

RESEARCH

Open Access



# The hybrid Cramér-Rao lower bound for simultaneous self-localization and room geometry estimation

Maya Veisman, Yair Noam and Sharon Gannot\*

\*Correspondence:

[sharon.gannot@biu.ac.il](mailto:sharon.gannot@biu.ac.il)

Sharon Gannot is a Member of  
EURASIP  
Bar-Ilan University, Faculty of  
Engineering, 5290002 Ramat Gan,  
Israel

## Abstract

This paper addresses the problem of tracking a moving source, e.g., a robot, equipped with both receivers and a source, that is tracking its own location and simultaneously estimating the locations of multiple plane reflectors. We assume a noisy knowledge of the robot's movement. We formulate this problem, which is also known as simultaneous localization and mapping (SLAM), as a hybrid estimation problem. We derive the extended Kalman filter (EKF) for both tracking the robot's own location and estimating the room geometry. Since the EKF employs linearization at every step, we incorporate a regulated kinematic model, which facilitates a successful tracking. In addition, we consider the echo-labeling problem as solved and beyond the scope of this paper. We then develop the hybrid Cramér-Rao lower bound on the estimation accuracy of both the localization and mapping parameters. The algorithm is evaluated with respect to the bound via simulations, which shows that the EKF approaches the hybrid Cramér-Rao bound (CRB) (HCRB) as the number of observation increases. This result implies that for the examples tested in simulation, the HCRB is an asymptotically tight bound and that the EKF is an optimal estimator. Whether this property is true in general remains an open question.

**Keywords:** SLAM, Speaker localization and tracking, Room mapping, Hybrid Cramér-Rao lower bound

## 1 Introduction

Consider a source that moves in an unknown area while simultaneously mapping the environment and tracking its own location, a problem known as SLAM. As a concrete example, we use a moving robot that emits sound in a room enclosure. Nevertheless, the analysis here applies to other similar scenarios, e.g., drone navigation [1], vehicles in urban areas [2], or marine vessels [3, 4] in ports. This problem is difficult since the robot must localize itself with respect to an incomplete map. The general SLAM problem has been the subject of substantial research with many different applications [5].

As implied by its name, SLAM consists of two components, localization and mapping. Source localization is a well-studied problem in acoustic signal processing. Several energy-based localization techniques exist [6, 7]. A common approach utilizes an array

of microphones and their TDOA [8, 9]. Another technique exploits both the spatial and temporal information to localize and track a moving speaker [10]. Additional approaches and techniques are reviewed in [11]. In the case of mapping, a general solution for room geometry estimation is based on spatial room-impulse-responses and the image source model [12]. In addition, echo labeling is used to map the shape of polyhedral rooms [13, 14]. The latter treat the acoustic echoes as nodes in a graph.

While the mapping and the localization sub-problems can be considered well understood when separately treated, the combined SLAM problem is much more challenging. The reason is that localization and mapping depend on each other and are executed simultaneously in an intertwined way. The most common philosophy in addressing SLAM is the Bayesian theoretical-based approach [5]. The reader is referred to the review [15] and references therein for further study on the subject. Notable is the work in [16], which employs particle filtering and uses a moving microphone-array. Kalman-based approaches are in wide use for the tracking and mapping. These approaches provide a recursive solution to the localization problem and consistent estimates for mapping landmarks. This is done using statistical models for the speaker motion and the environment's features [17]. A Fast-SLAM algorithm that integrates the Rao-Blackwell particle filter and the Kalman filter is presented in [18].

We consider a robot that emits an acoustic signal and is equipped with one microphone. While moving in a four-wall room, the robot tracks its own location and simultaneously estimates the room geometry (namely, mapping) using the acoustic echoes. For this problem, some solutions exist. Notable is [19] which estimates the shape of a room as well as determines the location and geometry of an array of microphones using Euclidean distance and matrix completion algorithms. In [20], a Bayesian filtering method is applied for movement tracking and room shape estimation. Most algorithms are based on particle filtering, which is computationally intense and may suffer from convergence problems.

The paper is organized as follows. In Section 2, we summarize the main methodology of the paper and briefly describe the experimental study. The problem is formulated in Section 3, which includes the robot's kinematic model, the observation model, and the unknown parameters. In Section 4, we present an online SLAM algorithm, and in Section 5, we derive the HCRB (cf. Proposition 1). Simulation results of the proposed algorithm with respect to the derived HCRB are presented in Section 6. Summary and conclusions are given in Section 7.

