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 a number of challenges, including the unavailability of 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 signals sent by robots propagate long distances and around corners in the pipe environment, and can be used to estimate the distance to reflective features. It is shown that the reverberant environment causes a number of echoes which contain useful information. A novel algorithm for using this information in pose-graph optimization is proposed, and is shown to be necessary in comparison to a naive approach. The algorithm is demonstrated to be accurate and robust by measuring the error rate, which is the proportion of estimate error that was greater than a target threshold of 0.5 metres. The median error rate was 0 for measurement uncertainty 2.5 times larger than that found experimentally, and for motion uncertainty 4 times that which could be tolerated by an idealised loop-closure approach used for comparison. This work is the foundation for new data fusion methods for robot localization in buried pipes.**

***Index Terms*—Localization, mapping, simultaneous localization and mapping (SLAM), acoustic sensing, robots, pipe networks**

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. There is a total of 250 billion pounds of investment in water infrastructure in the UK, with around 8 billion pounds spent each year [3]. Faults in wastewater pipes can cause pollution and danger to public health, with around 50 serious pollution incidents per year in the UK [4]. 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. Its output is useful for large-scale navigation, small-scale control, for mapping an unexplored environment, and for locating the robot in an environment. Robust, accurate localization is required for the feasibility of robotic inspection, which will improve the efficiency and effectiveness of buried infrastructure monitoring, having public health and economic impacts.

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 is with the Department of Mechanical Engineering, University of Sheffield, Sheffield, UK. e-mail: {yicheng.yu, k.horoshenkov}@sheffield.ac.uk

Robot localization in this application is challenging as the pipe network adds a number of constraints. 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 fundamentally there is a constraint on what kinds of perception can be made at all. GPS for measuring position and a magnetometer for measuring heading angle are unavailable or unreliable underground.

Despite these limitations, common robot sensing has been used in pipes. Vision has been used for odometry and mapping [5] and for detecting features such as corners and junctions using shadows [6], [7], light [8], projected laser arrays [9], and depth images [10]. These features can also be detected with inertial sensing [11], [12] and scanning rangefinders [13]. While effective, cameras and rangefinders in the pipe environment can only observe nearby features so *loop-closure* measurements will be infrequent. Therefore, visual localization in pipes relies on odometry, which suffers from accumulating uncertainty and drift in the position estimate. As well as this, the environment has few distinguishing features, which can cause perceptual aliasing. These typical sensing methods also have drawbacks in the pipe environment compared to a typical application. Cameras require lights, which add to the power consumption, and are sensitive to lighting conditions, scanning rangefinders add to the size, and both methods suffer from a lack of detectable features in a typical pipe.

Sensing methods specialised for buried pipes have also been deployed on robots. Radio waves have been used to produce a recognisable signal which varies spatially along the length of a pipe [14], and radio signals has been used in pipes with *RFID tags* [15]. The time of travel of an acoustic signal has been used on a tethered robot to estimate the distance from a separate transmitter [16]. These methods estimate the robot's position relative to a known location, but require separate hardware for signal transmission or a tether, which limit the robot's mobility and will be avoided in this work. High frequency acoustic sensing has been made using a hydrophone in water-filled pipes to measure a recognisable spatially varying property of the pipe [17], [18]. This is used for localization where there are otherwise no observable features, but is limited to observing only nearby features like vision-based sensing.

There is a research gap, therefore, motivating the development of sensors for robots which can perceive distant features in pipes, increasing the perspective of the robot, but that do not require any hardware that is not on the robot, such as tethers and beacons. This paper proposes the use of acoustic echoes

for robot localization in pipes, which overcomes many of the limitations of other sensing modes such as vision.

Acoustic echoes can be used to measure the absolute distance to distant acoustically reflective features. This will not suffer from accumulated uncertainty and therefore avoids drift in the position estimate. However, while acoustic echo localization can only estimate a one-dimensional robot pose along the pipe, vision-based localization can estimate in a higher number of dimensions. Future work could address the complementary fusion of information from these sensing modes, while this paper aims to first establish the utility of acoustic echoes for robot localization in pipes.

The aim of this paper is to develop a novel acoustic echo localization system for robots in pipes, and in particular to:

1) develop a mathematical model for acoustic perceptions in the pipe environment (Section III-A to Section III-D);
2) develop a novel pose-graph localization algorithm specialized for acoustic echo sensing in the pipe environment (Section III-E);
3) demonstrate that acoustic echoes can be used by a robot to make a useful perception of the environment and investigate the uncertainty in perception over a range of sensing conditions (Section IV-A, IV-B);
4) show the necessity of a novel localization algorithm designed for acoustic sensing in this environment (Section IV-E);
5) evaluate the effectiveness of this localization method in the pipe environment over a range of uncertainty in inputs (Section IV-F).

In our previous work, acoustic echo localization was done in a single pipe using Kalman filtering [19], [20]. This paper takes a different approach to localization based on pose-graph optimization in a more realistic and challenging environment, and more thoroughly investigates and demonstrates the use of this method of sensing using new experimental data.

II. Background

A. Related work in acoustic echo localization

The use of acoustic echoes for robot localization requires an understanding of the acoustic response of the robot's surroundings. Typically this acoustic response is considered as the transfer function between transmitter and receiver. In principle, measurements of distances between features in the environment and the robot can be made using this acoustic response, for the purpose of robot localization. Estimation of the acoustic *room impulse response* (RIR) in different environments with variation in the position of acoustic transmitter and receiver has been studied for decades [21].

Recent work has demonstrated robot localization in a room with multiple sources of sound moving in an unknown way, using measurements of direction of arrival of sounds [22]. In contrast, here both the direction and time of arrival of sounds will be considered, which can be done using echoes rather than separate sound sources. Other recent work has shown the use of a single co-located acoustic source and receiver to estimate the position of a robot in a structured room using the RIR with a filtering approach [23] and an optimization approach [24]. There it was assumed that each echo is clear and unique, that there are no multi-path echoes, and that each detected echo can be associated with a specific wall in the room. These assumptions cannot be made in the pipe environment, where measurement of acoustic echoes is more challenging due to noise from the robot and the environment, multi-path echoes are likely, and data association is not trivial.

The estimation the impulse response in the presence of noise can be improved [25], [26] by using a *maximum likelihood* (ML) (or *expectation maximization* (EM), or *nonlinear least-squares* (NLS)) optimization method [27]. This method has been applied to localization using a single microphone [28] and an array of microphones, and the effects of signal to noise ratio, hardware transfer function [29], correlated and coloured noise, and faulty microphones [25] have been considered. Recent work has used acoustic interference to localize walls close to a small, resource constrained robot [30], showing precise localization at a distance up to around 0.5 metres. While these methods are designed for room environments, the results are useful for developing similar methods in pipes.

However, in the pipe environment, features will be less reflective acoustically, while pipes will act as a waveguide for acoustic waves, increasing the distance that sound travels. Wastewater pipe networks are constructed with pipes up to a typical length of 90 metres between connections [1], so the focus of this work is on detection of features up to a longer distance than in work outside of pipes. Work outside of robotics has used acoustic echoes recorded from pipe networks to detect features up to a range of 10 kilometres in gas pipelines [31]. This detection of features has been done automatically [32], manually [33], or by comparison to signals recorded previously in the same pipe network [34] or simulated signals [35].

While development of acoustic echo sensing for robot localization in pipe networks is informed by a range of existing literature, autonomous localization in pipe networks requires further development of novel algorithmic aspects and evaluation of the sensing and localization approaches.

B. Background on acoustic wave propagation in pipes

The acoustic field in a pipe comprises multiple *modes* of propagation, which exist at different frequencies and travel at different velocities. For a pipe with cylindrical coordinates (ρ, α, x) for radial position, azimuth angle, and axial position, the total acoustic field in the frequency domain [36] is

$$F(\rho, \alpha, x, \omega) = \sum A_{mn} \Gamma_{mn}(\rho, \alpha) e^{i(\omega k - \gamma_{mn} x)} \qquad (1)$$

where ω is the angular frequency, k is the time, m and n are the mode indices for the (m, n) mode, and γ_{mn} is the axial wavenumber. A_{mn} is the mode amplitude, which depends on the position of the source, the strength of the source, and the reflection from features in the environment. $\Gamma_{mn}(\rho, \alpha)$ is the mode shape function, which describes how the pressure varies over the pipe cross-section. This means that the position of the transmitter, receiver, and of the source of reflection in the pipe cross-section will impact the detected acoustic field, as illustrated in Figure 2(a). $e^{i(\omega t - \gamma_{mn} z)}$ describes the wave

