Motion planning for robot audition

van Quan Nguyen, Francis Colas, Emmanuel Vincent, François Charpillet

Autonomous Robots manuscript No.
(will be inserted by the editor)

Motion planning for robot audition

Quan V. Nguyen · Francis Colas · Emmanuel Vincent · François Charpillet

Received: date / Accepted: date

Abstract Robot audition refers to a range of hearing capabilities which help robots explore and understand their environment. Among them, sound source localization is the problem of estimating the location of a sound source given measurements of its angle of arrival with respect to a microphone array mounted on the robot. In addition, robot motion can help quickly solve the front-back ambiguity existing in a linear microphone array. In this article, we focus on the problem of exploiting robot motion to improve the estimation of the location of an intermittent and possibly moving source in a noisy and reverberant environment. We first propose a robust extended mixture Kalman filtering framework for jointly estimating the source location and its activity over time. Building on this framework, we then propose a long-term robot motion planning algorithm based on Monte Carlo tree search to find an optimal robot trajectory according to two alternative criteria: the Shannon entropy or the standard deviation of the estimated belief on the source location. These criteria are integrated over time using a discount factor. Experimental results show the robustness of the proposed estimation framework to false angle of arrival measurements within $\pm 20^\circ$ and 10% false source activity detection rate. The proposed robot motion planning technique achieves an average localization error 48.7% smaller than a one-step-ahead method. In addition, we compare the correlation between the estimation error and the two criteria, and investigate the effect of the discount factor on the performance of the proposed motion planning algorithm.

Keywords Robot audition · Motion planning · Sound source localization · Extended mixture Kalman filter · Monte Carlo tree search

Quan V. Nguyen
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, 38000 Grenoble, France
* Institute of Engineering Univ. Grenoble Alpes
E-mail: nguyenquan012@gmail.com

Francis Colas, Emmanuel Vincent, and François Charpillet
Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
E-mail: francis.colas@inria.fr, emmanuel.vincent@inria.fr, francois.charpillet@inria.fr

1 Introduction

For an autonomous assistive robot, perceiving the sound environment is an important function. This is necessary for understanding human voices and gathering information about the environment such as the location and the activity of the sound sources. This hearing capability is called robot audition [41,44,49]. The auditory knowledge from robot audition completes the information delivered by other sensors like cameras or laser range-finders. When the robot gathers more knowledge about the environment, it knows better what to do next and human-robot interaction becomes more efficient.

Although still recent compared to static microphone array processing, robot audition has brought new advances in audio signal processing, especially in sound source localization. Using a static array, source localization in the far-field could only estimate the source angle of arrival (AoA) and even suffered from the so called front-back ambiguity in the case of a linear microphone array [41,30,45]. By exploiting robot motion, i.e., active audition [41], more information can be obtained. For instance, the front-back ambiguity can be eliminated by simply turning towards the target sound

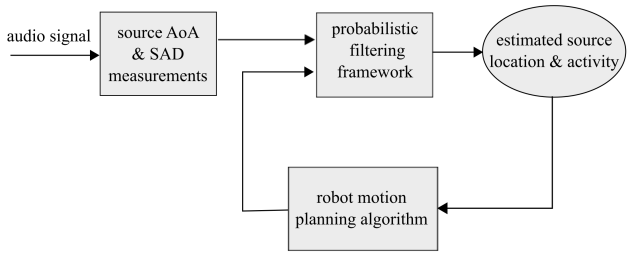

Fig. 1 General framework of the source localization for robot audition.

source [43,6]. Using body movements of the robot, the distance to the sound source [51,67] can also be estimated. Finally, robot audition provides the means of getting closer to the target source for increasing the signal-to-noise ratio (SNR) [39,61]. Based on this source location information, robots can separate a mixture of sounds, thereby increasing the performance of speech recognition. Based on that, they can process the data to extract other useful information (e.g., source identity, emotion, ...) and better understand their environment. In addition, by knowing the sound source location and its characteristics, they can make better decisions, e.g., decide to move closer for further investigation, or move away from a noisy area.

Generally speaking, detecting and precisely localizing a sound source is not an easy task due to noise in the measurements. Besides the direct signal from the target sources, the robot also receives reflected signals and noises, e.g., fan noise. Such an environment is qualified as reverberant and noisy. Reverberation and background noise severely degrade the performance of source localization [66,8] and also of source activity detection (SAD) [54,71]. In addition, the sound sources in the real world, e.g., speech, are not always active, making the estimation even more difficult. The silence intervals and the transitions between activity and silence intervals increase the uncertainty in detecting and localizing the source and induce false measurements. Most studies have tried to avoid this by assuming a priori knowledge of the source activity (i.e., whether the speaker is active or inactive in a given time frame) or by evaluating performance in an anechoic environment [51]. Obviously, these solutions do not perform well in a real-world situation where false measurements will happen. Moreover, when the source is not static or in the context of multiple sound sources, the uncertainty significantly increases. Therefore, conquering uncertainty is the key challenge for inferring precise information about the source location.

The general framework of source localization for robot audition is depicted in Fig. 1. The audio signal is captured by a microphone array on the robot. This audio signal could be any source that is active for a sufficient amount of time, e.g., a human speaker. The source AoA and SAD measurements are uncertain due to reverberation and the noisy background. By fusing these measurements with the robot motion information using a probabilistic filtering framework, we can deal with uncertainties in the observations and estimate the source location and activity over time. The estimation of the source location can be improved by implementing a robot motion planning algorithm. The optimal actions are selected and executed to minimize the estimation uncertainty.

In this paper, we address the uncertainty in the localization of a single, intermittent, and possibly moving source by exploiting robot motion. Our first contribution lies in the development of an extended mixture Kalman filter (MKF) framework for jointly estimating the location and activity of an intermittent and moving source in a reverberant environment. As an extension from our preliminary work [47], in this paper we provide a detailed analysis of the robustness of the extended MKF framework to both false SAD and false AoA measurements. The second main contribution concerns the adaptation of the Monte Carlo tree search (MCTS) method to the problem of long-term robot motion planning for improving source localization. A preliminary result of this method has been introduced in [48]. In this article, we provide significant improvements and a thorough analysis compared to our previous work. Specifically, we define a cost function for long-term robot motion planning by introducing two alternative criteria: the Shannon entropy or the standard deviation of the estimated belief. Moreover, we show and compare the correlation between the estimation error and these two criteria. In the cost function, the effect of the discount factor, i.e., the factor for tuning the trade-off between short vs. long term, on the performance of planning algorithm is also investigated. Finally, we compare our method with other motion planning methods for sound source localization.

The rest of this paper is organized as follows. In Section 2 we briefly revisit approaches related to the source localization and motion planning for robot audition. Section 3 details an extended mixture Kalman filter for robustly estimating source location. In Section 4, the proposed method for long-term robot motion planning is introduced. In Section 5, we report the experimental evaluation for the proposed algorithms. Finally, we provide concluding remarks in Section 6.

2 Related work

2.1 Sequential filtering for source localization

2.1.1 Frame-level localization

Source localization for static microphone arrays has been investigated since the beginning of research on signal processing. That is the case for a far field source, when the distance from the microphone array to the source is larger than the array size. In this situation, classical source localization methods can only estimate the source AoA [17].

These methods can be classified into three main classes [17]. The first class includes the techniques that exploit time difference of arrival (TDOA) information [31]. In the second class are the techniques which are based upon maximizing the steered response power of a beamformer [24,65,28]. The third class involves localization methods that are adapted from the field of high resolution spectral analysis [56]. Experimental comparisons of these algorithms are detailed in the literature [17,5,8].

After the source AoA has been obtained, the exact source location is often estimated by triangulation, either instantaneously or between time instants. For the instantaneous triangulation method, several microphone arrays are assumed to be in different known locations in the room environment and the AoA measurements obtained from different arrays at a given time are triangulated [2,22]. However, distributing multiple arrays over the room is often not feasible. The triangulation in time method is performed based on AoA measurements obtained at different times [51,67]. This method requires microphone arrays to move with a known motion. Even so, the triangulation result would be still not reliable if the AoA measurements are wrong. In that situation, we need to take into account the uncertainty of the AoA measurements.

2.1.2 Frame-level source activity detection

In the real world, sound sources often contain silent intervals which can degrade the performance of the source localization. Therefore, to achieve a robust source localization, the audio system should acquire information about the activity of the target sources. SAD refers to the problem of identifying the active or silent state of a single or multiple target sources in an audio signal. SAD is often used as an important front-end technique for many speech and audio processing systems [29,55]. The different detection principles include those based on SNR, energy thresholds [69], pitch detection [13], spectrum analysis [40], periodicity, dynamics of speech, zero-crossing rate or combinations of different features [62]. Some SAD techniques use statistical models to detect the activity of the source based on the average of the log-likelihood ratios between the observed signal and background noise in individual frequency bins [60]. Research in recent years has focused on developing robust SAD systems [55,21,1]. However, in a noise environment that has low SNR, SAD is still a serious challenge. In this work, we will not go into detail about the SAD technique. We assume that the SAD is not always perfect, and there can be 0% to 10% false detections, corresponding to false positive or false negative.

2.1.3 Sequential filtering

Over the last two decades, robot audition has started getting more attention in robotics and signal processing [25,64,42]. Most research still utilizes robot audition in a similar way as with a static microphone array [41,30,45]. Some research has focused on exploiting robot motion for improving signal processing results, especially in source localization [63,39]. Robots provide a mobile platform to continuously take AoA measurements at different locations and orientations of the microphone array. By using a sequential filtering algorithm for integrating robot motion with AoA measurements over time, we can reduce the uncertainty in the source AoA. In addition, the distance to the source, which is unavailable for a static microphone array, can be attained with robot audition. The estimated source location can be improved over time using an occupancy grid [38,16,67] or a particle filter [37,63,20]. The extended mixture Kalman filter (MKF) is especially suitable for this task [52,51], since it achieves the same performance as a particle filter at a lower computational cost[46]. However, the target audio source in most of these methods is continuously active [52,38] and static [67]. This is not true for many natural sound sources, e.g., speech: between successive words, successive sentences, and speaker turns, there often are silence intervals. A method for tracking one intermittent source is introduced in [51]. In that work, the SAD measurements are assumed to be perfect. In anechoic conditions, the experimental evaluation showed promising results [51]. However, in the real world, reverberation and noise severely degrade the measurement of AoA [66,8] and SAD [54,71]. A method to deal with false AoA measurements is presented in [53]. The estimation of the source activity from AoA measurements alone is presumably poor compared to exploiting additional SAD measurements. For this reason, we must jointly es-

timate the source activity and location from SAD and AoA measurements to achieve a robust estimation.

2.2 Robot motion planning

2.2.1 Overview

In general, robot motion can improve source localization [15,35,19]. For example, head rotation can resolve the ambiguity on the angle of arrival [43,6]. Moving towards a source can increase the signal-to-noise ratio [39,61]. However, no robot motion can be optimal in all situations. Therefore, robot motion planning must be considered in order to locate an audio source as precisely as possible.

Motion planning for mobile robot platforms has received extensive interest in robotics [33,34]. Classically, motion planning is considered as the problem of searching a feasible trajectory which guides the robot from its current position to a target position [33]. There are wide application prospects for robot motion planning, such as automation, driverless cars, robotic surgery, search and rescue [7,18,14], just to name a few. The development of robot motion planning includes planning with different constraints or objectives, planning under uncertainty, etc. In classical motion planning, avoiding collision is the most common constraint. However, besides that, there are objectives such as minimizing the time for reaching the goal, minimizing the consumed energy, or acquiring information about the environment. Gathering information about the environment through sensing is an important task for surveillance and mapping of unknown environments. This task is also called robot exploration.

In typical robot exploration, the general target of motion planning is to guide the robot moving to the direction that can maximize the explored area. One strategy is to move the robot to the nearest frontier [70], a border between empty and unknown environment. In [23], the next move is decided based on a next-best view approach. A multi-objective strategy to find the optimal exploration path under multiple constraints and objectives by modeling the environment constraints in the cost function is introduced in [4]. Most of these approaches consider a static environment in which the robot finds a destination for gathering information and then moves there without additional planning during the travel. For the case of dynamic environments with moving objects, continuous planning during movement is necessary.

2.2.2 Motion planning for robot audition

Recently, motion planning for robot audition has started attracting attention. The main objective is to find a robot trajectory that will maximize the information about the source location or minimize the uncertainty on the source location. In a simple strategy, the localization accuracy can be improved by following a fixed patrol loop and covering a potential maneuver area [39]. This approach is suboptimal because it often takes a lot of time for finding a potential area. In addition, if the target source is moving, the detected area will change. Therefore, it is not an efficient planning method and not posed as a formal optimization problem.

More sophisticated motion strategies based on information-theoretic criteria, e.g., Shannon entropy, have been proposed [67,12,57]. The target is to minimize the estimation error. Under the assumption that the chosen sequential filtering method is well-tuned and provides a reliable estimate of the belief on the source location, the expected error with respect to the (unknown) true source location can be replaced by a measure of the uncertainty of this belief. For instance, the uncertainty on the source location at a given time can be quantified as the entropy of the belief. The lower the entropy, the greater the amount of information. In [12], a gradient descent method was proposed to find the robot movement that minimizes the entropy of the belief on the position of a static sound source one time step ahead. This method yields a locally optimal robot motion but, in the long run, this sequence of local optima is generally not globally optimal. In addition, in [12] the belief about the source and robot positions is represented by a mixture of Gaussians but the determination of the best possible move is based on the single Gaussian with the largest weight.

An approach using Monte Carlo exploration to sample and select the next optimal action was proposed in [57]. In this approach, the final destination is predefined. The optimal sequence motion will minimize both the estimation uncertainty and the distance to the destination. This motion planning technique with a predefined destination is suboptimal when the target source is far from the robot's final destination. A long-term motion planning method was introduced by using a dynamic programming algorithm to find the optimal trajectory for localizing a static source in [67]. This method approximates the sum of entropies over a finite time horizon by assuming that the entropy at each future pose does not depend on the trajectory used to reach that pose. This assumption is required for dynamic programming, but in practice the entropy actu-

ally depends on the trajectory followed by the robot. This method is therefore also potentially suboptimal. Another long-term robot motion planning algorithm for a binaural sensor was proposed in [10]. An optimal motion of the robot is obtained by solving a constrained optimization problem involving the gradient of the reward function. An alternative motion planning approach, the so-called aural servo, was recently introduced [36]. This approach actually does not localize sound sources. It selects and executes a robot motion that satisfies given conditions characterized by a set of auditory measurements. Therefore, to generate an optimal robot motion in the context of intermittent and possibly moving source, we need an efficient long-term robot motion planning method with an appropriate cost function. The cost function with the entropy as the criterion should depend on the traveled trajectory. In addition, the destination of the robot trajectory would be flexible and depends on the possible movement of the target source.