## 2 Methods

In this paper, we address the problem of self-localizing a source and, simultaneously, inferring the room geometry using a single microphone and a single receiver that are mounted on a moving platform, e.g., a robot. The robot moves around the room according to given movement commands. The estimation is based on noisy distance measurements, corresponding to the first-order echoes, taken at finite set of points along the trajectory. While the movement commands are known, the overall robot trajectory is unknown. This follows from the fact that room geometry is unknown (and therefore, the robot position with respect to the room). In addition, we assume that the robot movement is inaccurate w.r.t. the known movement commands. This can occur due to imperfect execution of the robot commands. In general, if nothing is known about the trajectory, this problem is not identifiable for some important classes of room geometries, i.e., parallelogram room-geometry

cannot be uniquely determined from the distances between walls and trajectory points (cf. [21, 22], and for a comprehensive analysis, see [23]). Nevertheless, the robot movement commands provide partial information about the robot trajectory. We use these movement commands, as well as the distance measurements from each wall, to simultaneously map the room and localize the robot with respect to it. In contrast to particle filtering, we adopt a more analytic, computationally efficient approach, the EKF, which simultaneously tracks the robot moment and estimates the room geometry.

We formulate the problem as a hybrid estimation task that includes both random and deterministic unknown parameters. This combines both the classic and the Bayesian estimation philosophies, an approach that is very suitable for the problem at hand in which the room's features are considered static and deterministic, while the robot movement as a random state-space process. To apply the Bayesian EKF approach to our hybrid tracking-mapping estimation problem, we carefully incorporate randomness into the deterministic room parameters (cf. Section 4).

Moreover, we derive the HCRB, which is an important tool in estimation theory [24–29], for the SLAM problem at hand. We provide a recursive expression for the HCRB (see [30] for an overview of recursive CRB-type bounds) which constitutes, at each time instance, a lower bound on the estimation error of the current location and of the room geometry. The advantage of the HCRB is that it is valid for both random and deterministic parameters. A related bound, the Bayesian CRB (BCRB), was derived for a different SLAM problem in [21]. While the BCRB is a valid bound on the random parameters<sup>1</sup> (e.g., receiver location), it is generally not useful with respect to the deterministic parameters in hybrid setups. To apply the BCRB, one must transform the hybrid estimation problem into a pure Bayesian one. This can be done by incorporating some prior distribution to the deterministic parameters, which makes the BCRB susceptible to mismodelling. It is also possible to treat the deterministic parameters as random with zero variance. However, this will violate the BCRB regularity conditions [30]. A different approach would be to discard the prior distribution on the random parameters, thus transforming the problem into a pure classic one, i.e., an estimation problem where all unknown parameters are deterministic unknown. The drawback is that by discarding the prior distribution, one loses the useful information that is embedded in the prior. Moreover, it may render the estimation problem ill-posed (see, e.g., [24]). In this paper, we consider the HCRB since it is more suitable for hybrid estimation problems (see the discussion about the HCRB and other CRB-type bounds in [25], Sect. I-A and Sect. III).

Finally, we evaluate the performance of the proposed algorithm by simulation results examination and by comparing them with the obtained HCRB. We demonstrate a very good match between the Monte-Carlo simulation and the analytical bound, i.e., we show in simulation that EKF approaches the HCRB.

*Notation:* Boldface lowercase letters denote vectors. We use boldface uppercase letters, e.g.,  $\mathbf{F}$ ,  $\mathbf{C}$ ,  $\mathbf{J}$ , to denote matrices. All vectors are column vectors and  $(\cdot)^\top$  denotes the transpose operation. Let  $g(\mathbf{x})$  be scalar function of a vector  $\mathbf{x}$ , then,  $\partial g(\mathbf{x})/\partial \mathbf{x}$  denotes the gradient operation which is organized as a row vector. In the case where  $\mathbf{g}(\mathbf{x})$  is a vector,  $\partial \mathbf{g}(\mathbf{x})/\partial \mathbf{x}$  is the Jacobian matrix of  $\mathbf{g}(\mathbf{x})$ . Let  $\mathbf{A}$  be a matrix, then,  $[\mathbf{A}]_{ij}$  denotes its  $i, j$ th entry. The operator spectral norm operation is denoted by  $\|\cdot\|$ . Let  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times m}$ ,

<sup>1</sup>The BCRB, unlike its classic counterpart the CRB, is in general, not an asymptotically tight bound.

then  $\text{blockdiag}(\mathbf{A}, \mathbf{B})$  is the block-diagonal matrix whose upper-left and lower-right  $n \times m$  blocks are equal to  $\mathbf{A}$  and  $\mathbf{B}$ , respectively.

