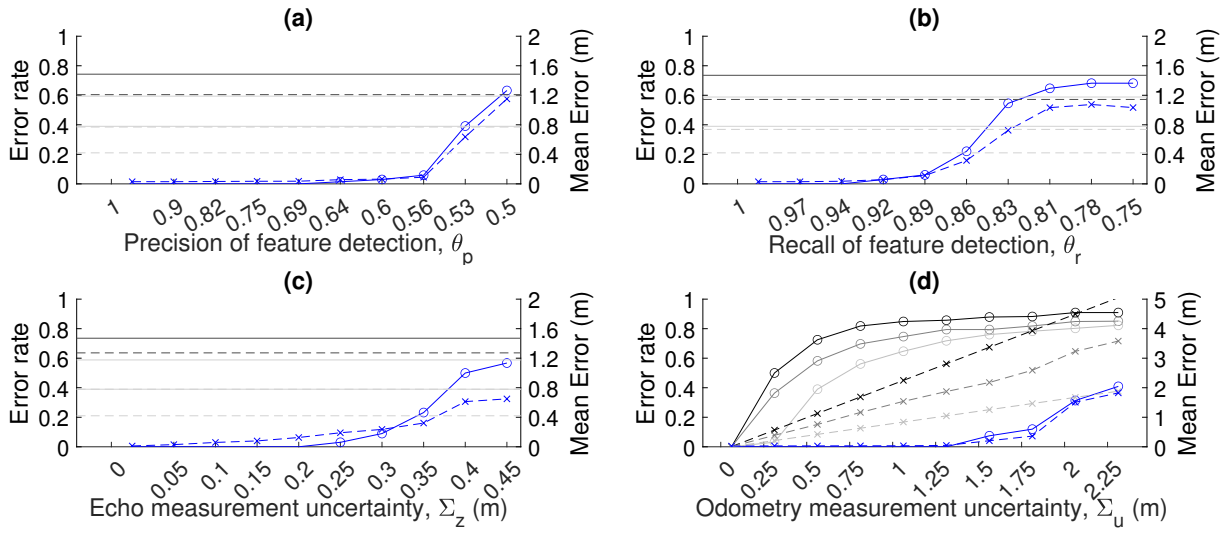


Fig. 10: Analysis of the sensitivity of localization algorithms to uncertainty in the input measurements. In all cases, the median estimate error and error rate over 50 tests are shown. The black data is the result using only odometry measurements, not echo measurements. The dark grey data is the result using prior knowledge of the position of the ends of the pipe with odometry. The light grey data is the result using loop-closure measurements, prior knowledge of the position of the ends of the pipe, and odometry. The blue data is the result using pose-graph optimization to incorporate the echo measurements, along with odometry and prior knowledge of the position of the ends of the pipe.

manholes at each end of the pipe can be observed and located from above ground, and can be detected reliably using vision sensing [9]. A *loop-closure*-based method emulating the effect of using vision to recognise previously observed locations was also compared; the effectiveness of this in the pipe environment has been measured previously [49]. It was assumed that this recognition is perfect, although in practice there will be uncertainty. This vision-based observation was not used in the acoustic localization algorithm. Figure 10(d) shows the sensitivity of the odometry-based methods to odometry measurement noise, and shows that error rate with *loop-closure* increases above 0.25 m of odometry uncertainty despite the ideal location recognition.

The odometry measurement noise in practice will depend on the odometry sensing approach and the environment. As described in section IV-D, measurements of error per unit of distance travelled are 1.6% [22] and 5.3% [20] in the literature. These would correspond to error of 0.47 m and 1.5 m in the 28.2 m long pipe used in this experiment. From Figure 10(d), these correspond to an odometry uncertainty of around 0.25 and 0.75, and an error rate of around 0.5 and 0.8 respectively, when applied to the odometry model used in this experiment. Odometry measurement using vision has been shown in literature to result in an average localization error relative to the total distance travelled of 2.17% [8] and of 2.27% using an alternative method where the position of each of the pipe is also known [9]. These would correspond to around 0.6 metre error in the 28.2 m long pipe, and from Figure 10(d) this corresponds to an odometry uncertainty of 0.25 m and 0.5 m respectively, and therefore would give an error rate of 0.5 and 0.4.

Figure 10(a) and (b) show that the PGO2nd algorithm has a low sensitivity to feature detection precision, and some sensitivity to feature detection recall, as the estimate error and error rate increase substantially only when precision and recall (as calculated in equations 21 and 22) are reduced to 0.53 and 0.86 respectively. In Figure 7, a precision of 0.53 is only found for sensing duration of less than 0.1 s, a sensor position at the very edge of the pipe cross-section, and a signal-to-noise ratio of -18 dB. A recall of 0.86 is only found for a sensor position at the edge of the pipe cross section, and a signal-to-noise ratio less than -22 dB, and is not found for even the lowest measured sensing duration. The localization result therefore shows low sensitivity to the sensing duration and sensor position, and some sensitivity to signal-to-noise ratio. A feature detection function would have some balance between probability of false positive and false negative detection. To improve the subsequent localization estimate, such a function should aim to increase recall at the cost of decreased precision.