3 Extended mixture Kalman filter

In this section, we present our contribution on developing an extended MKF framework which will be used later in our motion planning technique. This extended MKF framework jointly estimates the location and activity of an intermittent and possibly moving source. By explicitly estimating the source activity in addition to the source location, the filtering framework can deal with false measurements in both the SAD and the AoA.

3.1 State vector and dynamical model

3.1.1 State vector

Most the methods in the literature only track the source location when localizing a continuous or intermittent source [53,51]. By contrast, we also take the source activity into account in the state vector. Therefore, we define the state vector as follows:

$$\begin{bmatrix} X \\ a \end{bmatrix} = \begin{bmatrix} X_{\mathrm{r}} \\ X_{\mathrm{s}} \\ a \end{bmatrix} = \begin{bmatrix} x_{\mathrm{r}} \\ y_{\mathrm{r}} \\ \theta_{\mathrm{r}} \\ x_{\mathrm{s}} \\ y_{\mathrm{s}} \\ \theta_{\mathrm{s}} \\ v_{\mathrm{s}} \\ w_{\mathrm{s}} \\ a \end{bmatrix}, \tag{1}$$

where X_{r} is the pose of the robot, i.e., its absolute position $[x_{\mathrm{r}}, y_{\mathrm{r}}]$ and its orientation θ_{r} w.r.t. the x-axis; X_{s} is the continuous state of the sound source, i.e., its absolute position $[x_{\mathrm{s}}, y_{\mathrm{s}}]$, its orientation θ_{s} w.r.t. the x-axis, and its linear and angular velocities $[v_{\mathrm{s}}, w_{\mathrm{s}}]$; a is the source activity which is a discrete variable, where $a = 1$ indicates that the source is active, otherwise $a = 0$. It is to be noted that we have both the continuous variable X and the discrete variable a in the state vector.

In the present work, we assume that the pose of the robot X_{r} is known (e.g., through standard robot localization of SLAM techniques) and we focus on the estimation of the location X_{s} and the activity a of the source. In our work, we only consider the estimation and the AoA measurement in the horizontal plane.

3.1.2 Dynamical model of the robot

The dynamical model of the robot at time k can be written as

$$X_{\mathrm{r}k} = f_{\mathrm{r}}(X_{\mathrm{r}k-1}, u_k) + d_{\mathrm{r}k}, \tag{2}$$

where u are the robot commands which consist of the angular speeds v_{l} and v_{r} of the left and right wheel of the robot, and d_{r} is the process noise of the robot. Due to the above assumption that we know the pose of the robot, the process noise d_{r} is set to zero. We model the state transition function f_{r} as

$$f_{\mathrm{r}}(X_{\mathrm{r}}, u) = X_{\mathrm{r}} + \begin{bmatrix} \frac{r}{2}(v_{\mathrm{r}} + v_{\mathrm{l}})\cos(\theta_{\mathrm{r}}) \\ \frac{r}{2}(v_{\mathrm{r}} + v_{\mathrm{l}})\sin(\theta_{\mathrm{r}}) \\ \frac{r}{l}(v_{\mathrm{r}} - v_{\mathrm{l}}) \end{bmatrix} dt, \tag{3}$$

where l is the distance between the two robot wheels, r is the wheel radius, and dt is the time step or the sampling period [58].

Note that the control input u for the robot is given during the estimation, so for simplicity it is later omitted from the equations.

3.1.3 Dynamical model of the sound source

The dynamical model of the target sound source at time k is defined as follows:

$$X_{\mathrm{s}k} = f_{\mathrm{s}}(X_{\mathrm{s}k-1}) + d_{\mathrm{s}k}, \tag{4}$$

where d_{s} denotes the process noise of the sound source that has a Gaussian distribution with zero mean and covariance matrix Q_{s}, and f_{s} is modeled as

$$f_{\mathrm{s}}(X_{\mathrm{s}}) = X_{\mathrm{s}} + \begin{bmatrix} v_{\mathrm{s}}\cos(\theta_{\mathrm{s}}) \\ v_{\mathrm{s}}\sin(\theta_{\mathrm{s}}) \\ w_{\mathrm{s}} \\ 0 \\ 0 \end{bmatrix} dt. \tag{5}$$

3.1.4 Full dynamical model

We can write the dynamical model of the continuous state at time k as follows:

$$X_k = f(X_{k-1}) + d_k, \tag{6}$$

where $f(X) = \begin{bmatrix} f_\mathrm{r}(X_\mathrm{r}) \\ f_\mathrm{s}(X_\mathrm{s}) \end{bmatrix}$ and $d = \begin{bmatrix} d_\mathrm{r} \\ d_\mathrm{s} \end{bmatrix}$ is the process noise with covariance matrix $Q = \begin{bmatrix} 0 & 0 \\ 0 & Q_\mathrm{s} \end{bmatrix}$.

Since the source activities and the continuous states are conditionally independent, the transition probability between time steps is given by:

$$P(X_k, a_k | X_{k-1}, a_{k-1}) = P(a_k | a_{k-1}) P(X_k | X_{k-1}). \tag{7}$$

The state transition probability $P(X_k|X_{k-1})$ follows the dynamical model of the robot (2) and the source (4).

The source activity transition probability $P(a_k|a_{k-1})$ is defined by $P_\mathrm{appear} = P(a_k = 1|a_{k-1} = 0)$ which is the source appearance probability and $P_\mathrm{disappear} = P(a_k = 0|a_{k-1} = 1)$ which is the source disappearance probability.

3.2 Measurement vector and observation model

Audio source localization techniques can estimate the source AoA but not its distance. Therefore, we assume that the observation vector Z_k in a given time frame k consists of one AoA measurement Z_k^l (obtained via a localization technique) and one source activity measurement Z_k^a (obtained via an SAD technique).

The likelihood of the state vector w.r.t. this observation can be expressed as

$$P(Z_k | X_k, a_k) = \begin{cases} P_\mathrm{sn}(Z_k^\mathrm{l}|X_k) P(Z_k^\mathrm{a}|a_k) & \text{for } a_k = 1 \\ P_\mathrm{n}(Z_k^\mathrm{l}) P(Z_k^\mathrm{a}|a_k) & \text{for } a_k = 0, \end{cases} \tag{8}$$

with P_sn and P_n denoting the distribution of the measured AoA when the source is active or inactive, respectively. In the latter case, it is supposed that the recorded signal consists of spatially diffuse background noise, so P_n will have a uniform distribution and does not depend on X_k.

The observation model $P_\mathrm{sn}(Z_k^\mathrm{l}|X_k)$ represents the likelihood that a given localization technique applied to one signal frame provides the AoA measurement Z_k^l given the source position $X_{\mathrm{s}k}$ and the robot pose $X_{\mathrm{r}k}$. This likelihood does not depend on the orientation θ_s of the sound source, which appears solely in the dynamical model of the source and cannot be estimated from a single AoA measurement. Instead, θ_s is estimated by tracking the source movements and employed in the prediction step.

The distribution of the observation model differs for different microphone array geometries. For a linear microphone array, the distribution is bimodal. This is due to the front-back ambiguity: both the true AoA and its symmetric w.r.t. the microphone axis result in the same TDOAs and therefore cannot be distinguished [68]. As a result, the probability density concentrates around the true AoA and its symmetric w.r.t. the microphone axis, while the probability density for other AoAs is nonzero, but much smaller. Therefore we can approximate the observation model by a mixture of two Gaussians:

$$\begin{aligned} P_\mathrm{sn}(Z_k^\mathrm{l}|X_k) &= \sum_{j=1}^{2} \frac{1}{2} \mathcal{N}(Z_k^\mathrm{l}; \hat{Z}_k^{\mathrm{l},j}, R_k^j) \\ &= \sum_{j=1}^{2} \frac{1}{2} \mathcal{N}(Z_k^\mathrm{l}; h^j(X_k), R_k^j), \end{aligned} \tag{9}$$

where h^j is the observation model such that $h^j(X_k) = \hat{Z}_k^{\mathrm{l},j}$. $\hat{Z}_k^{\mathrm{l},j}$ and R_k^j are the mean and variance of component j, respectively. The two Gaussian densities in this expression are computed by wrapping the differences between the observed angle and the mean in $(-\pi, \pi]$.

Fig. 2 shows an example of the observation model $P_\mathrm{sn}(Z_k^\mathrm{l}|X_k)$ for a linear microphone array at different AoAs and distances. The observation model is built by applying the MUSIC-GSVD method to simulated data (see Section 5.1.1). At large distances, spurious peaks appear at 0° and 180° due to lower SNR.

For a planar but not linear microphone array, no front-back confusion would occur in the horizontal plane anymore, and the observation model would become unimodal.

3.3 Recursive Bayesian estimation

At the initial step, the belief about the source location is uniformly distributed across the room.

The posterior probability of the state vector can be recursively computed by alternating these two steps:

– **prediction step**: compute $P(X_k, a_k|Z_{1:k-1})$ given the previous belief $P(X_{k-1}, a_{k-1}|Z_{1:k-1})$ and the state transition model:

$$\begin{aligned} P(X_k, a_k | Z_{1:k-1}) = \sum_{a_{k-1}} \int & P(a_k|a_{k-1}) P(X_k|X_{k-1}) \\ & P(X_{k-1}, a_{k-1}|Z_{1:k-1}) dX_{k-1}. \end{aligned} \tag{10}$$

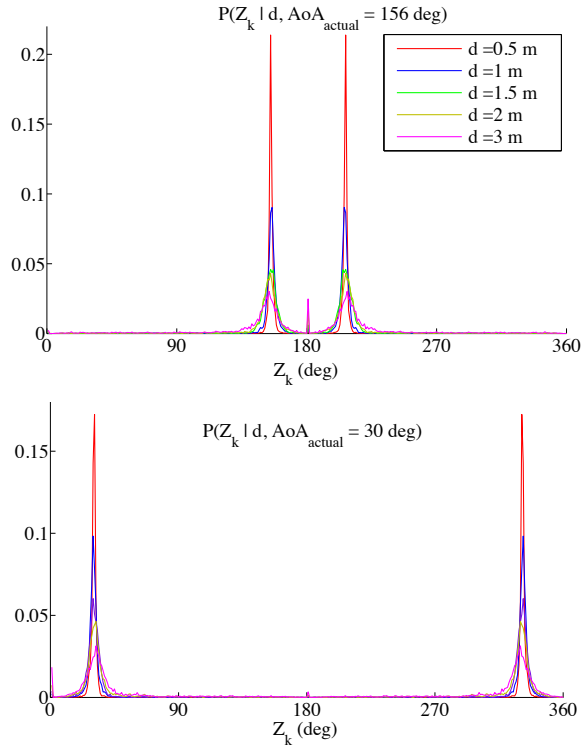

Fig. 2 Distribution of the measured AoA when the actual source is at different angles 156° (top) and 30° (bottom) and different distances from the microphone array.

– **update step**: recompute the belief $P(X_k, a_k|Z_{1:k})$ given the prediction and the new measurement Z_k:

$$P(X_k, a_k|Z_{1:k}) = \eta P(Z_k|X_k, a_k)P(X_k, a_k|Z_{1:k-1}), \tag{11}$$

where η is a normalizing constant.

Since the state vector includes both continuous and discrete variables and the observation model is a mixture of Gaussians, we propose an extended MKF to address these two issues. In the following, we present in detail the two steps above in the extended MKF.

3.4 Prediction step

In the first time step, we can assume that the estimated belief is a mixture of Gaussian. It is also true for the next time step, therefore it is true for all time steps and we can present the estimated belief by a mixture of Gaussians for every time steps. At the previous time step $k-1$, the belief about the state (X_{k-1}, a_{k-1}) is given by a mixture of $N_{k-1|k-1}$ Gaussians indexed by i with mean $\hat{X}^i_{k-1|k-1}$, variance $P^i_{k-1|k-1}$ and activity a_{k-1}:

$$\begin{aligned} &P(X_{k-1}, a_{k-1}|Z_{1:k-1}) \\ &= \sum_{i=1}^{N_{k-1|k-1}} \omega^i_{k-1|k-1} \\ &\quad \mathcal{N}(X_{k-1}; \hat{X}^i_{k-1|k-1}, P^i_{k-1|k-1})\delta_{a^i_{k-1}}(a_{k-1}), \end{aligned} \tag{12}$$

with weights $\omega^i_{k-1|k-1}$ such that $\sum_i \omega^i_{k-1|k-1} = 1$ and δ_a is the Dirac delta function centered in a.

Due to the presence of the source activity, which is a discrete variable, the mixture will evolve in the prediction step by duplicating each component into one active component and one inactive component following the source activity transition model. The components also move due to the continuous state dynamical model. Therefore, applying the prediction rule (10) to this density yields the predicted density:

$$\begin{aligned} &P(X_k, a_k|Z_{1:k-1}) \\ &= \sum_{i=1}^{N_{k-1|k-1}} \omega^i_{k-1|k-1} \big[\\ &\quad P(a^i_k = 0|a^i_{k-1})\mathcal{N}(X_k; \hat{X}^i_{k|k-1}, P^i_{k|k-1})\delta_0(a^i_k) \\ &\quad + P(a^i_k = 1|a^i_{k-1})\mathcal{N}(X_k; \hat{X}^i_{k|k-1}, P^i_{k|k-1})\delta_1(a^i_k)\big]. \end{aligned} \tag{13}$$

The predicted density $P(X_k, a_k|Z_{1:k-1})$ is thus also a mixture of Gaussians, with $N_{k|k-1} = 2N_{k-1|k-1}$ components and it can be expressed as follows:

$$\begin{aligned} &P(X_k, a_k|Z_{1:k-1}) \\ &= \sum_{i'=1}^{N_{k|k-1}} \omega^{i'}_{k|k-1}\mathcal{N}(X_k; \hat{X}^{i'}_{k|k-1}, P^{i'}_{k|k-1})\delta_{a^{i'}_k}(a_k). \end{aligned} \tag{14}$$

Its means, variances and weights are given by

$$\hat{X}^{i'}_{k|k-1} = f(\hat{X}^i_{k-1|k-1}, u_k), \tag{15}$$

$$F^i_{k-1} = \frac{\partial f(X, u_k)}{\partial X}|_{X=\hat{X}^i_{k-1|k-1}}, \tag{16}$$

$$P^{i'}_{k|k-1} = F^i_{k-1}P^i_{k-1|k-1}F^{iT}_{k-1} + Q^i_{k-1}, \tag{17}$$

$$\omega^{i'}_{k|k-1} = P(a^i_k|a^i_{k-1})\omega^i_{k-1|k-1}, \tag{18}$$

where index i is the index of the component at the previous time step on which component i' is based.