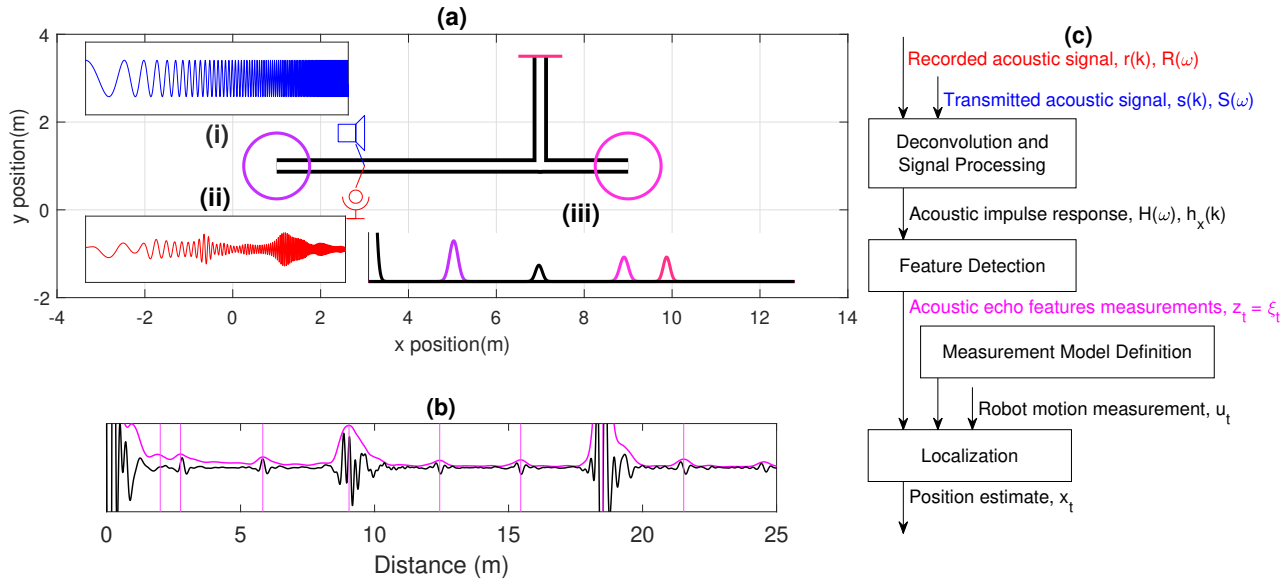

Fig. 1: (a) An illustration of acoustic echo sensing in a pipe network. A robot positioned at (x, y) = (3,1) emits a signal illustrated in (i) and receives a signal illustrated in (ii). The computed impulse response is in (iii), where each impulse and source of each echo are colour-coded. (b) An experimental impulse response in black, with the detected envelope in magenta, and the detected impulses given by the vertical lines. (c) An illustration of the system described in this paper, showing the link between the *signal processing*, *feature detection*, *measurement modelling*, and *localization*.

propagation along the pipe axis, where different modes have frequency dependant propagation velocity [36].

The presence of multiple modes of propagation gives a complex pipe impulse response. This is because each mode has a different propagation velocity, and higher-order modes have frequency dependant velocities, which spreads the signal from a source in the pipe environment across the time domain. Therefore, only the *plane-wave* mode is desired here. This is where the acoustic wave propagates directly along the pipe axis, so the field is equal across the pipe cross-section which produces a simple impulse response. The plane-wave mode is the only mode which exists at low frequencies, below a threshold determined by the pipe diameter.

III. Methods

Robot localization is performed by estimating the acoustic impulse response of the pipe environment using deconvolution, detecting features in the pipe environment from the response, modelling the uncertainty in the feature measurement, and estimating the robot position using optimization methods, as illustrated in Figure 1(c). This section describes the methods which form the overall localization algorithm.

A. Problem Definition

Robot localization is the problem of estimating a robot's state with respect to the environment, and the solution is typically based on a probabilistic approach [37]. Here, a robot moves to state $\mathbf{x}_t$ from state $\mathbf{x}_{t-1}$, modelled by the probability $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t)$, where $\mathbf{u}_t$ is measured by odometry. The robot moves from one end of a pipe to the other over a series of time steps from $t = 0$ to $t = T$.

Here, it is assumed that the pipe is part of a wastewater pipe network. Therefore, at each end of the pipe there will be a *manhole*, and along the main pipe there will be an arbitrary number of *lateral connections*, which are smaller pipes [1]. This is illustrated in Figure 1(a). In practice, pipes vary in length but are an average of 90 metres long [1]. It is assumed that the localization algorithm has prior knowledge of the position of each end of the pipe, as in practice these points will be at manholes which are observable from above ground. However, there is no knowledge of the position of any connections to other pipes, as these cannot be seen from above ground and are often not recorded in asset data [1].

A continuous-discrete state definition [38] is used, suitable for the pipe network environment [39], defined as

$$\mathbf{x}_t = \left(q_t, a_t, x_t\right) \tag{2}$$

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 main pipe, which is the variable of interest for localization. A robot makes observations $\mathbf{z}_t$ from the state, modelled by the probability $p(\mathbf{z}_t|\mathbf{x}_t)$. The measurements $\mathbf{z}_t$ and $\mathbf{u}_t$ always have uncertainty Σ_t^z and Σ_t^u, which motivates the problem and the probabilistic approach.

The problem addressed here is the estimation of posterior probability distribution $p(\mathbf{x}_{0:T}|\mathbf{u}_{1:T}, \mathbf{z}_{1:T}, \mathbf{x}_0, \mathbf{x}_T)$ of the trajectory of the robot from odometry measurements $\mathbf{u}_t$, measurements derived from acoustic echo signals $\mathbf{z}_t$, and prior knowledge of the position of the ends of the pipe $\mathbf{x}_0$ and $\mathbf{x}_T$. The following sections define how the measurements $\mathbf{z}_t$ are extracted from acoustic echo sensing information, and the optimization algorithms for estimating the posterior.

The manholes at each end of the pipe and each end of the lateral connections will all cause acoustic echoes. A map describing the position of these reflective features can be written for pipe q as

$$\mathcal{M}^q = \left\{\mathcal{M}_c^q, \mathcal{M}_l^q\right\} \tag{3}$$

where $\mathcal{M}_l^q$ and $\mathcal{M}_c^q$ describe the length and position of each connection or manhole along the pipe.

For example, a map could be given by

$$\mathcal{M}^q = \left\{\mathcal{M}_c^q = \left\{0, 7, 15, 25\right\}, \mathcal{M}_l^q = \left\{0, 2, 2, 0\right\}\right\}$$

for a 25 metre long pipe with a manhole at either end, and 2 metre long lateral connections at 7 and 15 metres along the pipe. Relative to this, the robot's position would be given in the same one-dimensional coordinates, for example $x_t = 10$.

B. Deconvolution and Signal Processing

A simple model of acoustic echoes in a pipe network is described here, illustrating their utility for robot localization. A robot with a loudspeaker and a microphone can emit a sound and listen to the response, and use this information to make perceptions about its surroundings. The acoustic impulse response, or transfer function, between the loudspeaker and microphone is desired. This is the signal the microphone would detect if the speaker emitted an infinitesimal impulse of sound. The received signal $r(k)$ is given in discrete time k [21] by

$$r(k) = s(k) * h_{\mathbf{x}}(k) = \sum_{j=0}^{k} s(j) h_{\mathbf{x}}(k-j) \tag{4}$$

which is the convolution of the transmitted signal $s(k)$ and the pipe transfer function between the loudspeaker and microphone $h_{\mathbf{x}}(k)$ at position $\mathbf{x}_t$, which depends on the state to be estimated.

In practice it is not possible to find the impulse response directly as an infinitesimally short impulse cannot be produced by a speaker. Instead, it can be estimated by sending a chirp signal from the loudspeaker, and using deconvolution between the measured and emitted sound, which is described by

$$H_{\mathbf{x}}(\omega) = \frac{R(\omega)}{S(\omega)} \tag{5}$$

where each term is the Fourier transform of the terms above.

The impulse response $h_{\mathbf{x}}(k)$ is input to a band-pass filter which removes low-frequency oscillation, high frequency noise, and high frequency high-order modes of acoustic propagation. This band-pass filter is characterised by parameters f_l^h and f_u^h which describe the cut-on and cut-off frequencies.

C. Feature Detection

The impulse response $h_{\mathbf{x}}(k)$ at position $\mathbf{x}$ will be a sum of a sequence of impulses, with a time delay given by the distance travelled between loudspeaker, source of reflection, and microphone [24], [40], [41]. This is represented as

$$h_{\mathbf{x}}(k) = \sum_{\psi=1}^{\Psi} g^{\psi} \delta(k - \kappa^{\psi}) \tag{6}$$

Each of N components is a Dirac delta impulse with magnitude g^{ψ} and delay κ^{ψ}.

The distance to the source of reflection ξ^{ψ} is found using

$$\xi^{\psi} = \frac{1}{2} c \kappa^{\psi} \tag{7}$$

where c is the wave speed. It is assumed that the wave speed is known, requiring calibration in the case of operation in varying temperature and humidity.

The discrete components of the impulse response must be estimated from the acoustic impulse response. Figure 1(b) shows an example of the acoustic echo detection process. The filtered impulse response signal is seen to be made up of a sequence of oscillatory, noisy components, each likely corresponding to a different source of reflection. *Envelope detection* smooths the signal by squaring the signal and applying a low-pass filter characterised by cut-off frequency f_u^h giving an impulse response $\hat{h}_{\mathbf{x}}(k)$. The result of this is illustrated in Figure 1(b).

The smooth signal is closer to the ideal impulse response, and discrete echo components can be found using *peak detection*, which finds salient points in the signal which have an amplitude greater than a threshold ϵ^g. A value of ξ_t^{ψ} is found for each peak in the signal $k = \kappa^{\psi}$ using equation 7.

Overall, $\mathbf{z}_t$ is made up of values ξ_t^{ψ} for each feature ψ in the current measurement.

$$\mathbf{z}_t = \boldsymbol{\xi}_t = \left\{\xi_t^1, \ldots, \xi_t^{\psi}, \ldots, \xi_t^{\Psi}\right\} \tag{8}$$

It is not known which feature in the environment ϕ corresponds to which observed feature ψ.