Figure 10(c) shows the algorithm's error rate begins to increase when the measurement uncertainty becomes greater than 0.2 m. The median error rate is zero for measurement uncertainty up to 2 times the uncertainty in experimental measurement of 0.09 m reported in previous work [25]. While in practice, measurements may be less precise than in a laboratory experiment, there are advances in sensing and signal processing for this application [26] that could improve the measurement precision. Figure 10(d) shows that the median error rate remains zero for odometry uncertainty up to 1 m in magnitude, which is approximately equal to the robot's travelled distance at each time step. This is 1.3 to 4 times the equivalent amount found from reported values in the literature, as described earlier in this section, and is 4 times the amount for which the median error rate of the loop-closure method remains close to zero.

V. Conclusion

This paper investigated the use of acoustic echo sensing for robot localization in pipes. Experimental sensor data has shown that the sensing approach can measure the distance from the robot to distant reflective features, such as pipe joints and manholes, achieving useful distance measurements without the need for hardware external to the robot such as a tether cable or radio beacon. The acoustic sensing performance has been measured in terms of precision and recall, and evaluated over three sensor parameters: sensing duration, sensor position, and sensor signal-to-noise ratio. It is seen that the sensing performance remains constant for a sensing duration greater than 0.5 seconds and a sensor position within the middle 70% of the pipe cross section, while the signal-to-noise ratio has a greater effect on the sensing output, motivating the development of a robust localization algorithm.

A novel pose-graph optimization algorithm for acoustic echo sensing for localization in pipes was developed, with bespoke aspects shown to be necessary to incorporate acoustic echo information from the highly reverberant environment. The algorithm gives a relative estimate error of 0.23% of distance travelled, which surpasses that of alternative sensing approaches except those which require external hardware. This algorithm has been evaluated over a range of values of different sources of uncertainty in robot motion and sensing. It gives a low estimate error rate for values of measurement uncertainty beyond those expected in practice based on experimental measurements and results from the literature. Results show a median error rate of zero for echo measurement uncertainty 2 times that measured experimentally, and an error rate of zero for motion measurement uncertainty 1.3 to 4 times larger than the equivalent uncertainty reported in the literature. The median error rate remains zero for a sensing duration as low as 0.1 seconds, for a sensor position up to 70% of the pipe radius away from the ideal position at the pipe cross-section centre, and for a signal-to-noise ratio as low as -12 dB.

Overall, the experimental evaluation of the sensing system and the evaluation of the localization system over a variation in sensing system performance gives a good understanding of the effectiveness of acoustic echo sensing for robot localization in pipes. The system overall is shown to be accurate and robust, and the limits of this robustness have been found, which can be used for further development and deployment of this approach.

Acknowledgment

This work is supported by UK Engineering and Physical Sciences Research Council (EPSRC) Programme Grant No. EP/S016813/1. The authors thank Gavin Sailor for help with the experimental hardware. For the purpose of open access, the authors have applied a creative commons attribution (CC BY) licence to any author accepted manuscript version arising.

References

[1] R. Sterling, J. Simicevic, E. Allouche, W. Condit, and L. Wang, "Epa, state of technology for rehabilitation of wastewater collection systems," *U.S. Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Edison, NJ*, 2010.

[2] GOV.UK, "Water and treated water," 2015. [Online]. Available: https://www.gov.uk/government/publications/water-and-treated-water/water-and-treated-water

[3] The Consumer Council for Water, "Water, water everywhere. Delivering resilient water and waste water services everywhere (2017-2018)," 2017. [Online]. Available: https://www.ccwater.org.uk/research/water-water-everywhere-delivering-resilient-water-and-wastewater-services-2017-18/

[4] H. Jang, T. Yu, K. Ye, C. Lee, Y. Soo, K. Jooyong, K. Hae, Y. Lee, and H. Ryeol, "A Review : Technological Trends and Development Direction of Pipeline Robot Systems," *Journal of Intelligent & Robotic Systems*, pp. 1–20, 2022. [Online]. Available: https://doi.org/10.1007/s10846-022-01669-2

[5] J. M. Aitken, M. H. Evans, R. Worley, S. Edwards, R. Zhang, T. Dodd, L. Mihaylova, and S. R. Anderson, "Simultaneous Localization and Mapping for Inspection Robots in Water and Sewer Pipe Networks: A Review," *IEEE Access*, vol. 9, pp. 140 173–140 198, 2021.