3.5 Update step

We obtain the new belief at time step k by applying the update rule (11) to the predicted density with the observation model. For the active components in the mixture,

the observation will be a mixture of two Gaussians as in (9). In this case, the components will be duplicated and each of them will be updated based on one of the two Gaussians in the observation. The inactive components in the mixture will be updated based on the uniform distribution, which is equivalent to assuming that no measurement has been assimilated during the update step. Note that this results in a different number of active vs. inactive components. Therefore, the updated belief at time k can be expressed as

$$\begin{aligned}
&P(X_k, a_k|Z_{1:k}) \\
&= \eta \sum_{i'=1}^{N_{k|k-1}} \omega_{k|k-1}^{i'} P(Z_k|X_k, a_k) \\
&\qquad \mathcal{N}(X_k; \hat{X}_{k|k-1}^{i'}, P_{k|k-1}^{i'}) \delta_{a_k^{i'}}(a_k) \\
&= \eta \sum_{i=1}^{N_{k-1|k-1}} \omega_{k|k-1}^{i} \Big[\\
&P(Z_k^{\mathrm{a}}|a_k^i = 0) P_{\mathrm{n}}(Z_k^{\mathrm{l}}) \mathcal{N}(X_k; \hat{X}_{k|k-1}^i, P_{k|k-1}^i) \delta_0(a_k^i) \\
&+ P(Z_k^{\mathrm{a}}|a_k^i = 1) \sum_{j=1}^{2} \frac{1}{2} \mathcal{N}(Z_k^{\mathrm{l}}; h^j(\hat{X}_{k|k-1}^i), R_k^{i,j}) \\
&\qquad \mathcal{N}(X_k; \hat{X}_{k|k-1}^i, P_{k|k-1}^i) \delta_1(a_k^i) \Big] \qquad (19)
\end{aligned}$$

In order to compute the product of every two Gaussians $\mathcal{N}(Z_k^{\mathrm{l}}; h^j(\hat{X}_{k|k-1}^i), R_k^{i,j}) \mathcal{N}(X_k; \hat{X}_{k|k-1}^i, P_{k|k-1}^i)$ we first linearize $h^j(\hat{X}_{k|k-1}^i)$ and then compute the product in closed form to obtain $\lambda^{i,j} \mathcal{N}(X_k; \hat{X}_{k|k}^i, P_{k|k}^i)$ where $\lambda^{i,j}$ is a constant defined as:

$$\begin{aligned}
\lambda^{i,j} = &\frac{1}{\sqrt{\left|2\pi(H_k^i P_{k|k-1}^i {H_k^i}^T + R_k^{i,j})\right|}} \\
&e^{-\frac{1}{2}\left[Z_k^{l,j} - h^j(\hat{X}_{k|k-1}^i)\right]^T {S_k^i}^{-1} \left[Z_k^{l,j} - h^j(\hat{X}_{k|k-1}^i)\right]},
\end{aligned} \qquad (20)$$

where $H_k^i = \frac{\partial h^j(X)}{\partial X}\big|_{X=\hat{X}_{k|k-1}^i}$ and $S_k^i = H_k^i P_{k|k-1}^i {H_k^i}^T + R_k^{i,j}$.

Therefore, the new belief can be expressed as a mixture of Gaussians:

$$P(X_k, a_k|Z_{1:k}) = \sum_{i=1}^{N_{k|k}} \omega_{k|k}^i \mathcal{N}(X_k; \hat{X}_{k|k}^i, P_{k|k}^i) \delta_{a_k^i}(a_k), \qquad (21)$$

where its number of components is $N_{k|k} = 3N_{k-1|k-1}$ and its weights:

$$\omega_{k|k}^i = \begin{cases} \omega_{k|k-1}^i P_{\mathrm{n}}(Z_k^{\mathrm{l}}) P(Z_k^{\mathrm{a}}|a_k) \eta \text{ if } a_k = 0 \\ \frac{1}{2} \omega_{k|k-1}^{i'} P(Z_k^{\mathrm{a}}|a_k) \lambda^{i',j} \eta \text{ if } a_k = 1. \end{cases} \qquad (22)$$

Its means and variances are computed as follows.
If $a_k = 0$, then $\hat{X}_{k|k}^i = \hat{X}_{k|k-1}^i$ and $P_{k|k}^i = P_{k|k-1}^i$.
If $a_k = 1$, then

$$\hat{X}_{k|k}^i = \hat{X}_{k|k-1}^{i'} + G_k^{i'} \left[Z_k^{l,j} - h^j(\hat{X}_{k|k-1}^{i'})\right], \qquad (23)$$

$$P_{k|k}^i = P_{k|k-1}^{i'} - G_k^{i'} H_k^{i'} P_{k|k-1}^{i'}, \qquad (24)$$

$$G_k^i = P_{k|k-1}^{i'} {H_k^{i'}}^T (S_k^{i'})^{-1}, \qquad (25)$$

where i' is the index of the component at the prediction step on which component i is based.

3.6 Pruning

From (19), we can see that the number of hypotheses in the extended MKF increases over time, which will consume a lot of memory and computational time. To deal with this problem, when the number of hypotheses is larger than $N_{\max}$ we simply keep the $N_{\max}$ hypotheses with the largest weights and prune the other hypotheses. The weights are re-normalized after pruning.

4 Motion planning for robot audition

This section presents our motion planning method for finding an optimal robot motion to minimize the uncertainty on the source localization. The belief about the source location is represented by a mixture of Gaussians and propagates by the extended MKF presented in Section 3. We present our contribution on defining the cost function for long-term motion planning with two alternative criteria: the Shannon entropy or the standard deviation of the estimated belief on the source location. Similarly to Markov Decision Processes, the two criteria are integrated over time using a discount factor. This is in contrast with [10] in which only the N-step ahead entropy is considered. After that, we present our contribution on adapting the MCTS method for efficiently finding the optimal robot motion which optimizes the cost function. We also investigate the effect of the discount factor on the performance of the motion planning algorithm.

4.1 Cost function

The goal of optimal motion planning for robot audition is to find the sequence of robot motions that will minimize the uncertainty on the estimated source location. We quantify this uncertainty by the Shannon entropy or the standard deviation of the estimated belief distribution.

Let us assume that the robot has taken measurements up to a certain time step k. All the knowledge about the source location, source activity and robot pose at time k is represented by the belief $P(X_k, a_k|Z_{1:k})$. Now, we consider moving the robot to a new pose at time $k+1$. To do so, we consider possible poses not only at $k+1$ but also at future times $k+2$, $k+3$, etc., up to horizon $k+T$. In the following, we use the terms "motion" and "action" interchangeably.

4.1.1 Shannon entropy criterion

For every possible motion sequence $u_{k+1:k+T}$ up to horizon $k+T$, we can quantify the uncertainty about the estimated source location by the entropy of the belief at each future time step $k+i$, $1 \leq i \leq T$. We sum these entropies with a discount factor $\gamma \in \mathbb{R}^+$ so that we can investigate the tradeoff between short vs long term. These entropies depend on future observations $Z_{k+1:k+i}$ which are unknown at current time k. By considering their expectation, the cost function can be expressed as

$$J_T = \sum_{i=1}^{T} \gamma^{i-1} \mathbb{E}_{Z_{k+1:k+i}|Z_{1:k}}[H(P(X_{k+i}|Z_{1:k+i}))], \quad (26)$$

where $\mathbb{E}$ and H respectively stand for the expectation and the entropy. In practice, we approximate the expectation by drawing N_s random samples Z^s from the distribution $P(Z_{k+1:k+T}|Z_{1:k})$. One sample from this distribution is obtained by first drawing a sample Z^s_{k+1} from $P(Z_{k+1}|Z_{1:k})$, followed by a sample Z^s_{k+2} from $P(Z_{k+2}|Z_{1:k}, Z^s_{k+1})$, and so on. A sample Z_s of $P(Z_{k+1:k+T}|Z_{1:k})$ is thus the concatenation of individual one-step-ahead observation samples, each generated based on the observation model and predicted state distribution. Then, we have:

$$J_T \approx \sum_{i=1}^{T} \gamma^{i-1} \frac{1}{N_s} \sum_{Z^s} H(P(X_{k+i}|Z_{1:k}, Z^s_{k+1:k+i})). \quad (27)$$

The entropies of the estimated beliefs cannot be computed in closed form for mixtures of Gaussians. An approximation based on a second-order Taylor series was proposed in [27] which we adopt hereafter. Assume that we have a probability density function $f(X)$ which is a mixture of Gaussians:

$$f(X) = \sum_{i=1}^{N} \omega_i \mathcal{N}(X; \mu_i, P_i). \quad (28)$$

The entropy of $f(X)$ can be computed as

$$\begin{aligned} H(f(X)) &\approx H_0(f(X)) - \sum_{i=1}^{N} \frac{\omega_i}{2} \int \mathcal{N}(X; \mu_i, P_i) \\ &F(\mu_i) \bullet (X-\mu_i)(X-\mu_i)^T dX \\ &= H_0(f(X)) - \sum_{i=1}^{N} \frac{\omega_i}{2} F(\mu_i) \circ P_i, \end{aligned} \quad (29)$$

where $\bullet$ is the matrix inner product which consists of an element-wise matrix multiplication and a subsequent summation of all matrix elements, $H_0(f(X))$ is computed based on the zeroth-order Taylor-series expansion

$$\begin{aligned} H_0(f(X)) &\approx -\sum_{i=1}^{N} \int \omega_i \mathcal{N}(X; \mu_i, P_i) \log f(\mu_i) dX \\ &= -\sum_{i=1}^{N} \omega_i \log f(\mu_i), \end{aligned} \quad (30)$$

and

$$\begin{aligned} F(X) = \frac{1}{f(X)} \sum_{j=1}^{N} \omega_j P_j^{-1} \Big(&\frac{1}{f(X)} (X-\mu_j)(\nabla f(X))^T + \\ &(X-\mu_j)(P_j^{-1}(X-\mu_j))^T - I \Big) \mathcal{N}(X; \mu_j, P_j). \end{aligned} \quad (31)$$

4.1.2 Standard deviation criterion

As an alternative criterion, for all possible motion sequences $u_{k+1:k+T}$ up to horizon $k+T$, we quantify the uncertainty by the standard deviation of the belief at each future time step $k+i$, $1 \leq i \leq T$. The standard deviation of the belief is the square root of its variance. In the case of a mixture of Gaussians, the first and second-order moments hence the variance can be computed in closed form. We also sum these standard deviation terms with a discount factor. The future observations are unknown at current time k but by considering the expectation, the cost function can be computed as

$$J_T = \sum_{i=1}^{T} \gamma^{i-1} \mathbb{E}_{Z_{k+1:k+i}}[\sigma(P(X_{k+i}|Z_{1:k+i}))], \quad (32)$$

where σ is the standard deviation of the estimated source location belief. Similarly, we can approximate the expectation over $Z_{k+1:k+T}$ by drawing N_s random samples Z^s from the distribution $P(Z_{k+1:k+T}|Z_{1:k})$:

$$J_T \approx \sum_{i=1}^{T} \gamma^{i-1} \frac{1}{N_s} \sum_{Z^s} \sigma(P(X_{k+i}|Z_{1:k}, Z^s_{k+1:k+i})). \quad (33)$$

The method to draw a sample from this distribution is the same as in Section 4.1.1.

4.2 Adapting MCTS for robot audition

In practice, considering all possible motion sequences $u_{k+1:k+T}$ is intractable. We propose a method to adapt the MCTS algorithm to solve that problem. In the standard MCTS for games, the outcome of each simulation is often a binary value which represents "win" or "lose". We adapt MCTS for robot audition by defining the outcome or the reward as a function of the entropy or the standard deviation of the estimated belief. In addition, each node in the tree will contain the expected belief on the source location. We use MCTS with the upper confidence bound for trees (UCT) algorithm [32] as the selection criterion to address the exploration-exploitation dilemma in motion planning. The dilemma is between choosing an action that results in an immediate reward (exploitation) and an action that might learn more information from the environment in the longer term (exploration).

In this section, we briefly describe how the MCTS algorithm [9] can be adapted to efficiently search the tree of possible sequences to obtain the optimal action and minimize the uncertainty on the source location.

4.2.1 Formulation

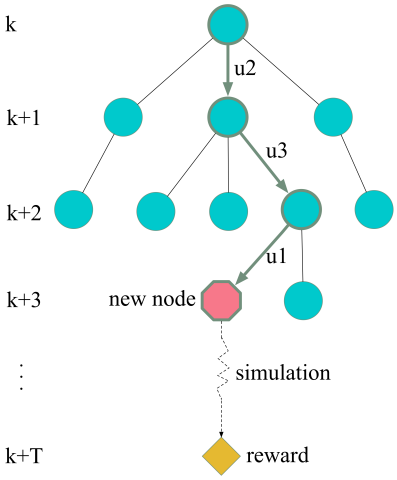

Fig. 3 An iteration of the MCTS algorithm.

Fig. 3 shows an iteration of the MCTS algorithm at time k. Each level of the tree corresponds to one future time step. Each node n contains the information about: the pose of the robot, the estimated belief $b(n)$ on the source location, the untried actions among a finite set of possible actions, the accumulated reward $\bar{Q}(n)$ (see below), and the visit count $N(n)$. The root node n_0 represents the pose of the robot at time k and carries the estimated belief $P(X_k, a_k|Z_{1:k})$. The links between a node and its child nodes represent different actions.

Starting from the root, a tree is built iteratively by selecting a node, adding a child corresponding to an untried action from this node, and following a random robot trajectory from this new node up to time $k+T$. The negative entropy or standard deviation Q corresponding to this trajectory is propagated upwards the tree to update the accumulated reward $\bar{Q}$ and the visit count N.

4.2.2 Selection

We select a node in the tree by applying the UCT algorithm [32]. For MCTS, UCT is a popular algorithm to balance between exploitation and exploration. UCT selects the child of a node that satisfies:

$$\text{node}_{\text{selected}} = \underset{n' \in \text{ children of } n}{\arg\max} \frac{\bar{Q}(n')}{N(n')} + C_p \sqrt{\frac{2 \log N(n)}{N(n')}}, \tag{34}$$

where $\bar{Q}(n')$ is the accumulated reward of the child node, $N(n)$ is the number of times the current (parent) node has been visited, $N(n')$ the number of times child n' has been visited, and $C_p > 0$ is a constant.

This UCT criterion derives from the Chernoff-Hoeffding inequality presented in Appendix A.

4.2.3 Expansion

The UCT criterion is iteratively used at each level to select a node until depth t where an untried action is chosen to create a new node at depth $t+1$. The belief at this new node is computed as follows.