D. Measurement Model Definition

From an acoustic echo impulse response, the distance to sources of reflection, or features, in the environment can be estimated. The measurement of distance does not correspond to the euclidean distance to the feature, but instead to the distance through the pipe network to the feature. The distance measurement is therefore one-dimensional, given by ξ_t^{ϕ}, which is the measurement of distance to feature ϕ at time t.

A model $p(\mathbf{z}_t|\mathbf{x}_t)$ can be constructed which gives the measurements that are expected at a given state $\mathbf{x}_t$ in pipe q The measurement model consists of three sets of components. Here they will be described generally with respect to the map defined in equation 3 and illustrated with an example relating to the pipe network described in Section III-A where the robot is at position $x_t = 10$. The first subset would describe the echoes from each of the lateral connections and manholes, and would be a set of distributions centred around the distance to features along the main pipe

$$\left\{|m^{\phi} - x_t| : m^{\phi} \in \mathcal{M}_c^q\right\} = \left\{10, 3, 5, 15\right\}$$

The second subset would describe the echoes from the features at the end of each of the lateral connections, and is given by

$$\left\{|m_l^{\phi} - x_t| + m_l^{\phi} : \phi \in \left\{1, \ldots, \Phi\right\}, m_c^{\phi} \in \mathcal{M}_c^q, m_l^{\phi} \in \mathcal{M}_l^q\right\} = \left\{10, 5, 7, 15\right\}$$

Algorithm 1 Echo Localization Algorithm

Get input variables
input $\mathbf{z}_{1:T} = \boldsymbol{\xi}_{1:T}, \mathbf{z}_t = \boldsymbol{\xi}_t = \left\{\xi_t^1, \xi_t^2, \ldots, \xi_t^\psi, \ldots, \xi_t^\Psi\right\}$
input $\mathbf{u}_{1:T} = \left\{u_1, u_2, \ldots, u_t, \ldots, u_T\right\}$
input Σ_t^z, Σ_t^u

Estimate measurement location
$\left[\tau_{0:J}\right] =$ `estimate-measurements-location` $\left(\boldsymbol{\xi}_{1:T}, \Sigma_t^z\right)$

Classify static measurements
$\left[\boldsymbol{\xi}_{1:T}^\Lambda, \boldsymbol{\xi}_{1:T}^z\right] =$ `estimate-measurement-class` $\left(\boldsymbol{\xi}_{1:T}, \tau_{0:T}, \Sigma_t^z\right)$

Estimate feature positions
$\left[(\mathbf{x}_{1:T}^\xi, \mathbf{l}_{1:T}^\xi)\right] =$ `estimate-feature-positions` $\left(\boldsymbol{\xi}_{1:T}^z, \mathbf{u}_{1:T}, \tau_{0:T}, \Sigma_t^u, \Sigma_t^z\right)$

Estimate feature relations
$\left[\mathbf{u}_{1:T,1:T}^z\right] =$ `estimate-feature-relations` $\left((\mathbf{x}_{1:T}^\xi, \mathbf{l}_{1:T}^\xi), \mathbf{u}_{1:T}, \tau_{0:T}, \Sigma_t^u, \Sigma_t^z\right)$

Estimate trajectory with pose-graph optimization
$\left[\hat{\mathbf{x}}_{0:T}, (\hat{\mathbf{x}}_{1:T}^\xi, \hat{\mathbf{l}}_{1:T}^\xi)\right] =$ `pose-graph-optimization` $\left(\mathbf{u}_{1:T,1:T}^z, \mathbf{u}_{1:T}, x_0, \Sigma_t^u, \Sigma_t^z\right)$

Estimate map
$\left[\hat{\mathcal{M}}^q\right] =$ `estimate-map` $\left((\hat{\mathbf{x}}_{1:T}^\xi, \hat{\mathbf{l}}_{1:T}^\xi), \tau_{0:T}, \Sigma_t^z\right)$

The third subset describes the 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 this *static* echo which travels a total distance equal to twice the distance between those two features. In this case this is given by

$$\begin{aligned}&\Big\{|m_c^{\phi_1} - m_c^{\phi_2}| + n_l^{\phi_1} m_l^{\phi_1} + n_l^{\phi_2} m_l^{\phi_2} : \\ &\Big\{\phi_1, \phi_2 \in \left\{1, \ldots, \Phi\right\}, m_c^{\phi_1}, m_c^{\phi_2} \in \mathcal{M}_c^q, m_l^{\phi_1}, m_l^{\phi_2} \in \mathcal{M}_l^q : \\ &m_c^{\phi_1} < x_t, m_c^{\phi_2} > x_t\Big\}, n_l^{\phi_1}, n_l^{\phi_2} \in \left\{0, 1\right\}\Big\} \\ &= \left\{15, 25, 8, 18, 10, 20, 17, 12\right\}\end{aligned}$$

While this description of the set of components is complicated mathematically, it is simple to implement algorithmically.

E. Localization using Pose-graph Optimization

Given the measurements $\mathbf{z}_t$ and the probability distribution $p(\mathbf{z}_t|\mathbf{x}_t)$, a practical means of acquiring the posterior probability distribution $p(\mathbf{x}_{0:T}|\mathbf{u}_{1:T}, \mathbf{z}_{1:T}, \mathbf{x}_0, \mathbf{x}_T)$over a sequence of robot poses described in Section III-A is needed. Two approaches are described here. The *simple* approach naively uses typical pose-graph optimization, which is applied by making some simplifying assumptions, but is seen to perform poorly when using measurements from realistic pipe networks. This paper proposes the *advanced* algorithm, which usefully incorporates information from the acoustic echo sensing in this type of environment into pose-graph optimization.

1) Simple Algorithm: For comparison to the *advanced* algorithm proposed in this paper, a *simple* algorithm is implemented which is based on a typical robot localization approach. This simple algorithm is based on pose-graph optimization, but is designed with the assumption that *static* echoes have been detected and removed from the measurement sets. It also assumes that the robot does not pass lateral connections along the pipe axis.

A pose-graph optimization algorithm is used to estimate the robot's trajectory and the position of reflective features, given by $\mathbf{x}_{0:T}$, using all the measurements $\mathbf{z}_{1:T}$ and $\mathbf{u}_{1:T}$ simultaneously. The algorithm aims to find a value for $\mathbf{x}_{0:T}$ which minimizes, in a weighted least squares sense, the difference between the measurements $\mathbf{z}_{1:T}$ and $\mathbf{u}_{1:T}$, and the equivalent values from the models $p(\mathbf{z}_t|\mathbf{x}_t)$ and $p(\mathbf{x}_t|\mathbf{x}_{t-t}, \mathbf{u}_t)$.

A *scan-matching* [42] approach is used to estimate correspondence between features detected in different measurement sets. The implementation of this here has been specialised to the acoustic echo measurements. This allows for some simplification as the space has one dimension rather than the typical two, but also requires some additional complexity as the measurements are directionless.

2) Advanced Localization Algorithm: This section describes in detail the implementation outlined in Algorithm 1, which performs localization and mapping using a sequence of odometry and acoustic echo measurements.

As in the *simple* algorithm, a pose-graph optimization algorithm is used to estimate the robot's trajectory $\mathbf{x}_{0:T}$ and the position of reflective features, using all the measurements $\mathbf{z}_{1:T}$ and $\mathbf{u}_{1:T}$ simultaneously.

There is a *data association* problem of estimating the correspondence between measurements and features in the environment. As described in Section III-D, the sets of acoustic measurements contain simple acoustic echoes from both from features along the main pipe and from the far end of lateral connections, and *static* acoustic echoes which need to be classified and incorporated differently.

First, poses in the same section of pipe between two lateral connections are grouped as they share their perspective of features in the pipe. Algorithm 2 describes a function `estimate-measurements-location` which does this using only acoustic echoes. To do this, a pose-graph is constructed, where the edges between each node correspond to the probability that the two nodes, or poses, are within the same section of pipe. Measurements at poses in the same section of pipe will have the same *static* acoustic echoes, so the sets of measurements will have some similarity independent of their corresponding pose. A probability distribution over the measurement space is created from each set of measurements. The desired probability is then estimated by finding the *Kullback-Liebler Divergence* between sets of measurements at different poses, giving a matrix of divergences $D_{0:T,0:T}^\Lambda$. The divergence values here tend to fall into two groups d_1^Λ and d_2^Λ, which are identified using *k-means* clustering, and from which a threshold value of divergence ϵ^d is found. This threshold is used to define a *sigmoid* function $\mathcal{S}$ to which

Algorithm 2 Estimate Measurement Location

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

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

Measure the difference between each distribution
Get the Kullback-Liebler Divergence $D(p_1(x)||p_2(x))$ for each pair of time steps, creating matrix $D_{0:T,0:T}$
for $(t_1, t_2) = \{(1,1), (1,2), \ldots, (2,1), \ldots, (T,T)\}$ **do**
$D_{t_1,t_2} = D(p(\Lambda_{t_1}|\boldsymbol{\xi}_{t_1})||p(\Lambda_{t_2}|\boldsymbol{\xi}_{t_2}))$
$= \sum_\Lambda p(\Lambda_{t_1}|\boldsymbol{\xi}_{t_1}) \log\left(\frac{p(\Lambda_{t_1}|\boldsymbol{\xi}_{t_1})}{p(\Lambda_{t_2}|\boldsymbol{\xi}_{t_2})}\right)$
end for
Find the mean divergence between each pair of time steps
$D^\Lambda_{0:T,0:T} = \frac{1}{2}\left(D_{0:T,0:T} + {D_{0:T,0:T}}^{\mathrm{T}}\right)$