[6] M.-G. Li, H. Zhu, S.-Z. You, and C.-Q. Tang, "Uwb-based localization system aided with inertial sensor for underground coal mine applications," *IEEE Sensors Journal*, vol. 20, no. 12, pp. 6652–6669, 2020.

[7] P. Hansen, H. Alismail, P. Rander, and B. Browning, "Visual mapping for natural gas pipe inspection," *International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 532–538, 2015.

[8] R. Zhang, R. Worley, S. Edwards, J. Aitken, S. R. Anderson, and L. Mihaylova, "Visual simultaneous localization and mapping for sewer pipe networks leveraging cylindrical regularity," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3406–3413, 2023.

[9] S. Edwards, R. Zhang, R. Worley, L. Mihaylova, J. Aitken, and S. R. Anderson, "A robust method for approximate visual robot localization in feature-sparse sewer pipes," *Frontiers in Robotics and AI*, vol. 10, p. 1150508, 2023.

[10] D. H. Lee, H. Moon, J. C. Koo, and H. R. Choi, "Map building method for urban gas pipelines based on landmark detection," *International Journal of Control, Automation and Systems*, vol. 11, no. 1, pp. 127–135, 2013.

[11] A. Kakogawa, Y. Komurasaki, and S. Ma, "Anisotropic shadow-based operation assistant for a pipeline-inspection robot using a single illuminator and camera," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2017-Septe, pp. 1305–1310, 2017.

[12] W. Zhao, M. Kamezaki, K. Yoshida, M. Konno, A. Onuki, and S. Sugano, "Modeling and simulation of FLC-based navigation algorithm for small gas pipeline inspection robot," *IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM*, vol. 2018-July, pp. 912–917, 2018.

[13] A. Kakogawa, T. Yamagami, Y. Tian, and S. Ma, "Recognition of pathway directions based on nonlinear least squares method," *2015 IEEE International Conference on Robotics and Biomimetics, IEEE-ROBIO 2015*, pp. 1596–1601, 2015.

[14] J. T. Thielemann, G. M. Breivik, and A. Berge, "Pipeline landmark detection for autonomous robot navigation using time-of-flight imagery," *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops*, 2008.

[15] F. Kirchner and J. Hertzberg, "A prototype study of an autonomous robot platform for sewerage system maintenance," *Autonomous Robots*, vol. 4, no. 4, pp. 319–331, 1997.

[16] H. Sahli and N. El-Sheimy, "A novel method to enhance pipeline trajectory determination using pipeline junctions," *Sensors (Switzerland)*, vol. 16, no. 4, pp. 1–17, 2016.

[17] X. Dai, W. Song, Y. Wang, Y. Xu, Y. Lou, and W. Tang, "Lidar–inertial integration for rail vehicle localization and mapping in tunnels," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 17 426–17 438, 2023.

[18] R. Worley, K. Ma, G. Sailor, M. M. Schirru, R. Dwyer-joyce, J. Boxall, T. Dodd, R. Collins, and S. Anderson, "Robot localization in water pipes using acoustic signals and pose graph optimization," *Sensors (Switzerland)*, pp. 1–23, 2020.

[19] B. Park and H. Myung, "Resilient Underground Localization Using Magnetic Field Anomalies for Drilling Environment," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1377–1387, 2018.

[20] C. Rizzo, T. Seco, J. Espelosín, F. Lera, and J. L. Villarroel, "An alternative approach for robot localization inside pipes using RF spatial fadings," *Robotics and Autonomous Systems*, vol. 136, p. 103702, 2021. [Online]. Available: https://doi.org/10.1016/j.robot.2020.103702

[21] Y. Bando, H. Suhara, M. Tanaka, T. Kamegawa, K. Itoyama, K. Yoshii, F. Matsuno, and H. G. Okuno, "Sound-based online localization for an in-pipe snake robot," *SSRR 2016 - International Symposium on Safety, Security and Rescue Robotics*, pp. 207–213, 2016.