The predicted belief distribution $P(X_{t+1}, a_{t+1}|Z_{1:t})$ is first obtained:

$$\begin{aligned} &P(X_{t+1}, a_{t+1}|Z_{1:t}) \\ &\quad = \sum_{a_t} \int P(X_{t+1}, a_{t+1}|X_t, a_t) P(X_t, a_t|Z_{1:t}) dX_t \\ &\quad = \sum_{a_t} \int P(a_{t+1}|a_t) P(X_{t+1}|X_t) P(X_t, a_t|Z_{1:t}) dX_t. \end{aligned} \tag{35}$$

Given one observation Z^s_{t+1} sampled as in Section 4.1.1, the updated belief at time step $t+1$ is expressed as:

$$\begin{aligned} &P(X_{t+1}, a_{t+1}|Z_{1:k}, Z^s_{k+1:t+1}) \\ &\quad = \eta P(Z^s_{t+1}|X_{t+1}, a_{t+1}) P(X_{t+1}, a_{t+1}|Z_{1:k}, Z^s_{k+1:t}), \end{aligned} \tag{36}$$

where η is a normalizing constant. Note that, although a single sample may not be sufficient to represent the distribution properly, each node in the tree will be visited several times and a different sample will be drawn each time, thereby better representing the distribution.

4.2.4 Simulation

From the new expanded node, we perform a simulation until time step $k+T$ by selecting an action at random at each time step. The procedure to update the estimated belief after selecting one action is identical to that in the expansion step. At the end of the simulation, we evaluate the reward by summing the negative entropy or standard deviation of the future belief from time step $k+1$ until time step $k+T$ with a corresponding discount factor γ:

$$Q = -\sum_{i=1}^{T} \gamma^{i-1} H(P(X_{k+i}|Z_{1:k}, Z^s_{k+1:k+i}), \tag{37}$$

or

$$Q = -\sum_{i=1}^{T} \gamma^{i-1} \sigma(P(X_{k+i}|Z_{1:k}, Z^s_{k+1:k+i})). \tag{38}$$

4.2.5 Backpropagation

After finishing the simulation, the number of times a node has been visited and the accumulated reward value are updated for all ancestor nodes, up to the root node. For each ancestor node n, $N(n)$ is incremented by 1 and $\bar{Q}(n)$ is incremented by Q.

4.2.6 Decision

The iterations of building the tree terminate when the MCTS has performed a fixed number of simulations. The optimal pose n at time $k+1$ is then chosen based on the average reward $\frac{\bar{Q}(n)}{N(n)}$. The robot performs the corresponding action in order to move to this pose, takes a new real measurement Z_{k+1}, builds a new tree to find the next optimal pose, and so on.

5 Experimental evaluation

The experimental evaluation is twofold. First, we evaluate the robustness of our extended MKF to false SAD and AoA measurements when tracking a single intermittent moving source in a noisy, reverberant environment. Second, we evaluate the performance of the MCTS algorithm for long-term robot motion planning to minimize the uncertainty in the source location.

In order to obtain statistically meaningful results, a large number of experiments is needed that can hardly be conducted with a real robot. Therefore, we conducted simulated experiments mimicking the *smart room* at Inria Nancy, where the robot is a Turtlebot equipped with a 4-microphone Kinect sensor forming a linear array.

5.1 Robustness of the extended MKF

5.1.1 Data

We employed state-of-the-art techniques for the simulation of reverberation [3] and acoustic noise, whose parameters are fixed as in [67] and closely match the real conditions in that room. More specifically, the reverberation time (250 ms), the intensity of speech and noise, and the noise spectrum match those of the real environment.

The source AoA was estimated by MUSIC-GSVD. We construct the probability distribution of AoA measurements by applying the MUSIC-GSVD for each of 360 true AoAs (from 0° to 359°) and 7 distances (from 0.5 to 3 m) instead of using an approximate Gaussian sensor model as in [12]. For other distances, we can interpolate the AoA distribution based on the AoA distribution of these 7 distances and sample the AoA measurement from that. We considered an SAD error rate of 5% and it could be a false negative or a false positive.

The target source is a single intermittent, possibly moving source. We considered four different scenarios, depending whether the sound source is either static or mobile ($v_s = 0.07\,\mathrm{m/s}$, $w_s = 8°/\mathrm{s}$) and inactive for several short time intervals (0.5 s) or a long time interval (2 s). For each scenario, we randomly generated 100 source trajectories for a duration of 10 s. The robot trajectory was fixed in all experiments with a maximum speed of 0.38 m/s. The robot starts from the point $(0, 0)$ in the coordinate, it goes straight forward for 3 s then makes a slightly turn after that and stops at 10 s.

5.1.2 Algorithm settings

We set the parameter values of the extended MKF as follows.
The time step is $dt = 0.1\,\mathrm{s}$.
We set the remaining parameters based on the typical characteristics of moving sound targets. The covariance matrix Q is set as

$$Q = \mathrm{diag}(0, 0, 0, 0.00095\,\mathrm{m}^2, 0.00062\,\mathrm{m}^2, (6.2°)^2, 0, 0). \tag{39}$$

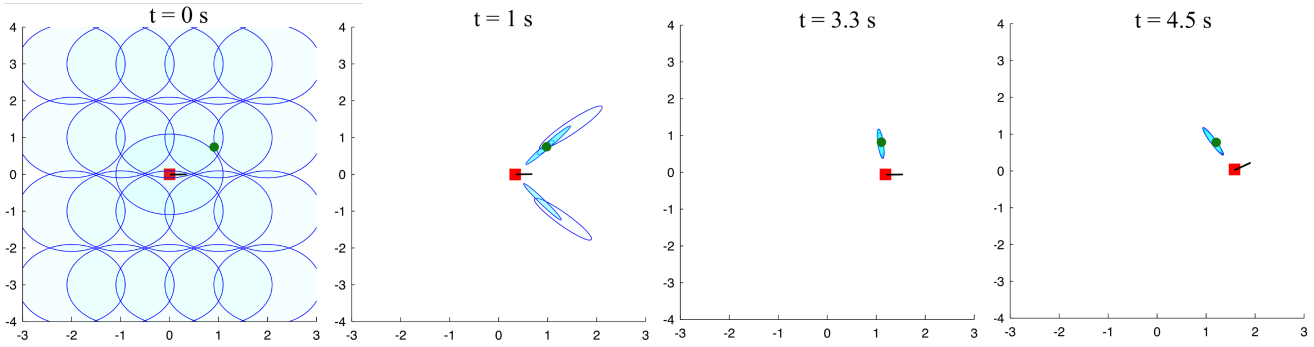

Fig. 4 Visualization of our extended MKF in an example localization. Robot positions are shown as red squares, and the actual source position as a green circle. Blue ellipses represent 95% confidence regions of source location estimation of various hypotheses in the mixture with a transparency proportional to the weight of the components.

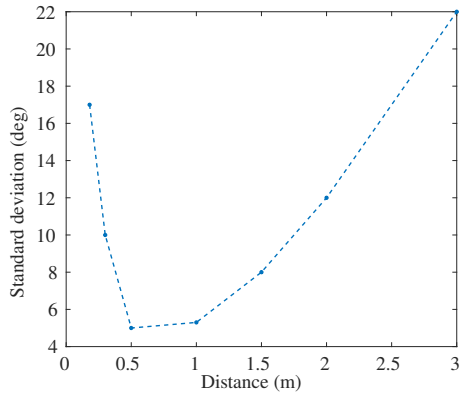

Fig. 5 The standard deviation of the observation noise as a function of the distance.

For the variance $R^{i,j}$, we obtain their values based on the histogram of the difference between the AoAs simulated in the reverberant room and the ground truth at 7 distances from 0.18 to 3 m. The variance values correspond to other distances are computed by a linear interpolation. Fig. 5 shows their standard deviation as a function of the distance. The source appearance/disappearance probabilities are set to $P_{\text{appear}} = 0.04$ and $P_{\text{disappear}} = 0.04$.
We set the number of hypotheses in the extended MKF to $N_{\max} = 50$.

5.1.3 Example localization

We show the performance of the extended MKF in tracking a single, intermittent, moving source in an example run. The first few seconds of tracking are showed in Fig. 4. At time $t = 0$ s, as there is no prior information about the source location, we initialize the mixture with several components that are evenly distributed over the room. After 1 s, there exists the front-back ambiguity as illustrated in Fig. 2 due to the linear microphone array. At this stage, half of the hypotheses for the source position are distributed along the direction from the source to the robot and the rest are symmetric w.r.t. the microphone axis. These symmetrical hypotheses become smaller and disappear after 3 s, thanks to the robot motion.

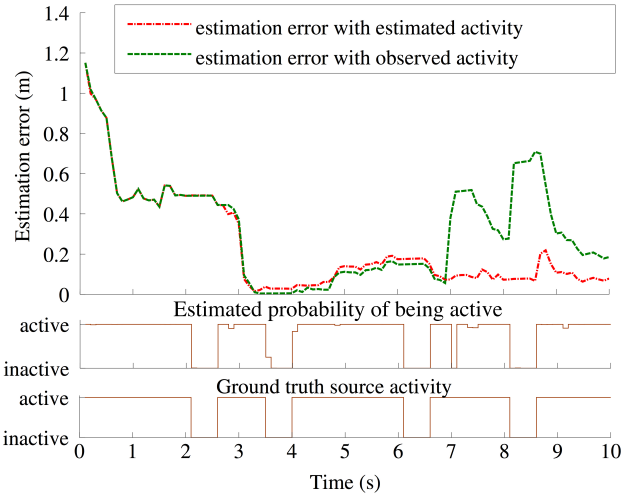

Fig. 6 Top: estimation error over time of our extended MKF with activity model vs the extended MKF without activity model. Middle: estimated probability of being active of the source over time. Bottom: ground truth source activity over time.

In Fig. 6, we plot the estimation error, that is the distance between the estimated source position and the true position, over the first 10 s of this example. We compare our extended MKF with estimated activity vs. an extended MKF based on the observed activity (assumed to be perfect). During the first 3 s when the front-back ambiguity still exists, both extended MKFs have a large estimation error. After 3 s, the front-back ambiguity disappears, and the estimation errors of the two extended MKFs reduce drastically. There are several times when false SAD happens. At time $t = 2.8$ s,

a false negative SAD measurement occurs: the source is active but the SAD method detects it as inactive. The estimation error of the extended MKF with observed activity becomes larger than ours but only for one time step. At time $t = 8.2\,\mathrm{s}$, a false positive SAD measurement occurs: the source is inactive but the SAD method detects it as active. Again, it appears that our extended MKF can handle this situation but the extended MKF with observed activity severely suffers from it. This was expected, since the tracking of the source activity implemented by our method was designed precisely to address this issue.

Fig. 7 shows the error in the AoA measurements over time. It is not a surprise that the errors are very large during the intervals when the source is inactive. However, even when the source is active, the AoA measurement error is slightly larger at several time steps. At time $t = 4.8\,\mathrm{s}$, the AoA difference between the observation and the ground truth is 9°, which is not a very large value. As a result, both the estimated error of our extended MKF and the extended MKF with observed activity increase slightly. At time $t = 7\,\mathrm{s}$, there is a high AoA measurement error: the ground truth AoA is 81°, but the measured AoA is 62° which is 19° difference. Although such a false measurement can occur with very low probability, it can have a major impact. Indeed, the estimation error of the extended MKF with observed activity increases drastically and remains large. By contrast, the estimation error of our extended MKF does not change much. This is an indirect benefit of the proposed approach: when a false AoA measurement occurs, the weight of the hypotheses corresponding to an inactive source increases, so that the belief is little affected. We can see this increasing weight in the plot of the estimated source activity at time $t = 7\,\mathrm{s}$ as the probability of the source being active is low.

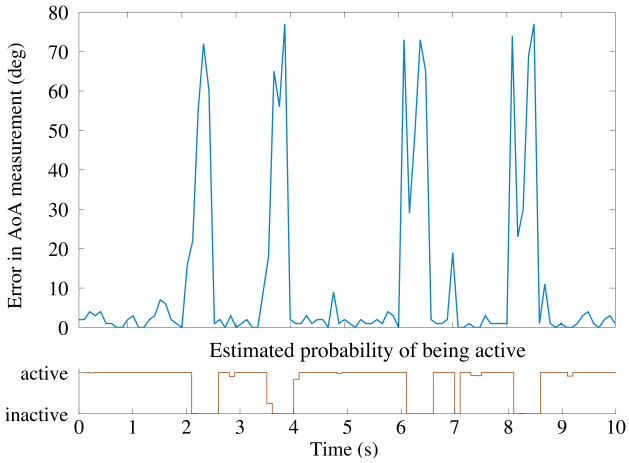

Fig. 7 Top: error in AoA measurements over time. Bottom: estimated probability of being active of the source over time.

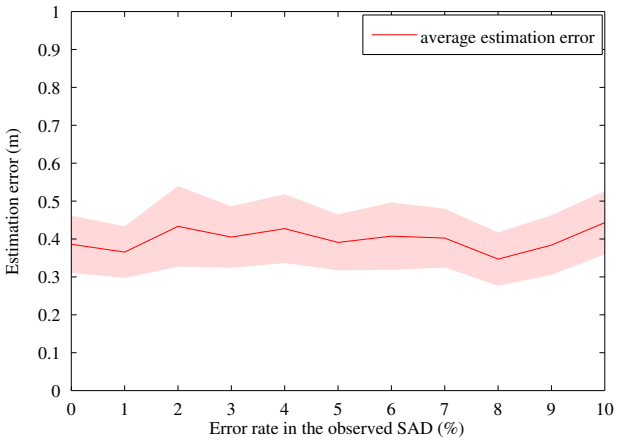

Fig. 8 Average estimation error on the estimated source location at the last time step and 95% confidence interval as a function of the error rate in the SAD observations.

5.1.4 Robustness of source location estimation

To evaluate the robustness of our proposed extended MKF framework to false SAD measurements, we change the error rate in the SAD observations from 0% to 10%. For each value of the SAD error rate, we randomly generate 100 initial source locations and run the extended MKF for 10 s.

Fig. 8 shows the average estimation error on the source location at the last time step as a function of the SAD error rate in the observation. The average values are around 0.4 m, and they are not significantly different when the SAD error rate varies. This confirms that our proposed extended MKF framework is robust to false SAD measurements.

To see how the SAD error rate assumed in the extended MKF model can affect the performance of the source localization, we tune the error rate in the extended MKF model between 0% and 10% and fix the error rate in the SAD observations $P(Z_k^a = 1|a_k = 0) = P(Z_k^a = 0|a_k = 1)$ to 5%. Similarly, we conduct 100 experiments with the same initialization as above for each error rate value.