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

Estimate probability that two measurement sets were taken in the same pipe section using sigmoid function $\mathcal{S}$
$\hat{p}^\Lambda_{t_1,t_2} = \hat{p}(\Lambda_{t_1} = \Lambda_{t_2}|\boldsymbol{\xi}_{t_1}, \boldsymbol{\xi}_{t_2})$
$= -\mathcal{S}(D^\Lambda_{t_1,t_2}, \mu = \epsilon^d, \sigma^2 = \Sigma_t^z)$

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

Do spectral graph partitioning
$L^\Lambda = \texttt{graph-laplacian}\left(\hat{P}^\Lambda_{0:T,0:T}\right)$
find $\mathbf{v}_2$ *using* $\lambda_2 = {\mathbf{v}_2}^{\mathrm{T}} L^\Lambda \mathbf{v}_2$ *where* λ_2 *is the second smallest eigenvalue of* L^Λ
Define subgraphs with time indices $\boldsymbol{\tau}$
$\Delta v_{2,t} = v_{2,t} - v_{2,t-1}, t \in \{1, 2, \ldots, T\}, v_{2,t}, v_{2,t-1} \in \mathbf{v}_2$
$\mathbf{t}^\Lambda_{j \in 0,1,\ldots,J} = \{\texttt{find-peaks}\left(\Delta \mathbf{v}_2\right)\} + \{0, T\}$
$\boldsymbol{\tau} = \left\{\mathbf{t}_j^\Lambda + 1, \mathbf{t}_j^\Lambda + 2, \ldots, \mathbf{t}_{j+1}^\Lambda : j \in 0, 1, \ldots, J\right\}$

Algorithm 3 Estimate Feature Positions

input $\hat{\mathbf{x}}_{0:T}$, $\boldsymbol{\xi}^z_{1:T} = \{\boldsymbol{\xi}^z_1, \boldsymbol{\xi}^z_2, \ldots, \boldsymbol{\xi}^z_t, \ldots, \boldsymbol{\xi}^z_T\}$

Create sets of hypothetical features $(\mathbf{x}_t, \mathbf{l}_t)$ for each measurement set through time
for $t \in \{1, 2, \ldots, T\}$ **do**
for $\psi_1 \in \{1, 2, \ldots, \Psi\}$ **do**
Project each measurement into the (x, l) space
$\hat{x}_t \pm \xi_t^{\psi_1} = x_t^{\psi_1} + l_t^{\psi_1}$ *defines subspace occupied by a measurement*
Hypothesise features in the (x, l) space using measurements pairs
for $\psi_2 \in \{1, 2, \ldots, \Psi\}, \psi_2 \neq \psi_1$ **do**
Compute hypothetical feature
$l_t^{\psi_2} = 0$, $\xi_t^{\psi_2} = x_t^{\psi_2} = x_t^{\psi_1}$, $l_t^{\psi_1} = \xi_t^{\psi_1} - x_t^{\psi_1}$
Create feature pair and add to vector $(\mathbf{x}_t, \mathbf{l}_t)$
$(x_t, l_t) = \left(x_t^{\psi_2,\psi_1}, l_t^{\psi_2}\right), \left(x_t^{\psi_2,\psi_1}, l_t^{\psi_1}\right)$
end for
end for
end for

the divergence values are input, giving the estimate of the probability $\hat{p}^\Lambda$ desired. *Spectral graph partitioning* methods are used to separate the pose-graph with adjacency matrix $\hat{P}^\Lambda$ into discrete subgraphs with different echo measurements.

With this estimation of relevant subgraphs with time indices $\tau_j \in \boldsymbol{\tau}$, the function `estimate-measurement-class` in Algorithm 1 estimates which measurements are the static acoustic measurements by comparing measurement sets corresponding to nodes 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 to static measurements as they have been observed multiple times. This is described by

$$\boldsymbol{\xi}^\Lambda_{\tau_j} = \texttt{find-peaks}\left\{\sum_{t \in \tau_j} p(\Lambda_t|\boldsymbol{\xi}_t)\right\} \tag{9}$$

The static measurements can be assigned as $\boldsymbol{\xi}^\Lambda_t = \boldsymbol{\xi}^\Lambda_{\tau_j}$ for $t \in \tau_j$. The normal measurements, which are distances to features, are therefore given by set subtraction $\{\boldsymbol{\xi}^z_t\} = \{\boldsymbol{\xi}_t\} - \{\boldsymbol{\xi}^\Lambda_t\}$.

The possible positions of the features corresponding to each of these echoes must be estimated. The measurements contain echoes from each end of the lateral connections. From measurements from only one section of the pipe, it cannot be determined whether a feature is the near or the far end of a lateral connection. Only after observing a lateral connection from both directions is it possible to make an estimate. Algorithm 3 describes a function `estimate-feature-positions` which makes this estimate. A two-dimensional space (x, l) is defined, where x corresponds to the position along the pipe axis, and l corresponds to distance from the pipe axis. Each echo is projected into this space, and occupies a linear subspace. A lateral connection will exist as a pair of features at the same position along the pipe axis, so these subspaces can be reduced to points in (x, l) which likely correspond to an echo from the far end of a lateral connection giving sets of points $(\mathbf{x}^\xi_{1:T}, \mathbf{l}^\xi_{1:T})$.

The function `estimate-feature-relations` compares each set of features $(\mathbf{x}_{t_1}, \mathbf{l}_{t_1})$, $(\mathbf{x}_{t_2}, \mathbf{l}_{t_2})$ at all times $t_1, t_2 \in \{1, 2, \ldots, T\}$ and finds the value of $u^z_{t_1,t_2}$, an estimate of the relation between poses at time steps t_1 and t_2, using

$$u^z_{t_1,t_2} = \arg\max_{\delta_x} \left\{\mathcal{C}\left(\left(\mathbf{x}_{t_1}, \mathbf{l}_{t_1}\right), \left(\mathbf{x}_{t_2} + \delta_x, \mathbf{l}_{t_2}\right)\right)\right\} \tag{10}$$

where $\mathcal{C}$ is the correlation function.

Subsequently, `pose-graph-optimization` can be done to estimate the sequence of robot poses using $\mathbf{u}^z_{1:T,1:T}$

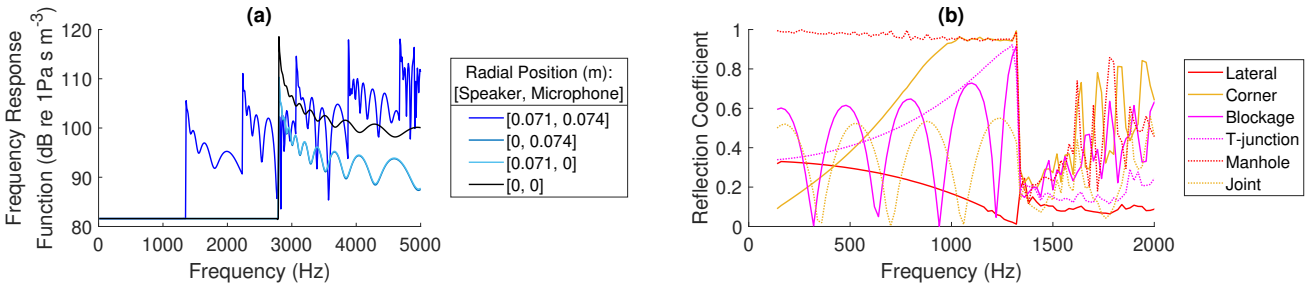

Fig. 2: (a) The analytical *frequency response function* in a 150 mm diameter pipe for different radial positions of the speaker and microphone, denoted in the legend as [speaker position, microphone position]. (b) The numerical acoustic *reflection coefficient* from different features in a 150 mm diameter pipe. The excitation source is placed at a radial position of 75 mm. Only the *plane wave* is shown, which is extracted from the simulated response by integration of the signal over the pipe cross-section.

and $\mathbf{u}_{1:T}$. The estimated values $u^z_{t_1,t_2}$ returned from equation 10 are typically accurate, but outlier inaccurate values are possible. Therefore, a robust pose-graph optimization approach [43] can be used where measurements in $\mathbf{u}^z_{1:T,1:T}$ which cause a large amount of residual error in the optimization are removed. Typically there are sufficiently many accurate measurements such that the estimation is not compromised if some are incorrectly removed.

Finally, mapping can be done, corresponding to the function `estimate-map` in Algorithm 1. For the set $\boldsymbol{\tau}$ of sets of times τ corresponding to different locations, estimated by Algorithm 2, the probability distribution of the map $\mathcal{M}$ given the estimated feature positions $(\mathbf{x}^\xi_{1:T}, \mathbf{l}^\xi_{1:T})$ is given by

$$p(\mathcal{M}|(\mathbf{x}^\xi_t, \mathbf{l}^\xi_t)t) = \prod_{\tau\in\boldsymbol{\tau}} \sum_{t\in\tau} \sum_n \mathcal{N}\left((\mathbf{x}^n_t, \mathbf{l}^n_t), \Sigma^z_t I\right) \qquad (11)$$

where $\mathcal{N}\left(\mu, \Sigma^2\right)$ is a normal distribution. This gives a result in two dimensions, in the space (x, l) defined above. Using the acoustic measurements described in Section III-D alone, it is not possible to determine the direction of each feature from the axis. This direction could be estimated by incorporating information from other sensors, or by using more advanced acoustic processing, but is beyond the scope of this work.