[22] A. Gunatilake, S. Kodagoda, and K. Thiyagarajan, "Battery-Free UHF-RFID Sensors-Based SLAM for In-Pipe Robot Perception," *IEEE Sensors Journal*, vol. 22, no. 20, pp. 20 019–20 026, 2022.
[23] S. Kumar and S. K. Das, "Target detection and localization methods using compartmental model for internet of things," *IEEE Transactions on Mobile Computing*, vol. 19, no. 9, pp. 2234–2249, 2020.
[24] J. Sun, W. Sun, J. Zheng, Z. Chen, C. Tang, and X. Zhang, "A novel uwb/imu/odometer-based robot localization system in los/nlos mixed environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–13, 2024.
[25] R. Worley, Y. Yu, and S. Anderson, "Acoustic echo-localization for pipe inspection robots," *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 2–7, 2020.
[26] Y. Yu, R. Worley, S. Anderson, and K. V. Horoshenkov, "Microphone array analysis for simultaneous condition detection, localization and classification in a pipe," *The Journal of the Acoustical Society of America*, vol. 153, pp. 367–383, 2023.
[27] F. Dümbgen, A. Hoffet, M. Kolundžija, A. Scholefield, and M. Vetterli, "Blind as a bat: Audible echolocation on small robots," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1271–1278, 2023.
[28] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *Journal of Sound and Vibration*, vol. 102, no. 2, pp. 217–228, 1985.
[29] C. Evers, S. Member, P. A. Naylor, and S. Member, "Acoustic SLAM," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 9, pp. 1484–1498, 2018.
[30] M. Krekovic, I. Dokmanic, and M. Vetterli, "EchoSLAM : Simultaneous Localization and Mapping with Acousitc Echoes," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 11–15, 2016.
[31] U. Saqib and J. R. Jensen, "A framework for spatial map generation using acoustic echoes for robotic platforms," *Robotics and Autonomous Systems*, vol. 150, p. 104009, 2022.
[32] M. Feder and E. Weinstein, "Parameter Estimation of Superimposed Signals Using the EM Algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.
[33] U. Saqib, M. Græsbøll Christensen, and J. Rindom Jensen, "Robust Acoustic Reflector Localization for Robots," Feb. 2021, preprint. [Online]. Available: https://hal.archives-ouvertes.fr/hal-03154438
[34] U. Saqib, S. Gannot, and J. R. Jensen, "Estimation of acoustic echoes using expectation-maximization methods," *Eurasip Journal on Audio, Speech, and Music Processing*, vol. 2020, no. 1, 2020.
[35] S. B. Beck, M. D. Curren, N. D. Sims, and R. Stanway, "Pipeline network features and leak detection by cross-correlation analysis of reflected waves," *Journal of Hydraulic Engineering*, vol. 131, no. 8, pp. 715–723, 2005.
[36] C. Xu, S. Du, P. Gong, Z. Li, G. Chen, and G. Song, "An improved method for pipeline leakage localization with a single sensor based on modal acoustic emission and empirical mode decomposition with hilbert transform," *IEEE Sensors Journal*, vol. 20, no. 10, pp. 5480–5491, 2020.
[37] Y. Zhu, X. Lang, L. Zhang, and Z. Cai, "Leak localization method of jet fuel pipeline based on second-generation wavelet transform and short-time energy time delay estimation," *IEEE Sensors Journal*, vol. 23, no. 3, pp. 2823–2832, 2023.
[38] X. Wang, K. M. Lewis, K. A. Papadopoulou, B. Lennox, and J. T. Turner, "Detection of hydrate and other blockages in gas pipelines using acoustic reflectometry," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 226, no. 7, pp. 1800–1810, 2012.
[39] Z. Liu, R. Chen, F. Ye, G. Guo, Z. Li, and L. Qian, "Improved toa estimation method for acoustic ranging in a reverberant environment," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 4844–4852, 2022.
[40] M. T. B. Ali, K. V. Horoshenkov, and S. J. Tait, "Rapid Detection of Sewer Defects and Blockages Using Acoustic Based Instrumentation," *Water Science & Technology*, vol. 64, no. 8, pp. 1700–1707, 2010.
[41] L. Tao, "Monitoring Gas Distribution Pipelines," Ph.D. dissertation, The University of Manchester, 2017.
[42] Y. Yu, K. V. Horoshenkov, and S. Tait, "Microphone array analysis of the first non-axisymmetric mode for the detection of pipe conditions," *The Journal of the Acoustical Society of America*, vol. 155, no. 1, pp. 575–587, 01 2024. [Online]. Available: https://doi.org/10.1121/10.0024360
[43] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2006.
[44] R. Worley and S. Anderson, "Robust Efficient Localization of Robots in Pipe Networks using a Particle Filter for Hybrid Metric-Topological Space," *2021 European Conference on Mobile Robots (ECMR)*, 2021.
[45] D. Murray, S. Panguluri, G. Skipper, and S. Donovan, "Demonstration of Innovative Sewer System Inspection Technology SewerBatt ™," Tech. Rep. January, 2014.
[46] F. Lu and E. Milios, "Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 18, no. 3, pp. 249–275, 1997.
[47] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, "Graduated Non-Convexity for Robust Spatial Perception: From Non-Minimal Solvers to Global Outlier Rejection," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1127–1134, 2020.
[48] Institution of Civil Engineers, "Specification for underground utility detection, verification and location (PAS128:2014)," Tech. Rep., 2014.
[49] M. H. Evans, J. M. Aitken, and S. R. Anderson, "Assessing the feasibility of monocular visual simultaneous localization and mapping for live sewer pipes: a field robotics study," pp. 1073–1078, 2021.