Fig. 9 shows the average estimation error in the source location at the last time step as a function of the error rate in the extended MKF model. As expected, the estimation error is larger when the error rate in the extended MKF model is equal to 0%: this corresponds to the extended MKF based on the observed activity considered for comparison in Section 5.1.3. When the error rate in the extended MKF model is equal or greater than 1%, the estimation error is not significantly af-

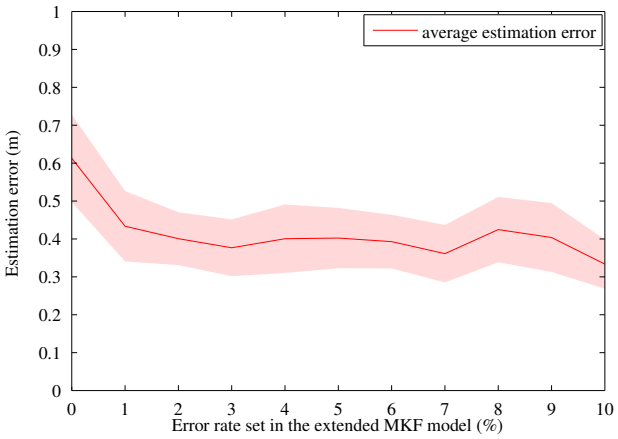

Fig. 9 Average estimation error on the estimated source location at the last time step and 95% confidence interval as a function of the error rate assumed in the extended MKF model.

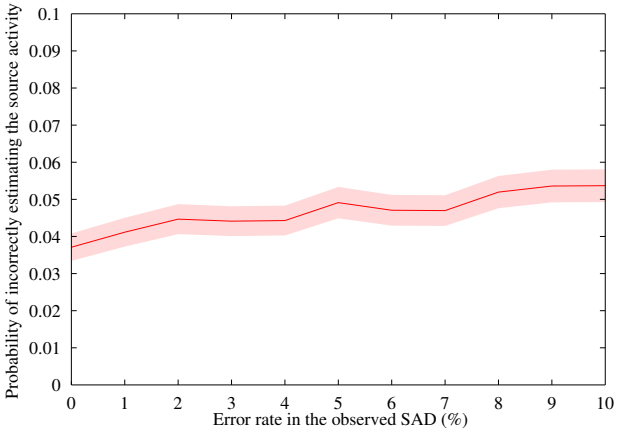

Fig. 10 Average probability of incorrectly estimating the source activity over all experiments and 95% confidence interval as a function of the error rate in the SAD observation.

fected and its value is around 0.4 m. From this, we can conclude that by taking into account the error rate of the SAD in the extended MKF model we can obtain better estimation of the source location.

5.1.5 Robustness of source activity estimation

Similar to the previous evaluation on the impact of the SAD error rate on the estimated source location, in this section, we evaluate the impact of the SAD error rate on the activity estimation.

At each time step, we compute the estimated source activity by summing the weights $\omega^i_{k|k}$ of all components for which $a_k = 1$. This results in a real value between 0 and 1, that is the estimated probability of the source being active. The probability of incorrectly estimating the source activity is the absolute value of the difference between the estimated source activity probability and the ground truth source activity.

Fig. 10 shows the average probability of incorrectly estimating the source activity over all experiments when we change the error rate of the SAD observations from 0% to 10%. We can see that this probability just increase slightly from 3.7% to 5.3% as the SAD error rate increases from 0% to 10%.

The probability of incorrectly estimating the source activity when we change the SAD error rate in the extended MKF model from 0% to 10% but keep the SAD error rate in the observation at 5% is presented in Fig. 11. This probability is slightly larger when the error rate in the extended MKF model equals 0%.

5.2 Consistency of the extended MKF

The filter is considered to be consistent if the true source location lies into the confidence region at least 95% of the time. We choose to take into account only the x and y coordinates of the source. The real source location $X_{\text{groundtruth}}$ lies into the confidence region if it satisfies [59]:

$$(X_{\text{groundtruth}} - X_{\text{mean}})^T \Sigma^{-1} (X_{\text{groundtruth}} - X_{\text{mean}}) \leq 2F^{-1}(0.95, 2) \quad (40)$$

where the right hand side of the equation is the inverse of the cumulative chi square distribution with 2 degrees of freedom, and X_{mean} and Σ are the mean and the covariance matrix of the belief on the source location.

We checked the consistency of the extended MKF at each time step over all our experiments. We found the empirical probability for the ground truth to be in the 95%-confidence ellipsis to be 95.19% over all time steps and all experiments. This value is close enough to 95% for us to conclude that the proposed extended MKF is consistent.

In the following section, we present the performance of MCTS and compare it with other motion planning methods.

5.3 Performance of MCTS based motion planning

5.3.1 Data

We simulated the robot and source movements and the resulting location and activity measurements. The source is silent from $k = 1.2$ s to $k = 2$ s. It is static

or mobile ($v_s = 0.07\,\mathrm{m/s}$, $w_s = 8°/\mathrm{s}$). For each motion planning approach, we generated 200 experiments with different random initial robot locations and source locations.

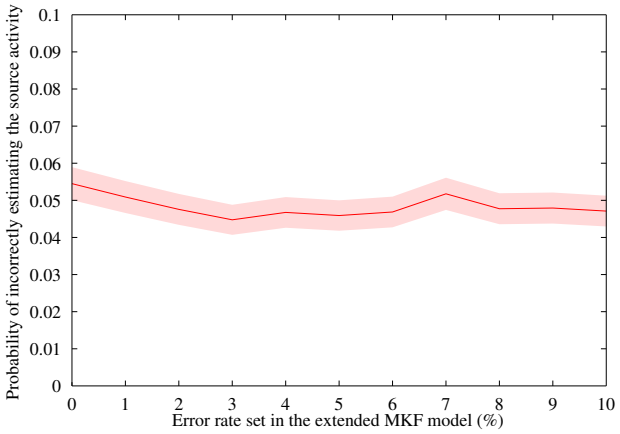

Fig. 11 Average probability of incorrectly estimating the source activity over all experiments and 95% confidence interval as a function of the error rate in the extended MKF.

5.3.2 Algorithm settings

If we run motion planning algorithms starting from a non-informative belief, all algorithms perform similarly because the robot does not know where to move when there is no information. Therefore, we start comparing them at a time when there is enough information such that they can make different decisions. To start from an informative belief, the robot first follows a fixed trajectory while updating the belief using the extended MKF for 3 s. After this, it follows the proposed MCTS algorithm with $T = 20$, and 700 tree nodes. The action set contains 13 discretized actions corresponding to different left velocity v_l and right velocity v_r values, as listed in Table 1.

Table 1 13 discretized actions of the robot

Index	1	2	3	4	5	6	7	8	9	10	11	12	13
v_l(m/s)	0.6	0.6	0.6	0.6	0.6	0.5	0.4	0.3	0.2	0.4	-0.6	0.6	-0.4
v_r(m/s)	0.6	0.5	0.4	0.3	0.2	0.6	0.6	0.6	0.6	-0.6	-0.6	-0.6	0.6

These actions basically correspond to moving forward, moving backward, turning left or right while moving forward or backward, turning with different speed and radius. One important remark here is that the robot is non-holonomous.

The optimal selected action is applied 5 times in a row, with 0.2 s time step. For comparison, we applied the same procedure to three other motion planning algorithms: a greedy algorithm inspired from [11,12,57], MCTS with an approximate cost function inspired from [67], and random motion. The greedy algorithm finds the optimal action that minimizes the expected entropy or the expected standard deviation of the belief one step ahead. The MCTS algorithm with approximate cost function computes the expected entropy or standard deviation for each future pose in (37) or (38) by recursively predicting the belief via (35) at all time steps, but updating it via (36) with an observation only for the last step. The random method simply chooses the action at random.

5.3.3 Example trajectory

In this section, we show an example scenario and compare the robot trajectory from the MCTS algorithm with the greedy algorithm.

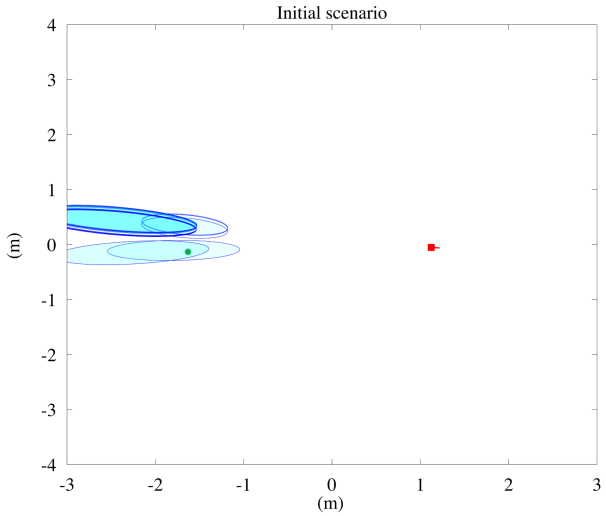

Fig. 12 Initial position of the robot and the source and estimated belief before running motion planning strategies. Blue ellipses represent 95% confidence regions of various hypotheses in the mixture with a transparency proportional to the weight of the components. The robot position is shown as a red square with a red line indicating the direction of the microphone array, and the actual source position is shown as a green circle.

The initial example scenario is depicted in Fig. 12. In this situation, the front-back ambiguity still exists in the source location belief. The distance from the center of the microphone array to the true robot position is around 3 m. At this point, the microphone array is in the endfire position, i.e., the microphones are arranged in line with the source position. In this endfire position, the localization accuracy will be lower compared to the

broadside position in which the microphones are arranged in a perpendicular line to the direction of sound propagation.

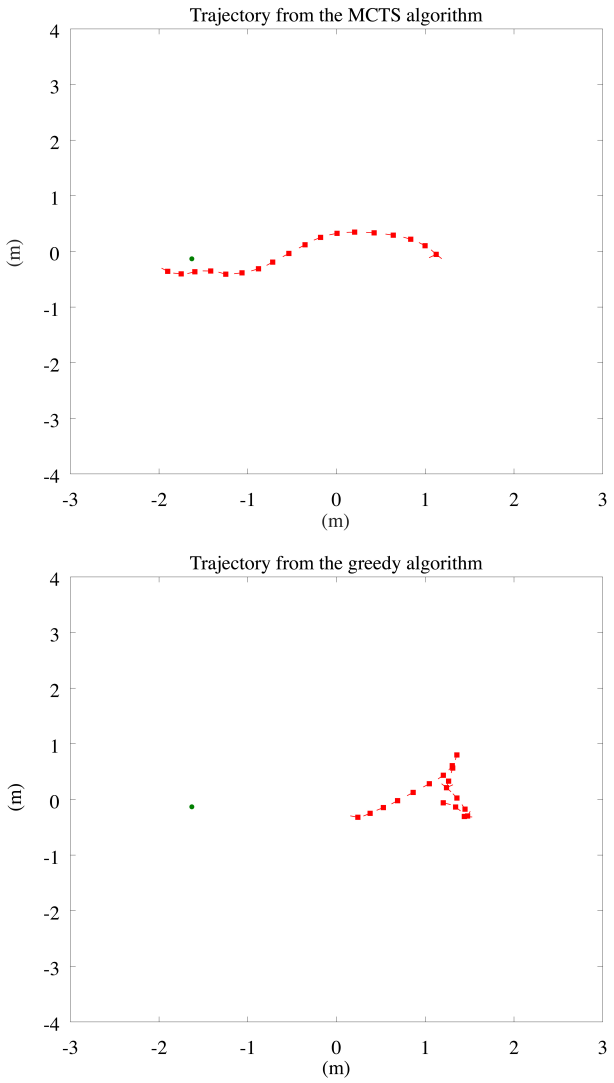

Fig. 13 Top: example robot motion sequence obtained from the MCTS algorithm with the entropy criterion. Bottom: example robot motion sequence obtained from the greedy algorithm with the entropy criterion. Robot positions are shown as red squares with a red line indicating the direction of the microphone array, and the actual source position is shown as a green circle.

From this initial robot-source position and estimate, we run separately the MCTS algorithm and the greedy algorithm with the entropy criterion to find the subsequent optimal robot trajectory. The resulting robot trajectories are shown in Fig. 13. In the robot trajectory obtained from the MCTS method, the robot first rotates around its position. This rotation does not minimize the entropy of the estimated belief in the short term because the microphones are still in endfire position. However, it helps to eliminate the front-back confusion in the estimated belief. After the rotation, the robot moves towards the source in order to increase the SNR which results in more accurate AoA estimation. Actually, from the figure, we see that the robot does not move in a straight line towards the source but it makes a slight turn when approaching it. Gradually, the microphone array rotates to the broadside position, a.k.a the auditory fovea, where the localization accuracy is higher.

In contrast, the robot trajectory obtained with the greedy method shows that the robot makes a turn down first then goes up and ends up closer to the source. The purpose of moving down or moving up is to put the microphone array in the auditory fovea and minimize the entropy of the estimated belief. However, the robot is still far from the source and the SNR does not change, or even decreases. After that, the robot decides to move closer to the source in the final steps.

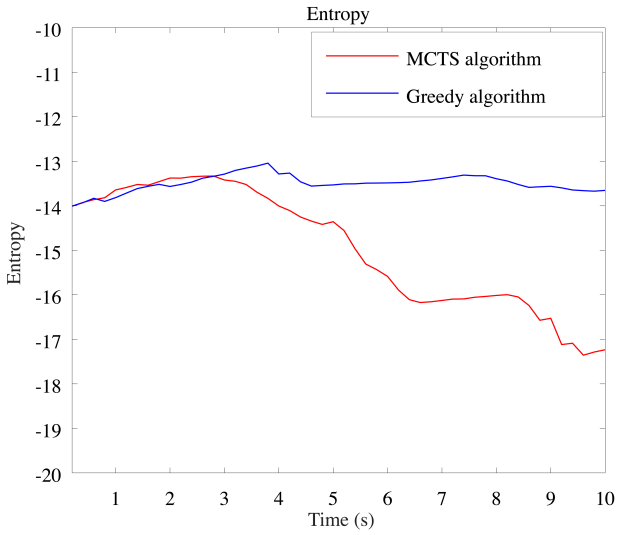

Fig. 14 Entropy over time for the MCTS algorithm and the greedy algorithm with the entropy criterion.

Fig. 16 a) and Fig. 16 b) show the estimated source location belief after 10 s when the robot follows the trajectory from the MCTS and greedy algorithm respectively. From these two figures, we can clearly see that the robot motion from the MCTS algorithm leads to better localization compared to the robot motion from the greedy algorithm, although they started from the same initial position. As a result, the accuracy of the localization result is higher and the estimated source location belief represented by the blue ellipses is smaller.