IV. Results

In this section, results from simulation and experimental measurements confirm the effectiveness of acoustic feature detection over a range of sensing parameters and conditions. Acoustic localization is demonstrated through laboratory experiments, the *simple* and *advanced* localization algorithms are compared, and the performance of the proposed localization algorithm is evaluated over a range of values describing the uncertainty in robot motion and perception.

A. Acoustic Echoes in Pipes

Two types of features are expected in the pipe environment:

- Lateral connections: Joints between pipes, which are either *T* or *Y* shaped, corresponding to an angle of 90 or 45 degrees between the *lateral* pipe and the *main* pipe.
- Manholes: Larger spaces connecting multiple pipes, which are cylinders where the central axis is oriented perpendicular to the pipe network plane.

Other features may be found in the pipe environment:

- Corners: A curved section of pipe connecting two pipes.
- Joints: A seam around the circumference of the pipe where two sections connect.
- T-junction: A lateral connection from the lateral pipe.
- Blockage: An unintended obstruction in the pipe.

1) Simulation Results: The investigation here shows the acoustic reflectivity of pipe features, which is required so that a robot can use them as features in the pipe environment for localization. A finite element method (*COMSOL* multiphysics simulation) was used to investigate the acoustic wave reflections from a range of features in pipes. A point source was used at the edge of the cross-section of a 150 mm diameter pipe, at a radial position of 75 mm. Modal orthogonality [36] was used to obtain the reflection coefficient from the different features for the plane wave mode described in Section II-B. This measure of reflection coefficient is independent of the distance to the feature, but dependent on frequency.

Figure 2(b) shows the plane wave reflection coefficient from a lateral connection, corner, blockage, T-junction (which is a lateral connection from a different direction), manhole, and joint between pipes in the frequency domain for the plane wave mode. The behaviour of each reflection coefficient changes substantially at around 1300 Hz, above which the acoustic propagation becomes more complex. There is a distinct variation in the detected echo from each feature.

2) Experimental Results: This section presents experimental acoustic echo measurements made in a variety of pipe configurations. It demonstrates the effectiveness of acoustic echo sensing in pipes, and allows a qualitative comparison of the measurements that can be made of each type of feature.

Figure 3 shows experimental acoustic echo measurements made in a range of pipe configurations. The sensing hardware (a Visaton 2242 loudspeaker and GRAS 46AE microphone) was placed as illustrated in each case, then the sensing procedure is followed as described in Section III-B and illustrated in Figure 1. The sensing hardware is placed at the pipe wall, and the signal processing band-pass filter parameters are $f^h_l = 400$ Hz and $f^h_u = 1200$ Hz.

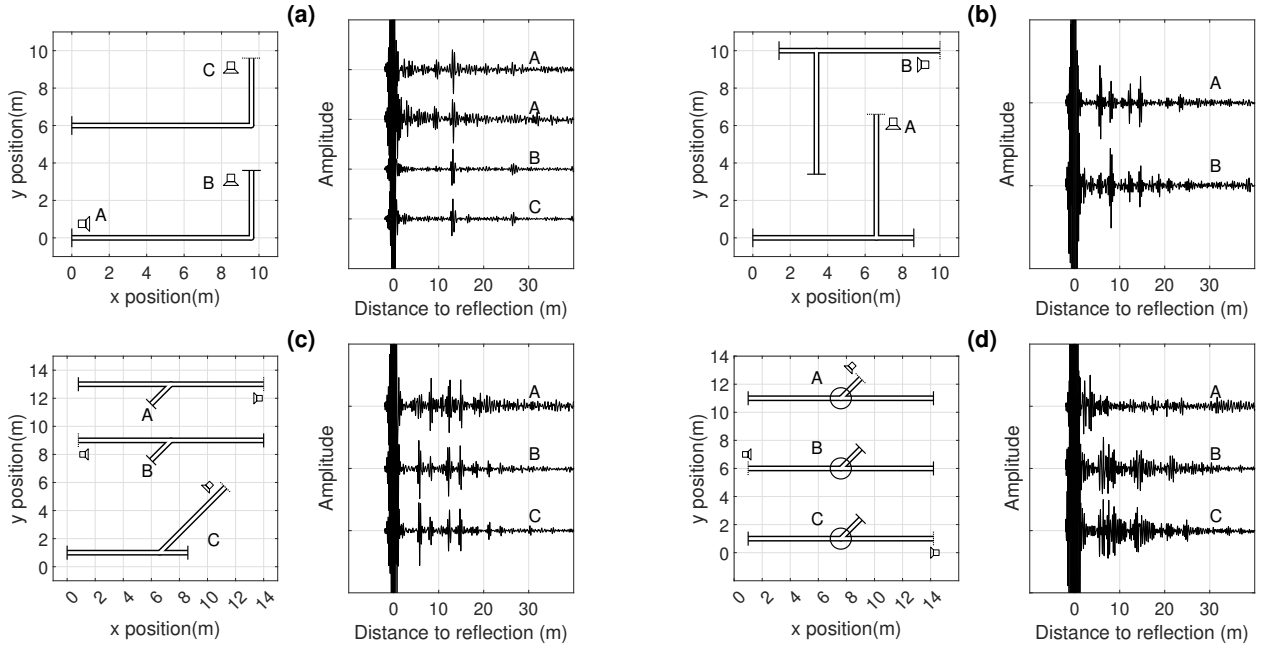

Fig. 3: Diagrams of different pipe network configurations and the impulse responses measured from each of them for 300 mm diameter pipes. The position of the speaker (with a colocated microphone) is indicated schematically, next to where the hardware was placed inside the pipes. Blockages depicted by a solid line at the end of a pipe are used to produce a reflection. Absorbers behind the speaker and microphone improve clarity by reducing reverberation. (a) A corner. (b) A lateral connection, or T-connection. (c) A lateral connection, or Y-connection. (d) A manhole.

The range of acoustically reflective features which are expected in a typical pipe environment that can be used for robot localization were investigated: Corners, lateral connections, and manholes. Reflective blockages were also used to investigate the acoustic transmission through the features. Figure 3 shows the differences between the impulse responses from these features in the time domain

The corner feature is reflective from both directions, but the reflectivity of the T-junction and Y-junction features depends on the relative position of the sensor. When the sensor was in the main pipe, and the lateral connection was at a $45°$ (Y) or $90°$ (T) angle, there was a smaller reflection from the connection, while there was a larger reflection from a $135°$ connection. There was a larger transmission into the lateral pipe for a $45°$ connection, and a smaller transmission for the $90°$ and $135°$ connections. Transmission into the main pipe depended on the angle of the connection; the transmission was equal in both directions for a $90°$ connection, while the transmission was larger in the direction of the more shallow angle for a $45°$ or $135°$ connection.

In all cases, an echo can be seen which has travelled the entire combined length of the pipe network twice, a phenomenon which has been observed previously and which can be both advantageous and disadvantageous to robot localization [19].

The manhole was reflective, as predicted by simulation, the result of which is shown in Figure 2(b), and any echoes returning from beyond the feature were unclear. The manhole was larger than the other pipe features, and was only one order of magnitude smaller than the distances between the sensor and features. Therefore, the echo returning from the manhole was longer than the echo from other features. The size of this feature limits the precision to which it can be located relative to the robot using acoustic sensing. The nonzero size of all of these features expected in the pipe network therefore would limit the benefits of increased sensing precision.

B. Acoustic Echoes for Different Sensing Parameters

This sensing system has a number of parameters which can impact the acoustic echo measurements. The sensitivity of the acoustic echo measurement 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. In all experiments here, a PYLE PDS122 loudspeaker and GRAS 46AE microphone were used.

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-C. Here, the envelope detection filter length is $f_u^{\hat{h}} = 50$ samples, for a sampling frequency of 16 kHz. The echo detection threshold is $\epsilon^g = 0.02$, 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-D, to calculate the precision and recall of the detection.