Fig. 14 shows the entropy value of the estimated belief over time for both the MCTS and the greedy

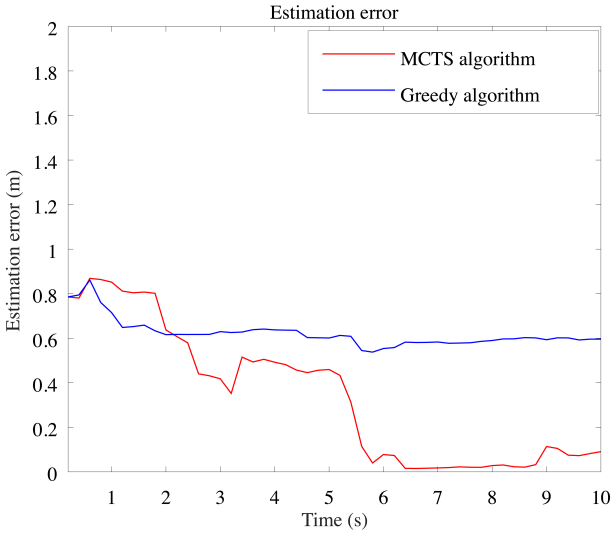

Fig. 15 Estimation error over time for the MCTS algorithm and the greedy algorithm with the entropy criterion.

method. During the first 3 s, the greedy method yields a lower entropy compared to MCTS. This result is due to the greedy movement made to minimize the entropy in the short term. However the entropy does not change much over time afterwards. By contrast, from 3 s to 6.5 s, the entropy value obtained by MCTS decreases drastically. This corresponds to the time when the robot gets closer to the source. The entropy value continues to decrease after 8.5 s when the robot makes a turn to put the microphone array in broadside position.

Fig. 15 presents the estimation error over time for both the MCTS and the greedy method, that is the distance between the estimated source position and the true position. After 2 s, it is not surprising that the estimation error of MCTS decreases and becomes much lower compared to the greedy method.

To summarize this example scenario, the robot in the greedy method turns sideways in the first few steps to immediately reduce the entropy, however, by doing that it does not yield the optimal trajectory in the long term. In contrast, with the long-term motion planning method using MCTS, at the beginning, the rotation to the direction of the source may not result in an optimal value. However, this is the first step to prepare for the later actions that move closer to the source. In the long run, we can obtain a better robot trajectory.

In the following, we compare the performance of the MCTS method with other motion planning methods in the long run and with two different criteria: entropy and standard deviation. In addition, we analyze the effect of the discount factor on the performance of the MCTS.

5.3.4 MCTS vs other motion planning approaches

We compare the performance of the MCTS approach with other motion planning approaches for two criteria: entropy and standard deviation. We use a discount factor $\gamma = 1$.

- *Entropy criterion*

Fig. 17 a) shows the entropy of the estimated belief over time for all algorithms, on average over all experiments. The entropy values of all methods decrease until the time $k = 1.2$ s. From that to time $k = 2$ s, they rise up. This is due to the silent interval in the sound source. However, after this period, when the source is active again and we have better information from AoA measurements, the entropy of all methods decreases over time except for the random algorithm. So, after the silent interval, the random motion approach cannot help to improve the performance.

From time $k = 4$ s, the two flavors of MCTS already yield significantly lower entropy compared to the greedy algorithm and the random algorithm. The entropy decreases drastically at further time steps. This is because MCTS optimizes the entropy in the long run. In contrast to that, random motion barely reduces it and the greedy algorithm only optimizes it one time step ahead.

This result is expected as our method is actually optimizing the entropy. However, the objective of this work is to use robots to better localize sound sources. Therefore the evaluation of the algorithms should be done according to the estimation error. The average estimation error over time is presented for all algorithms in Fig. 18 a). Again, the average estimation error of both MCTS methods is smaller compared to greedy and random motion in the long run, although the estimation error from the greedy algorithm is smaller but not significantly so in the first four seconds compared to both MCTS methods. During the silent interval of the source, the estimation error of all methods does not change, because the MKF method that we use for estimating the source location deals well with this intermittent source.

With a Wilcoxon signed-rank test, we can show that both MCTS variants yield a significantly smaller entropy and estimation error than the greedy and random algorithms ($p < 0.01$). The two MCTS methods are not significantly different in terms of both entropy and estimation error.

- *Standard deviation criterion*

Fig. 17 b) shows the standard deviation of the estimated belief over time for all algorithms, on average

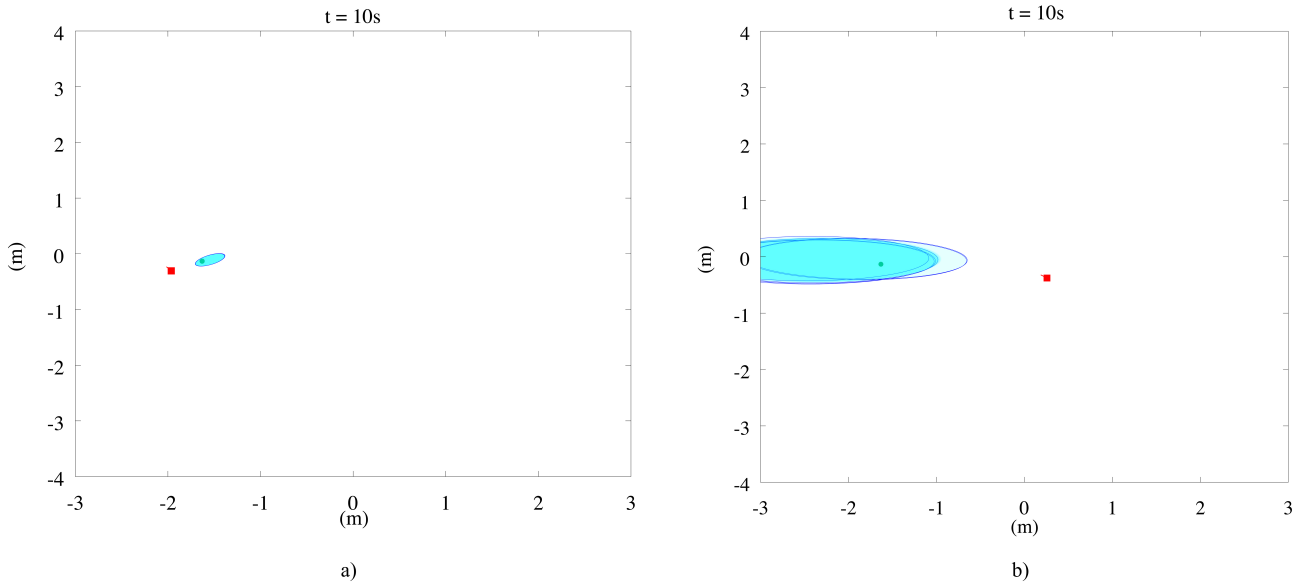

Fig. 16 a) Estimated source location belief after 10 s when the robot follows the trajectory from the MCTS algorithm with the entropy criterion. b) Estimated source location belief after 10 s when the robot follows the trajectory from the greedy algorithm with the entropy criterion. Blue ellipses represent 95% confidence regions of various hypotheses in the mixture with a transparency proportional to the weight of the components. Robot positions are shown as red squares, and the actual source position as a green circle.

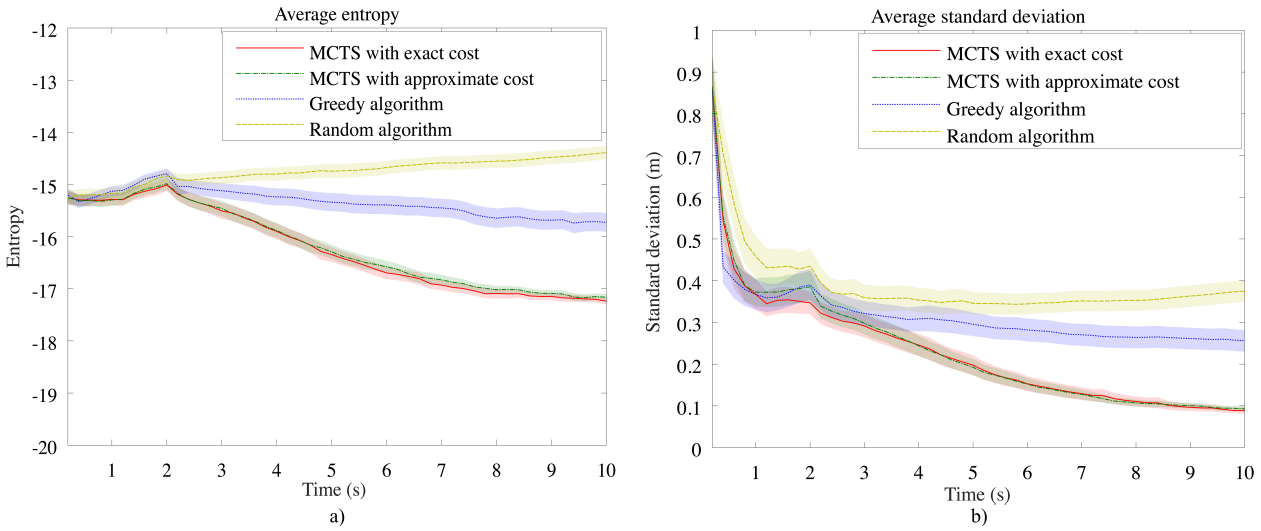

Fig. 17 a) Average entropy and 95% confidence interval over time of the 4 algorithms with the entropy criterion over all 200 experiments. b) Average standard deviation and 95% confidence interval over time with the standard deviation criterion of the 4 algorithms over all 200 experiments.

over all experiments. The average standard deviation for all methods falls drastically until $k = 1.2$ s. There is a slight rise in the standard deviation for all methods during the time when the source is inactive (from $k = 1.2$ to 2 s). Similarly with the entropy criterion, after time $k = 4$ s, the two MCTS approaches yield significantly lower standard deviation compared to the greedy algorithm and the random algorithm. The standard deviation decreases drastically at further time steps.

The average estimation error over time with the standard deviation criterion is presented for all algorithms in Fig. 18 b). In the long run, the average estimation error of both MCTS methods is smaller compared to greedy and random motion. In addition, the estimation error of all methods does not change during the silent interval of the source when we still use the MKF method for estimation.

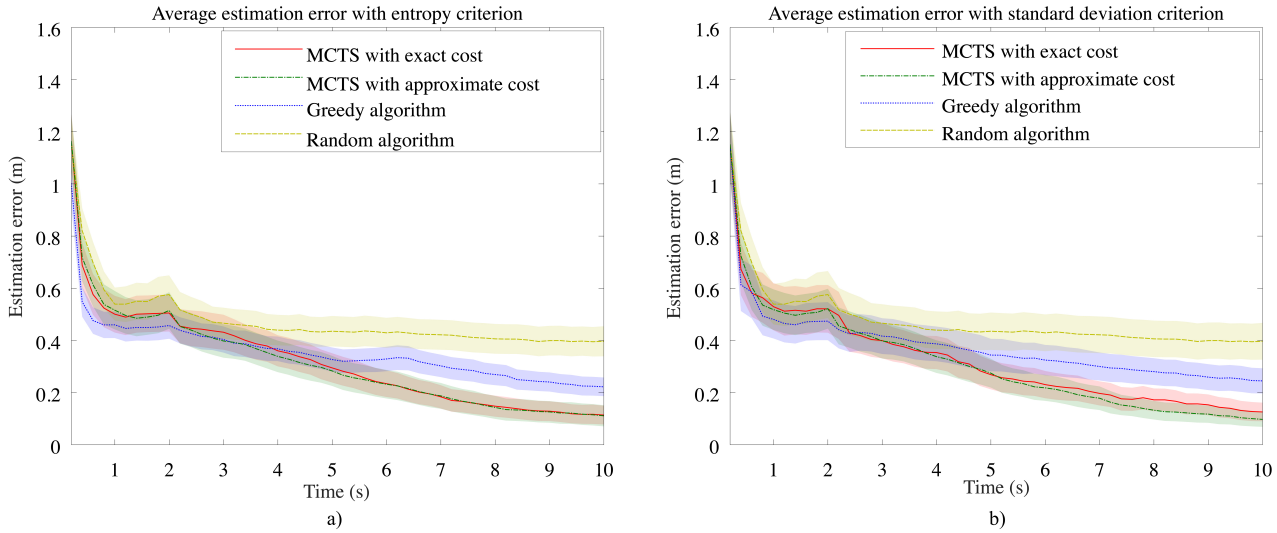

Fig. 18 a) Average estimation error and 95% confidence interval over time of the 4 algorithms with the entropy criterion over all 200 experiments. b) Average estimation error and 95% confidence interval over time of the 4 algorithms with the standard deviation criterion over all 200 experiments.

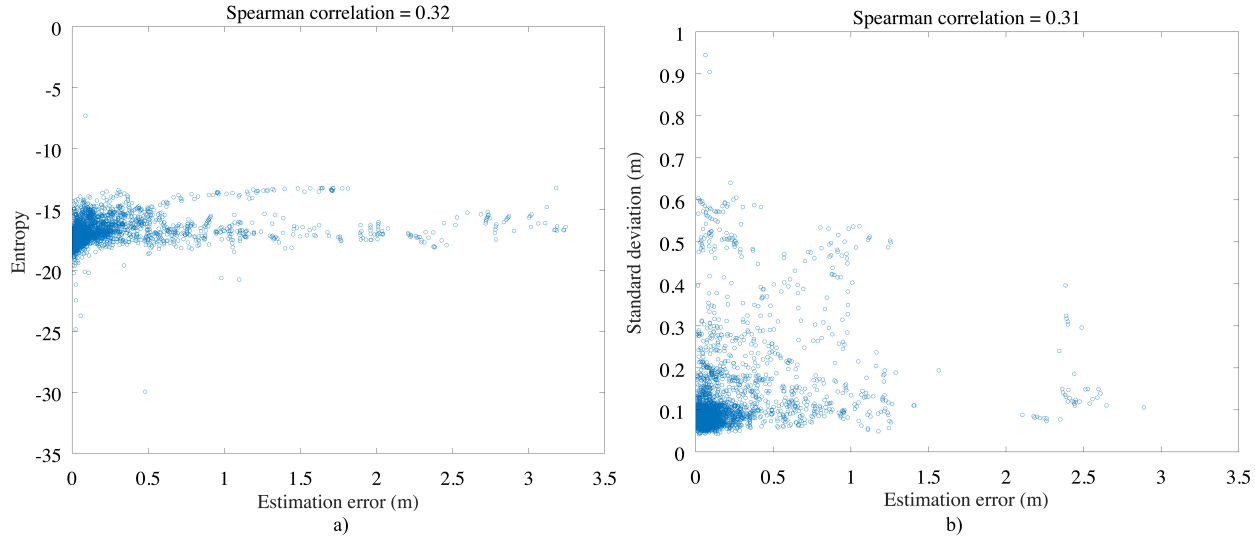

Fig. 19 a) Correlation between the entropy and the estimation error. b) Correlation between the standard deviation and the estimation error.