Figure 4(a) illustrates the acoustic echo measurements made for a range of lengths of sensing duration. This length impacts

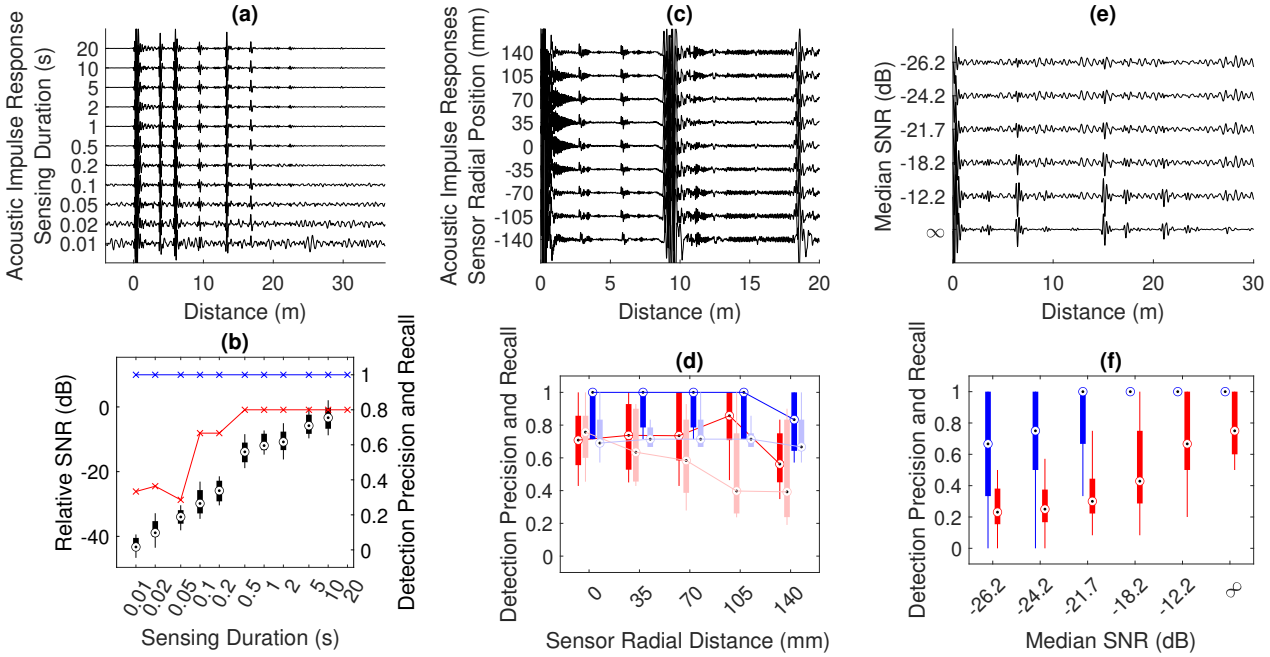

Fig. 4: The effect on acoustic echo sensing of three parameters: (a, b) the sensing duration, (c, d) the sensor position in the pipe cross-section, (e, f) the background noise. (a, c, e) show the measured acoustic impulse response. (b, d, f) show the precision and recall of the feature detection in red 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 $10\log_{10}\left(P_{signal}/P_{noise}\right)$, and the pipes were 150 mm in diameter.

the robot's operation and energy consumption as it should be stationary while outputting an acoustic signal for the duration. Figure 4(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 shorter sensing duration, while the recall remains at 1.

The total sensing duration is actually 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 4(c) illustrates the acoustic echo measurements made with the sensor at different positions across the diameter of the pipe. Section II-B described the complex acoustic propagation modes which exist in a pipe above a threshold frequency, which are to be avoided when a simple impulse response is desired as described in Section III-C. This threshold reduces when the sensor is moved away from the pipe axis, which may be required for 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 4(d) shows a reduction in feature detection precision when the sensor is positioned farther from the pipe axis, but shows that precision is reduced less 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 4(e) illustrates acoustic echo measurements made in the presence of different magnitudes of background noise. Intuitively, the echo impulses observable with no background noise diminish relative to the increasing background noise. The sensitivity of the feature detection system to background noise was measured by adding noise to the experimentally recorded data before the deconvolution process is used to find the acoustic impulse response. The background noise was applied at different magnitudes, and several sets of recorded background noise were used.

Figure 4(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, so a robust localization algorithm is needed. Improvements to sensing might increase the amplitude or duration of the transmitted acoustic signal in the presence of higher magnitude background noise or avoid

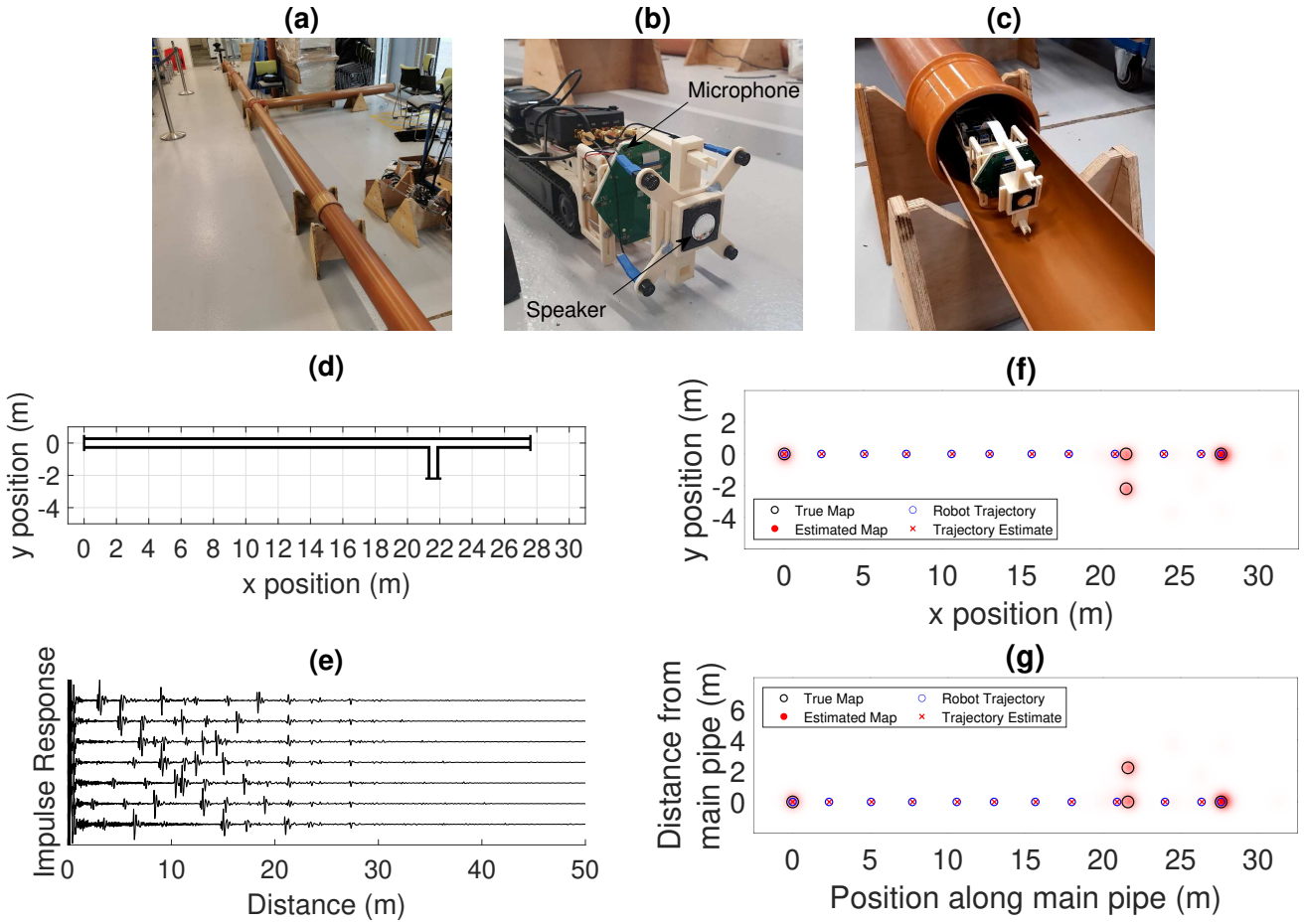

Fig. 5: Illustrative results from experimental data. (a) A photograph of the experimental apparatus, showing a main pipe with a lateral connection. (b, c) Photographs of the robot platforms used to record the experimental data. (d) A sketch of the pipe in (a), showing the position of the ends of the main pipe and the lateral connection. (e) Experimentally recorded echo impulse responses recorded at different positions along the main pipe shown in (a), offset in the y-axis. (f) Localization and mapping estimate transformed to (x, y) coordinates. (g) Localization and mapping estimate in coordinates relative to the pipe.

using acoustic sensing when there is substantial background noise.

C. Localization with Acoustic Echoes

This section shows robot localization and mapping using acoustic echo sensing with experimental and simulated measurements. A robotic platform shown in Figure 5(b) and (c) was moved along a pipe shown in Figure 5(a). The pipe was 27.6 m long, with a 2.2 m long lateral connection 21.6 m from one end. The pipe is made up of separate sections 3.6 m in length. The robot made an acoustic echo measurement approximately every 2 metres (using a Visaton 2242 loudspeaker and a MSM321A3729H9CP MEMSensing Microsystems Co., Ltd. microphone), and an odometry measurement was made between each acoustic measurement. The standard deviation of uncertainty was 0.1 m in the echo distance measurements, and 0.5 m in the odometry distance measurements. The measured acoustic impulse responses are shown in Figure 5(e).

Figure 5(f) shows the localization and mapping estimate transformed to (x, y) coordinates. This estimate requires knowledge or measurements of the below-ground parts of the environment, as well as the location of above-ground parts described in Section III-A. This could be achieved by sensing such as vision [6]–[8], or using general prior knowledge of the environment construction, such as that lateral connections are typically approximately $90°$ from the main pipe [1], or specific knowledge of the robot's environment, such as knowledge that a connection at a particular approximate location along the main pipe is on the north of the main pipe.

Estimating the direction of the lateral connection relative to the main pipe cannot be done using the acoustic echo sensing described in this paper. Figure 5(g) shows a localization and mapping estimate made using the experimental echoes in coordinates relative to the pipe, which is possible using this sensing method. The true map and the estimated map, which is a probability distribution, are compared, and the peaks in probability distribution are at the same points as the features in the true map.

Figure 6(a) shows a mapping estimate made using echoes produced in simulation for a longer pipe with multiple lateral connections which the robot passes. As in the experimental case, the true and estimated maps match.

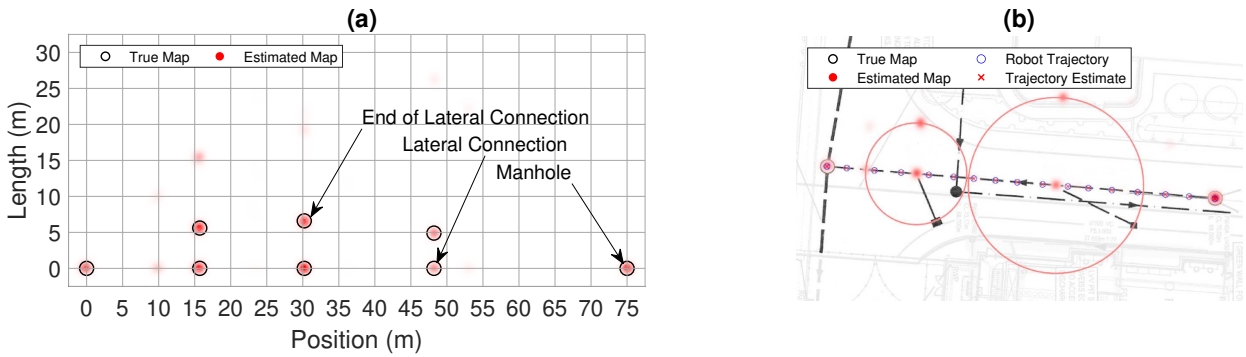

Fig. 6: Illustrative results based on simulated acoustic measurements. (a) Mapping results using simulated data for a pipe with a number of lateral connections. (b) Mapping results using another set of simulated data created with reference to a real buried pipe map, which is overlaid.

Figure 6(b) shows a mapping estimate made using echoes produced in simulation for a real pipe environment, indicating how this approach would be translated to a real application. Red circles show possible locations in (x, y) coordinates of lateral connection features detected in (x, l) coordinates.

D. 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. Sets of measurements $\boldsymbol{\xi}_t$ and $\mathbf{u}_t$ can be computed for a given trajectory $\mathbf{x}_{0:T}$. These can be used by the localization algorithms to estimate this trajectory.

The uncertainty in the robot's 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, an estimate of mean precision, $\tilde{\theta}_p$, and mean recall, $\tilde{\theta}_r$, can be made, from

$$\tilde{\theta}_p = \frac{N_{tp}}{N_{tp} + \frac{1}{2}\theta_{fp}} \tag{12}$$

$$\tilde{\theta}_r = \frac{N_{tp} - \frac{1}{2}\theta_{fn}}{N_{tp}} \tag{13}$$

where N_{tp} is the number of true positive measurements. The values of θ_{fp} and θ_{fn} can be varied to produce precision and recall values based on those seen in Figure 4.

The simulated measurements were made at intervals of around 2.5 m along a 75 m long pipe, so each trajectory is made of around 30 positions. The *error rate* of the algorithm 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 [44]. In Sections IV-E and IV-F, *error rate* is shown for 20 trajectories as a boxplot.

E. Localization Algorithm Comparison

The performance of the *Simple* and *Advanced* algorithms measured by extending the experimental data in simulation, as described in Section IV-D, is shown in Figure 7. The algorithms incorporating acoustic echo sensing are compared to an estimate made using only odometry. Here, the results are shown over a variation in the value of σ_x, with constant values of $\sigma_z = 0.1$, $\theta_{fp} = 1$ ($\tilde{\theta}_p \approx 0.97$), and $\theta_{fn} = 1$ ($\tilde{\theta}_r \approx 0.97$).

As described in Section III-E1, the *Simple* algorithm is designed with the assumptions that *static* measurements are removed and that lateral connections are not passed by the robot. Figure 7(a) shows the performance of the *Simple* and *Advanced* algorithms in a case where these two assumptions are valid. The two algorithms are seen to improve substantially on the estimate made using only odometry for nonzero odometry uncertainty. However, the *Simple* algorithm estimate error rate increases when odometry uncertainty becomes larger. Figure 7(b) shows the comparison in the case that static measurements are not removed from the measurement sets. In this case, the error rate of the *Simple* algorithm is still low for very low odometry uncertainty, but increases quickly and substantially as odometry uncertainty is increased, while the *Advanced* algorithm shows a low error rate. Figure 7(c) shows the comparison in the case that the robot passes lateral connections along its trajectory. In this case, the *Simple* algorithm error rate becomes substantial for even the lowest odometry uncertainty, and increases as odometry uncertainty increases. Figure 7(d) shows the comparison in the case that both of the assumptions are not valid, so static measurements are not removed and the robot passes lateral connections, which represents the most challenging and realistic case. Here the *Simple* algorithm error rate is consistently high, while the *Advanced* algorithm error rate remains low, although it does increase with increasing odometry uncertainty.

These results show the need for the *Advanced* algorithm when incorporating acoustic echo measurements in the most realistic case, and justifies the additional complexity when using the novel approach proposed here.

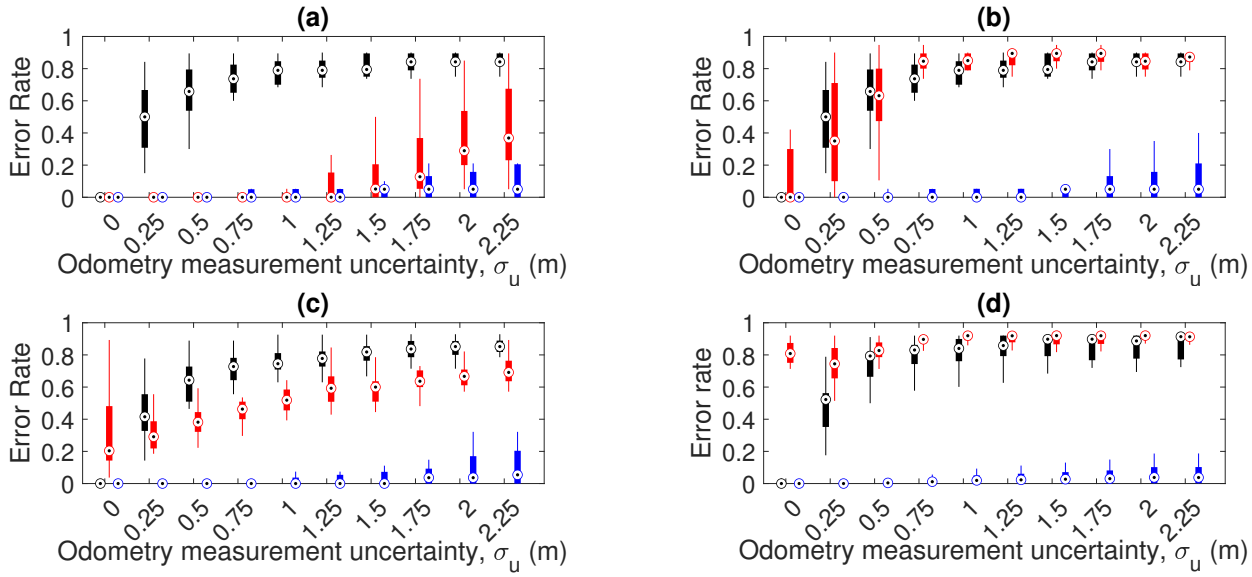

Fig. 7: A comparison of the *Simple* (red) and *Advanced* (blue) algorithms, in the case: (a) without static measurements or lateral connections; (b) with static measurements; (c) with lateral connections; (d) with both static measurements and lateral connections. In all cases, the variation of error rate over odometry measurement uncertainty is shown.

F. Localization Algorithm Evaluation

The *Advanced* algorithm described in Section III-E2 is evaluated over a variation in the uncertainty in the robot's motion and sensing. This is done by extending the experimental results in simulation as described in Section IV-D. The values of σ_x, σ_z, $\tilde{\theta}_p$, and $\tilde{\theta}_r$ are varied, with default values of $\sigma_z = 0.5$, $\sigma_z = 0.1$, $\theta_{fp} = 8$ ($\tilde{\theta}_p \approx 0.82$), and $\theta_{fn} = 2$ ($\tilde{\theta}_r \approx 0.94$).

For comparison, two odometry-based methods are also used. The first uses only odometry, and the second uses odometry and prior knowledge of the position of each end of the pipe, as described in Section III-A. A *loop-closure*-based method which emulates the effect of using vision to recognise previously observed locations was also compared. It was assumed that the recognition is perfect, although in practice there will be uncertainty. This vision-based *loop-closure* method was not used in the acoustic localization algorithm.

Figure 8(d) shows that the error rate of the two odometry-based methods increases quickly with increasing odometry uncertainty. The use of prior knowledge slightly reduces the error rate. The use of loop-closures is shown to give more robustness to odometry error than the use of odometry alone. However, above 0.25 m of odometry uncertainty, the error rate when using loop-closures quickly increases with increasing odometry error. Despite the ideal recognition of previously observed locations used for loop-closing, the uncertainty in motion still accumulates and causes estimate error.