With a Wilcoxon signed-rank test, we can show that both MCTS approaches yield a significantly smaller entropy and estimation error than the greedy and random algorithms ($p < 0.01$). As with the entropy criterion, the two MCTS approaches with standard deviation criterion are not significantly different from each other in terms of both standard deviation and estimation error. Therefore, the MCTS with the approximate cost can be an interesting alternative for the MCTS with exact cost as it can have faster computational time.

5.3.5 Relation of both criteria with estimation error

We investigate the relation between the entropy criterion and the estimation error as well as between the standard deviation criterion and the estimation error in the MCTS approach. We plot the entropy as a function of the estimation error during the 4 s interval from $k = 6$ s to $k = 10$ s. Similarly, we plot the standard deviation as a function of the estimation error in the same interval of time.

In Fig. 19 a) and b), most of the smaller values of entropy or standard deviation correspond to smaller values of estimation error and most of the larger values of entropy or standard deviation correspond to larger val-

ues of estimation error. There are some outliers whereby small values of entropy or standard deviation correspond to large values of estimation error however the number of such outliers is small.

To assess the relationship between each of the two criteria and the estimation error, we compute Spearman's rank correlation coefficient for each of them. The test shows that the Spearman correlation between the standard deviation and the estimation error is approximately equal (Spearman correlation = 0.31) to the Spearman correlation between the entropy and the estimation error (Spearman correlation = 0.32). From this assessment and the two figures above we can conclude that both the standard deviation and the entropy can be a good criterion for motion planning.

5.3.6 Effect of the discount factor

We also analyze the impact of different discount factor values in the entropy and the standard deviation criteria. Six discount factor values are tested. Values lower than 1 correspond to emphasis on short-term reward and are typical for Markov Decision Processes where infinite future reward is computed. A value of 1 corresponds to an equal contribution of all time steps. Finally, values higher than 1 correspond to emphasis on longer-term reward with a limit, when it grows, to consider only the last time step, as is done in [10]. For each discount factor value, we run 200 experiments for different random initial robot locations and source locations.

- *Entropy criterion*

Fig. 20 a) shows the average entropy over time with different discount factor values. From 2 s to 10 s, we see a significant difference between the entropy value of the two discount factors $\gamma > 1$ and the others. For the discount factors which are equal or smaller than 1, the entropy value decreases drastically after 2 s. For the discount factors that are greater than 1, the entropy value decreases more slowly.

We can see those decreases in Fig. 21 a) which shows the average estimation error over all experiments with different discount factors. However, there are not significant differences in the speed of decrease of the estimation error for different discount factor values.

Fig. 22 a) presents the error bars which show the average estimation error and 95% confidence interval for different discount factors at 1 s and 10 s with the entropy criterion. This figure is extracted from Fig. 21 a), with a closer look at two time steps: 1 s and 10 s. The error bars for all discount factors at 10 s are smaller compared to those at 1 s. This is as desired with the MCTS approach. We can see that at 10 s the estimation error with a discount factor greater than 1 is slightly larger than with a discount factor equal to or smaller than 1.

- *Standard deviation criterion*

Fig. 20 b) shows the average standard deviation over time and Fig. 21 b) shows the corresponding average estimation error over time in all experiments with different discount factors. We can see a similar shape between the two figures. After 2 s, both the values of standard deviation and estimation error decrease. With a discount factor greater than 1, both the standard deviation and the estimation error decrease faster compared with a discount factor equal to or smaller than 1.

Fig. 22 b) depicts the error bars which show the average estimation error and 95% confidence interval for different discount factors at 1 s and 10 s with the standard deviation criterion. This figure is extracted from Fig. 21 b), with a closer look at two time steps: 1 s and 10 s. At 1 s, all the discount factors have similar error bars. At 10 s, we have smaller error bars and lower error values compared to those at 1 s. In contrast to the entropy criterion, we see that discount factors smaller than 1 result in higher error than discount factors equal to or larger than 1.

6 Conclusions

We first presented an extended MKF framework for tracking a single intermittent, moving source. From the statistical result of the experiments, we showed that by jointly estimating the source location and its activity, the extended MKF framework is more robust to false SAD and AoA measurements. The localization framework can work well even when the false AoA measurement is in the range of $\pm 20^\circ$ or the false SAD rate is 10%.

With the above localization framework, we proposed a long-term robot motion planning algorithm to find an optimal robot motion to minimize the uncertainty in the source localization. Our main theoretical contributions concern the cost function with two criteria: the standard deviation and the entropy of the estimated belief. In addition, we adapted a practical MCTS algorithm for finding an optimal sequence of robot movements that will minimize the estimation uncertainty in the long run. The experiments showed that our long-term planning algorithm achieves better performance compared to greedy or random motion. In average, the estimation error in our method converges to a 20 cm smaller value compared to the greedy method with corresponds to a 48.7% reduction. We also evaluated the

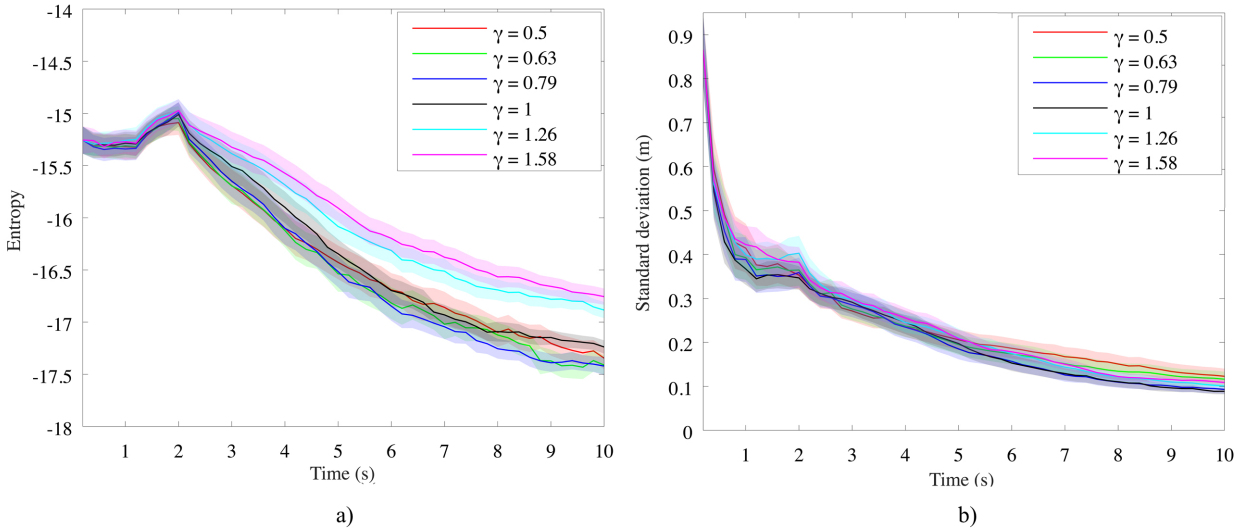

Fig. 20 a) Average entropy and 95% confidence interval over time with different discount factor values for the entropy criterion. b) Average standard deviation and 95% confidence interval over time with different discount factor values for the standard deviation criterion.

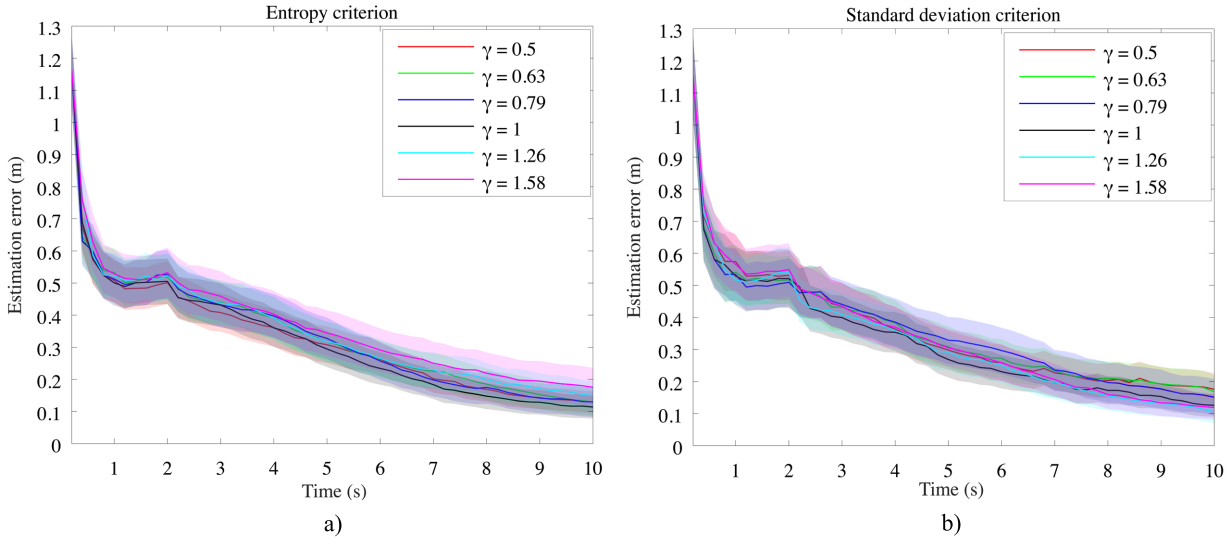

Fig. 21 a) Average estimation error and 95% confidence interval over time with different discount factor values for the entropy criterion. b) Average estimation error and 95% confidence interval over time with different discount factor values for the standard deviation criterion.

performance of the MCTS approach with different criteria and different discount factors. The analysis of all the results showed a coherent estimation error result when optimizing the standard deviation or the entropy of the estimated belief.

Future work will focus on improving the selection and simulation step in MCTS by providing prior knowledge when building the tree. Moreover, we can extend the motion planning algorithm using MCTS to the context of multiple sources. Assessing the robustness of the extended MKF to imprecise knowledge of the reverberation time and the room geometry will be also considered in future work.

Appendix A Chernoff-Hoeffding inequality

The UCT criterion derives from the Chernoff-Hoeffding inequality which is valid for a bounded reward function [26]. The Chernoff-Hoeffding inequality is stated in the theorem below.

Theorem: Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables whose values are within the range $[a, b]$. Denote $\mu_i = \mathbb{E}(Y_i)$ as their expected values, $Y = \frac{1}{n}\sum_i Y_i$ and $\mu = \mathbb{E}(Y) = \frac{1}{n}\sum_i \mu_i$. Then for all $\epsilon > 0$, we have:

$$P(|Y - \mu| > \epsilon) \leq 2e^{-2n\epsilon^2/(b-a)^2}. \quad (41)$$

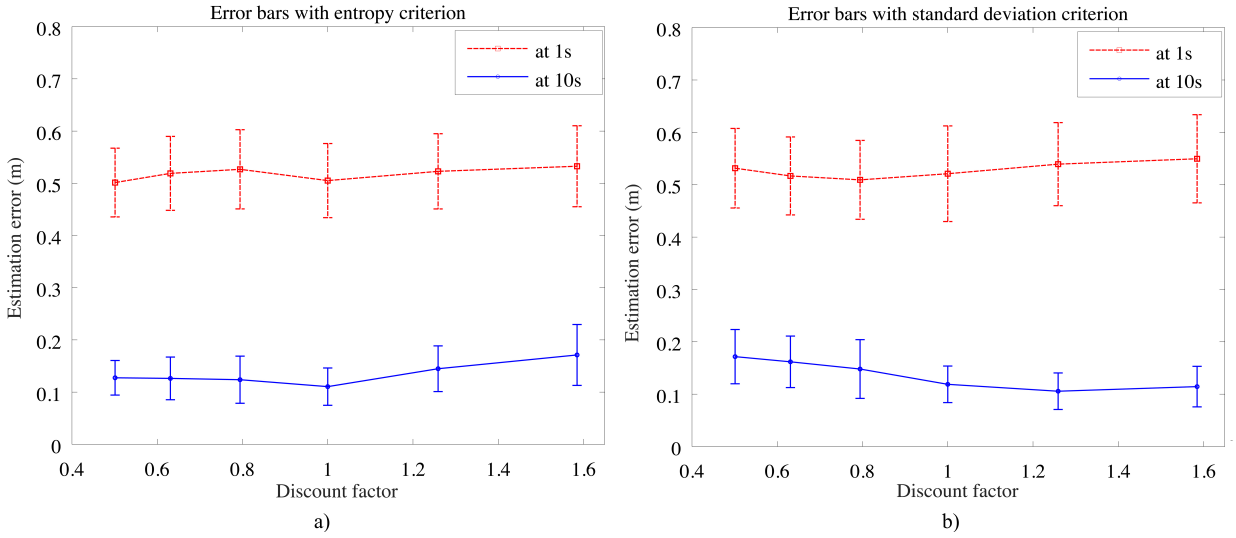

Fig. 22 a) Error bars with the entropy criterion at $t = 1$ s and $t = 10$ s with different discount factor values. b) Error bars with the standard deviation criterion at $t = 1$ s and $t = 10$ s with different discount factor values.

For us, Y_i is the reward Q at the end of each simulation and Y is the average reward $\frac{\bar{Q}(n')}{N(n')}$ at each child node in the tree. So, to satisfy the Chernoff-Hoeffding inequality, the two criteria: entropy and standard deviation must be bounded. We show that these two criteria are bounded in the following subsections.