Figure 8(a) shows that the *Advanced* algorithm has a low sensitivity to feature detection precision, as the error rate only increases substantially when precision (as calculated in equation 12) is reduced to 0.53, which means around half of the measurements made at each point in time are false positives. Figure 8(b) shows there is some sensitivity to false negative measurements. At a recall (as calculated in equation 13) of less than 0.8, the error rate increases substantially. A function which detects measurements in the echo impulse response would have some balance between probability of false positive and false negative detection. From these results, it can be concluded that such a function should aim to prevent false negatives at the cost of increased likelihood of false positives, to improve the subsequent localization estimate.

Figure 8(c) shows that the algorithm's performance in the presence of increasing measurement uncertainty. The error rate is seen to begin to increase when the measurement uncertainty becomes greater than 0.25 m. The median error rate is zero for measurement uncertainty up to 2.5 times the uncertainty in experimental measurement of 0.09 m reported in previous work [19]. This shows that there is a large margin in the variation of quality of sensing which can be used for successful localization. In practice, measurements may be less precise than in a laboratory experiment. However, there are advances in sensing and signal processing for this application [20] that could improve the measurement precision.

Figure 8(d) shows the algorithm's performance in the presence of increasing uncertainty in the motion measurement $\mathbf{u}_t$. The robot motion at each time step is around 2.5 m, and an odometry measurement of this travelled distance is made. The median error rate remains zero for uncertainty up to 1 m, which is 0.4 times the robot's travelled distance. This is 4 times the amount for which the median error rate of the loop-closure method remains zero. The localization error rate begins to increase substantially when the odometry uncertainty is increased over 0.6 times the travelled distance, but the median error rate remains close to zero.

Overall, the acoustic echo localization approach is seen to have a low error rate in the presence of false positive measurements, measurement uncertainty, and motion uncertainty. In the case of measurement and motion uncertainty, the proposed approach has a low error rate with larger magnitudes

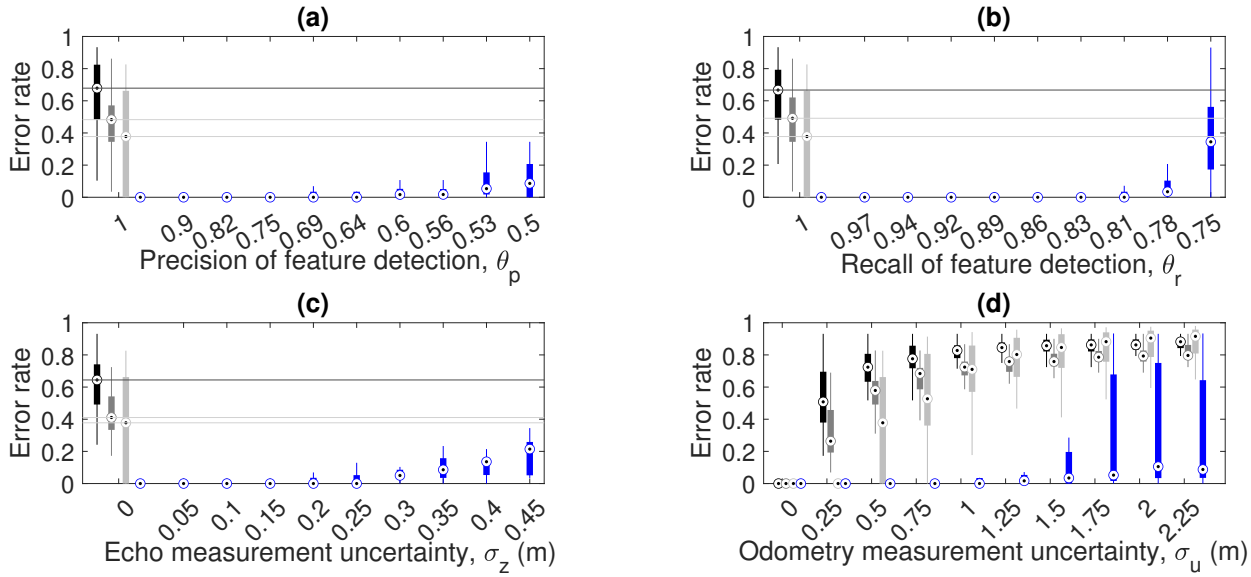

Fig. 8: Analysis of the sensitivity of localization algorithms to uncertainty in the input measurements. In all cases, the results over 25 tests are illustrated by boxplots. The black data is the result using only odometry measurements, not echo measurements. The dark grey data is the result using prior 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.

of uncertainty than those seen experimentally or reported in the literature, which gives some confidence in the approach.

V. Conclusion

This paper proposes a novel acoustic echo system for localization and mapping for mobile robots in pipes. Other sensing methods typically used in localization such as vision have limited scope and therefore accumulate uncertainty over time, and are unsuitable for the pipe environment which lacks features from their perspective. This paper shows that, in contrast, acoustic echo measurements can provide a drift-free location estimate of a robot along the length of a pipe with respect to typical features.

Experiments confirm the reflectiveness of different typical features in the pipe environment predicted by numerical simulation. The precision and recall of a feature detection algorithm have been characterized over variation in sensing duration, signal to noise ratio, and sensor position.

This paper developed a pose-graph optimization algorithm with bespoke aspects shown to be necessary to incorporate acoustic echo information from the reverberant environment. The performance of the algorithm has been evaluated over a range of values of different sources of uncertainty in robot motion and sensing. The algorithm has been shown to give a low estimate error rate for values of measurement uncertainty, precision, and recall beyond those expected in practice based on experimental measurements, with a median error rate is zero for echo measurement uncertainty 2.5 times that measured experimentally. The algorithm has also been shown to be more robust than an idealised typical loop-closure based approach, with an error rate of zero for motion measurement uncertainty 4 times larger than the value necessary to get the same error rate using the typical method.

One aspect of future work is more extensive experimental testing. The sensing hardware should be deployed on a mobile robot platform more similar to those used in the field, to allow measurement of any effects of robot motion, noise, or robot design, and should be done in buried pipes in the field with the noise and acoustic properties expected in practical application. Another aspect of future work is the fusion of information from multiple sensors, such as vision or lidar with the novel acoustic sensing method. Such experimental data would allow better investigation of the relative merits of these different sensing modes, and show the benefits of using the higher precision of vision sensing and the larger range of acoustic sensing.

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] Government Office for Science, "UK Water Research and Innovation Framework:Taking Responsibility for Water," 2012.

[4] 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/

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] L. Guan, X. Xu, Y. Gao, F. Liu, H. Rong, M. Wang, and A. Noureldin, "Micro-inertial-aided high-precision positioning method for small-diameter PIG navigation," *Advances in Human and Machine Navigation Systems*, 2018.

[13] 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.

[14] T. Seco, M. T. Lazaro, C. Rizzo, J. Espelosin, and J. L. Villarroel, "Graph-based robot localization in tunnels using RF fadings," in *Robot 2019: Fourth Iberian Robotics Conference*, 2019.

[15] A. Gunatilake, S. Kodagoda, and K. Thiyagarajan, "A Novel UHF-RFID Dual Antenna Signals Combined With Gaussian Process and Particle Filter for In-Pipe Robot Localization," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6005–6011, 2022.

[16] 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.

[17] K. Ma, M. M. Schirru, A. H. Zahraee, R. Dwyer-Joyce, J. Boxall, T. J. Dodd, R. Collins, and S. R. Anderson, "Robot mapping and localisation in metal water pipes using hydrophone induced vibration and map alignment by dynamic time warping," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 2548–2553, 2017.

[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] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] ——, "Look, no Beacons! Optimal All-in-One EchoSLAM," *ArXiv*, 2016.

[25] 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.

[26] 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.

[27] 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.

[28] U. Saqib and J. Rindom Jensen, "A model-based approach to acoustic reflector localization with a robotic platform," *IEEE International Conference on Intelligent Robots and Systems*, pp. 4499–4504, 2020.

[29] 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

[30] 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.

[31] 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.

[32] D. Murray, S. Panguluri, G. Skipper, and S. Donovan, "Demonstration of Innovative Sewer System Inspection Technology SewerBatt ™," Tech. Rep. January, 2014.

[33] 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.

[34] K. A. Papadopoulou, M. N. Shamout, B. Lennox, D. Mackay, A. R. Taylor, J. T. Turner, and X. Wang, "An evaluation of acoustic reflectometry for leakage and blockage detection," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 222, no. 6, pp. 959–966, 2008.

[35] L. Tao, "Monitoring Gas Distribution Pipelines," Ph.D. dissertation, The University of Manchester, 2017.

[36] P. Morse and K. Ingard, *Theoretical Acoustics*. Princeton university press, 1986.

[37] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2006.

[38] J.-l. Blanco, J.-a. Fernandez-Madrigal, and J. Gonzalez, "Toward a unified Bayesian approach to hybrid metric – topological SLAM," *IEEE Transactions on Robotics, VOL. 24*, vol. 24, no. 2, pp. 259–270, 2008.

[39] 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.

[40] U. Saqib and J. R. Jensen, "Sound-based distance estimation for indoor navigation in the presence of ego noise," *European Signal Processing Conference*, vol. 2019-Septe, pp. 1–5, 2019.

[41] D. Salvati, C. Drioli, and G. L. Foresti, "Sound Source and Microphone Localization from Acoustic Impulse Responses," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1459–1463, 2016.

[42] 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.

[43] 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.

[44] Institution of Civil Engineers, "Specification for underground utility detection, verification and location (PAS128:2014)," Tech. Rep., 2014.