A.1 Bounded entropy

For a scalar random variable X in the range $[a, b]$ with no other constraints, the maximum entropy distribution of X is the uniform distribution over this range. In that case the formula for calculating the maximum entropy is expressed as follows:

$$\begin{aligned} H_{\max,\mathrm{X}} &= -\int_a^b p(X)\log p(X)dX \\ &= -\log p(X)\int_a^b p(X)dX \\ &= -\log p(X) \\ &= -\log\frac{1}{(b-a)} \\ &= \log(b-a) \\ &= \log|\mathrm{range}_X| \end{aligned} \tag{42}$$

The formula for computing the maximum entropy of the belief is:

$$\begin{aligned} H_{\max} = &\log\left|\mathrm{range}_{x_\mathrm{r}}\right| + \log\left|\mathrm{range}_{y_\mathrm{r}}\right| + \log\left|\mathrm{range}_{\theta_\mathrm{r}}\right| + \\ &\log\left|\mathrm{range}_{x_\mathrm{s}}\right| + \log\left|\mathrm{range}_{y_\mathrm{s}}\right| + \log\left|\mathrm{range}_{\theta_\mathrm{s}}\right| + \\ &\log\left|\mathrm{range}_{v_\mathrm{s}}\right| + \log\left|\mathrm{range}_{w_\mathrm{s}}\right| \\ = &\ 20.3027. \end{aligned} \tag{43}$$

In theory, the minimum entropy with perfect knowledge is $-\infty$ but it is not achievable in practice. So the lower bound can be computed as follows. In order to find the minimum entropy of the estimated belief, we begin the belief propagation with perfect knowledge about the source position. In the nonlinear MKF, it is represented by one hypothesis of active source whose variance for the source position is equal to 0. After the prediction step, there will be one hypothesis of active source with a higher weight and one hypothesis of inactive source with a lower weight. The uncertainty will appear due to the process noise Q in the dynamic model. We then evaluate the uncertainty of estimation of the source location after the update step. We find the angle from the robot to the source, the distance from the robot to the source in the range from 0.18 m to 8 m, and the AoA observation such that the entropy of the belief after the update step above is minimum. As a result, the minimum entropy $H_{\min}$ is -38.7824 obtained for an AoA of 176°, which does not suffer from the front-back ambiguity, and a distance of 0.18 m from the robot to the source.

So, the entropy is bounded upwards by the entropy of the uniform distribution on the state vector, and downwards by the entropy of the probability distribution in the case when there is no front-back ambiguity and the sound source is as close as possible to the robot (0.18 m due to the size of the robot).

A.2 Bounded standard deviation

Let a and b be respectively the lower and upper bounds on the values of any random variable with a particular probability distribution. Then, according to Popoviciu's inequality on variances [50], its variance satisfies:

$$\sigma^2 \leq \frac{1}{4}(b-a)^2, \tag{44}$$

or its standard deviation is bounded as follows:

$$-\frac{1}{2}(b-a) \leq \sigma \leq \frac{1}{2}(b-a). \tag{45}$$

Acknowledgements Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (see `https://www.grid5000.fr`).

References

1. Alam, J., Kenny, P., Ouellet, P., Stafylakis, T., Dumouchel, P.: Supervised/unsupervised voice activity detectors for text-dependent speaker recognition on the RSR2015 corpus. In: Proc. Odyssey (2014)
2. Ali, A.M., Asgari, S., Collier, T.C., Allen, M., Girod, L., Hudson, R.E., Yao, K., Taylor, C.E., Blumstein, D.T.: An empirical study of collaborative acoustic source localization. Journal of Signal Processing Systems **57**(3), 415–436 (2009)
3. Allen, J.B., Berkley, D.A.: Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am. **65**(4), 943–950 (1979)
4. Amanatiadis, A.A., Chatzichristofis, S.A., Charalampous, K., Doitsidis, L., Kosmatopoulos, E.B., Tsalides, P., Gasteratos, A., Roumeliotis, S.I.: A multiobjective exploration strategy for mobile robots under operational constraints. IEEE Access **1**, 691–702 (2013)
5. Badali, A., Valin, J.M., Michaud, F., Aarabi, P.: Evaluating real-time audio localization algorithms for artificial audition in robotics. In: Proc. IROS, pp. 2033–2038 (2009)
6. Berglund, E., Sitte, J.: Sound source localisation through active audition. In: Proc. IROS, pp. 509–514 (2005)
7. Bhattacharyya, S.: Motion Planning and Constraint Exploration for Robotic Surgery. Vanderbilt University (2011)
8. Blandin, C., Ozerov, A., Vincent, E.: Multi-source TDOA estimation in reverberant audio using angular spectra and clustering. Signal Processing **92**(8), 1950–1960 (2012)
9. Browne, C.B., Powley, E., Whitehouse, D., Lucas, S.M., Cowling, P.I., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A survey of Monte Carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in Games **4**(1), 1–43 (2012)
10. Bustamante, G., Danès, P.: Multi-step-ahead information-based feedback control for active binaural localization. In: Proc. IROS (2017)
11. Bustamante, G., Danès, P., Forgue, T., Podlubne, A.: Towards information-based feedback control for binaural active localization. In: Proc. ICASSP, pp. 6325–6329 (2016)
12. Bustamante, G., Danès, P., Forgue, T., Podlubne, A., Manhès, J.: An information based feedback control for audio-motor binaural localization. Autonomous Robots (2017). DOI 10.1007/s10514-017-9639-8
13. Chengalvarayan, R.: Robust energy normalization using speech/nonspeech discriminator for German connected digit recognition. In: Proc. Eurospeech (1999)
14. Colas, F., Mahesh, S., Pomerleau, F., Liu, M., Siegwart, R.: 3D path planning and execution for search and rescue ground robots. In: Proc. IROS, pp. 722–727 (2013)
15. Cooke, M., Lu, Y.C., Lu, Y., Horaud, R.: Active hearing, active speaking. In: Proc. ISAAR, pp. 33–46 (2007)
16. DeJong, B.P.: Auditory occupancy grids with a mobile robot. Journal of Automation, Mobile Robotics & Intelligent Systems **6**(3), 3–12 (2012)
17. DiBiase, J.H., Silverman, H.F., Brandstein, M.S.: Robust localisation in reverberant rooms. In: Microphone Arrays: Signal Processing Techniques and Applications, pp. 157–180. Springer (2001)
18. Dolgov, D., Thrun, S., Montemerlo, M., Diebel, J.: Practical search techniques in path planning for autonomous driving. In: Proc. STAIR (2008)
19. Evers, C., Moore, A., Naylor, P.: Towards informative path planning for acoustic SLAM. In: Proc. DAGA (2016)
20. Fallon, M.F., Godsill, S.J.: Acoustic source localization and tracking of a time-varying number of speakers. IEEE Transactions on Audio, Speech, and Language Processing **20**(4), 1409–1415 (2012)
21. Germain, F.G., Sun, D.L., Mysore, G.J.: Speaker and noise independent voice activity detection. In: Proc. Interspeech (2013)
22. Girod, L., Lukac, M., Trifa, V., Estrin, D.: The design and implementation of a self-calibrating distributed acoustic sensing platform. In: Proc. SenSys, pp. 71–84 (2006)
23. Gonzalez-Banos, H.H., Latombe, J.C.: Navigation strategies for exploring indoor environments. The International Journal of Robotics Research **21**(10-11), 829–848 (2002)
24. Hahn, W., Tretter, S.: Optimum processing for delay-vector estimation in passive signal arrays. IEEE Transactions on Information Theory **19**(5), 608–614 (1973)
25. Hashimoto, S., Narita, S., Kasahara, H., Takanishi, A., Sugano, S., Shirai, K., Kobayashi, T., Takanobu, H., Kurata, T., Fujiwara, K., Matsuno, T., Kawasaki, T., Hoashi, K.: Humanoid robot-development of an information assistant robot hadaly. In: Proc. RO-MAN, pp. 106–111 (1997)
26. Hoeffding, W.: Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association **58**(301), 13–30 (1963)
27. Huber, M.F., Bailey, T., Durrant-Whyte, H., Hanebeck, U.D.: On entropy approximation for Gaussian mixture random vectors. In: Proc. MFI, pp. 181–188 (2008)
28. Johnson, D.H., Dudgeon, D.E.: Array Signal Processing: Concepts and Techniques. Simon & Schuster (1992)
29. Karray, L., Martin, A.: Towards improving speech detection robustness for speech recognition in adverse conditions. Speech Communication **40**(3), 261–276 (2003)
30. Kim, U.H., Kim, J., Kim, D., Kim, H., You, B.J.: Speaker localization using the TDOA-based feature matrix for a humanoid robot. In: Proc. RO-MAN, pp. 610–615 (2008)
31. Knapp, C., Carter, G.: The generalized cross-correlation method for estimation of time delay. IEEE Transactions

on Acoustics, Speech, and Signal Processing **24**(4), 320–327 (1976)
32. Kocsis, L., Szepesvári, C., Willemson, J.: Improved Monte-Carlo search. Tech. Rep. 1, University of Tartu (2006)
33. Latombe, J.C.: Robot Motion Planning. Kluwer (1991)
34. LaValle, S.M.: Planning algorithms. Cambridge University Press (2006)
35. Lu, Y.C., Cooke, M.: Motion strategies for binaural localisation of speech sources in azimuth and distance by artificial listeners. Speech Communication **53**(5), 622–642 (2011)
36. Magassouba, A.: Aural servo: towards an alternative approach to sound localization for robot motion control. Ph.D. thesis, Université Rennes 1 (2016)
37. Marković, I., Portello, A., Danès, P., Petrović, I., Argentieri, S.: Active speaker localization with circular likelihoods and bootstrap filtering. In: Proc. IROS, pp. 2914–2920 (2013)
38. Martinson, E., Schultz, A.: Auditory evidence grids. In: Proc. IROS, pp. 1139–1144 (2006)
39. Martinson, E., Schultz, A.: Discovery of sound sources by an autonomous mobile robot. Autonomous Robots **27**, 221–237 (2009)
40. Marzinzik, M., Kollmeier, B.: Speech pause detection for noise spectrum estimation by tracking power envelope dynamics. IEEE Transactions on Speech and Audio Processing **10**(2), 109–118 (2002)
41. Nakadai, K., Lourens, T., Okuno, H.G., Kitano, H.: Active audition for humanoid. In: Proc. AAAI, pp. 832–839 (2000)
42. Nakadai, K., Okuno, H.G., Kitano, H.: Real-time sound source localization and separation for robot audition. In: Proc. Interspeech, pp. 193–196 (2002)
43. Nakadai, K., Okuno, H.G., Kitano, H.: Robot recognizes three simultaneous speech by active audition. In: Proc. ICRA, pp. 398–405 (2003)
44. Nakadai, K., Takahashi, T., Okuno, H.G., Nakajima, H., Hasegawa, Y., Tsujino, H.: Design and implementation of robot audition system 'HARK' — open source software for listening to three simultaneous speakers. Advanced Robotics **24**(5–6), 739–761 (2010)
45. Nakamura, K., Nakadai, K., Ince, G.: Real-time super-resolution sound source localization for robots. In: Proc. IROS, pp. 694–699 (2012)
46. Nguyen, Q.V.: Mapping of a sound environment by a mobile robot. Ph.D. thesis, University of Lorraine (2018)
47. Nguyen, Q.V., Colas, F., Vincent, E., Charpillet, F.: Localizing an intermittent and moving sound source using a mobile robot. In: Proc. IROS, pp. 61–65 (2016)
48. Nguyen, Q.V., Colas, F., Vincent, E., Charpillet, F.: Long-term robot motion planning for active sound source localization with Monte Carlo tree search. In: Proc. HSCMA, pp. 61–65 (2017)
49. Okuno, H.G., Nakadai, K.: Robot audition: Its rise and perspectives. In: Proc. ICASSP, pp. 5610–5614 (2015)
50. Popoviciu, T.: Sur les équations algébriques ayant toutes leurs racines réelles. Mathematica (Cluj) (1935)
51. Portello, A., Bustamante, G., Danès, P., Piat, J., Manhès, J.: Active localization of an intermittent sound source from a moving binaural sensor. In: Proc. Forum Acusticum (2014)
52. Portello, A., Danès, P., Argentieri, S.: Acoustic models and Kalman filtering strategies for active binaural sound localization. In: Proc. IROS, pp. 137–142 (2011)
53. Portello, A., Danès, P., Argentieri, S.: Active binaural localization of intermittent moving sources in the presence of false measurements. In: Proc. IROS, pp. 3294–3299 (2012)
54. Ramírez, J., Górriz, J.M., Segura, J.C.: Voice activity detection. Fundamentals and speech recognition system robustness. In: Robust Speech Recognition and Understanding (2007)
55. Ramirez, J., Segura, J.C., Benitez, C., de la Torre, A., Rubio, A.J.: A new adaptive long-term spectral estimation voice activity detector. In: Proc. Eurospeech (2003)
56. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Transactions on Antennas and Propagation **34**(3), 276–280 (1986)
57. Schymura, C., Grajales, J.D.R., Kolossa, D.: Monte Carlo exploration for active binaural localization. In: Proc. ICASSP, pp. 491–495 (2017)
58. Siegwart, R., Nourbakhsh, I.R., Scaramuzza, D.: Introduction to autonomous mobile robots. MIT Press (2011)
59. Slotani, M.: Tolerance regions for a multivariate normal population. Annals of the Institute of Statistical Mathematics **16**(1), 135–153 (1964)
60. Sohn, J., Kim, N.S., Sung, W.: A statistical model-based voice activity detection. IEEE Signal Processing Letters **6**(1), 1–3 (1999)
61. Song, K., Liu, Q., Wang, Q.: Olfaction and hearing based mobile robot navigation for odor/sound source search. Sensors **11**, 2129–2154 (2011)
62. Tanyer, S.G., Ozer, H.: Voice activity detection in nonstationary noise. IEEE Transactions on Speech and Audio processing **8**(4), 478–482 (2000)
63. Valin, J.M., Michaud, F., Rouat, J.: Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. Robotics and Autonomous Systems **55**(3), 216–228 (2007)
64. Valin, J.M., Yamamoto, S., Rouat, J., Michaud, F., Nakadai, K., Okuno, H.: Robust recognition of simultaneous speech by a mobile robot. IEEE Transactions on Robotics **23**(4), 742–752 (2007)
65. Van Veen, B.D., Buckley, K.M.: Beamforming: A versatile approach to spatial filtering. IEEE ASSP Magazine **5**(2), 4–24 (1988)
66. Vermaak, J., Blake, A.: Nonlinear filtering for speaker tracking in noisy and reverberant environments. In: Proc. ICASSP, vol. 5, pp. 3021–3024 (2001)
67. Vincent, E., Sini, A., Charpillet, F.: Audio source localization by optimal control of a mobile robot. In: Proc. ICASSP, pp. 5630–5634 (2015)
68. Wightman, F.L., Kistler, D.J.: Resolution of front–back ambiguity in spatial hearing by listener and source movement. The Journal of the Acoustical Society of America **105**(5), 2841–2853 (1999)
69. Woo, K.H., Yang, T.Y., Park, K.J., Lee, C.: Robust voice activity detection algorithm for estimating noise spectrum. IET Electronics Letters **36**(2), 180–181 (2000)
70. Yamauchi, B.: A frontier-based approach for autonomous exploration. In: Proc. CIRA, pp. 146–151 (1997)
71. Zhang, X.L., Wu, J.: Deep belief networks based voice activity detection. IEEE Transactions on Audio, Speech, and Language Processing **21**(4), 697–710 (2013)