### 3 Problem formulation

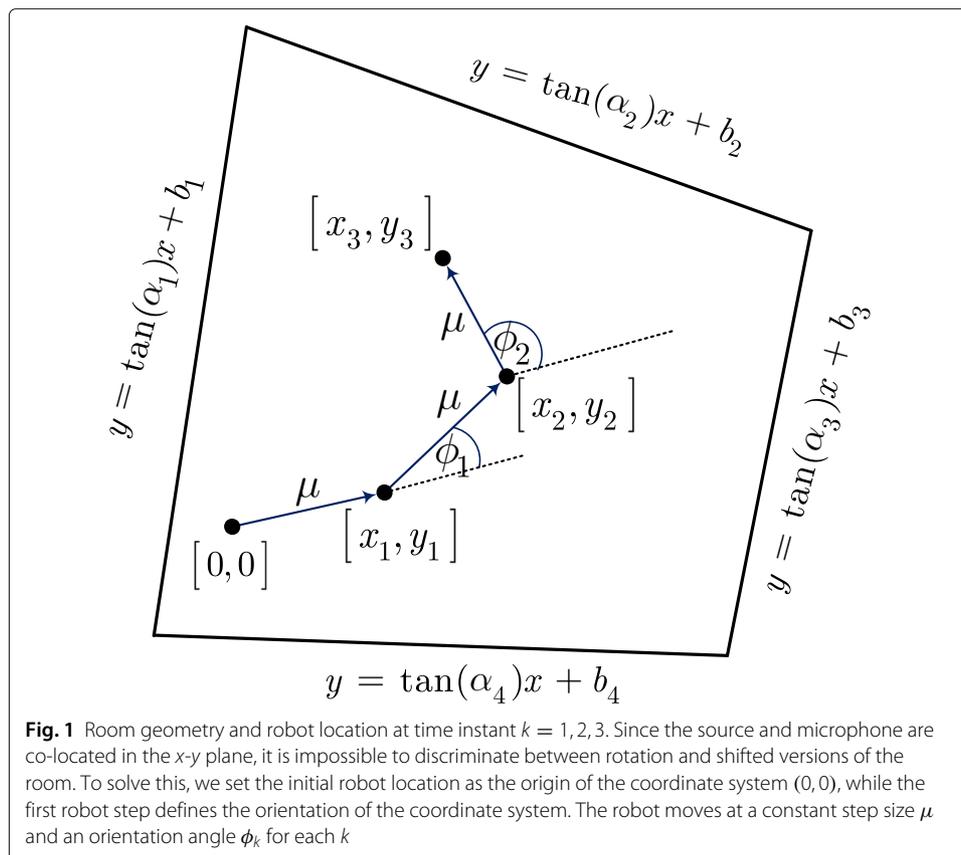
Consider a single omni-directional acoustic source and a microphone, both mounted on a robot at the same point on the  $x$ - $y$  plane, but at different heights, i.e., with different  $z$ -coordinates. A fully-exciting signal that simplifies the subsequent correlation operation would be better than a speech signal for echo-labeling.

The robot moves in a four-wall, closed-room according to a series of movement commands. In this work, we consider 2D environments. An extension to 3D environments is possible, and it involves a characterization of the third dimension. Such an extension can be made by adding characterization equations for additional walls.

Each command instructs the robot to take a single step of a given size and an orientation angle that indicate the direction and length of the movement as depicted in Fig. 1 and further discussed in the sequel. In each step, the robot takes a measurement (a full description is given in Section 3.2) and simultaneously estimates its location and the room geometry. The robot's location is denoted by

$$\mathbf{x}_k = [x_k, y_k]^T \tag{1}$$

where  $x_k$  and  $y_k$  denote the  $x$ -coordinate and the  $y$ -coordinate, respectively, at step  $k$ . The coordinate system is determined by the robot's first step, i.e., the positive direction of the  $x$ -axis coincides with the first step. In this paper, each time instant corresponds to a



step; thus, we use the terms time-instant  $k$  and step  $k$ , interchangeably. The room shape is determined by the parameter vector

$$\theta = [\alpha_1, b_1, \alpha_2, b_2, \alpha_3, b_3, \alpha_4, b_4]^T \tag{2}$$

where  $(\alpha_i, b_i)$  represent the angle and offset of the  $i$ th wall, as depicted in Fig. 2.

### 3.1 The robot kinematic model

We assume that the robot travels according to the following linear model

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{u}_k + \mathbf{w}_k, \tag{3}$$

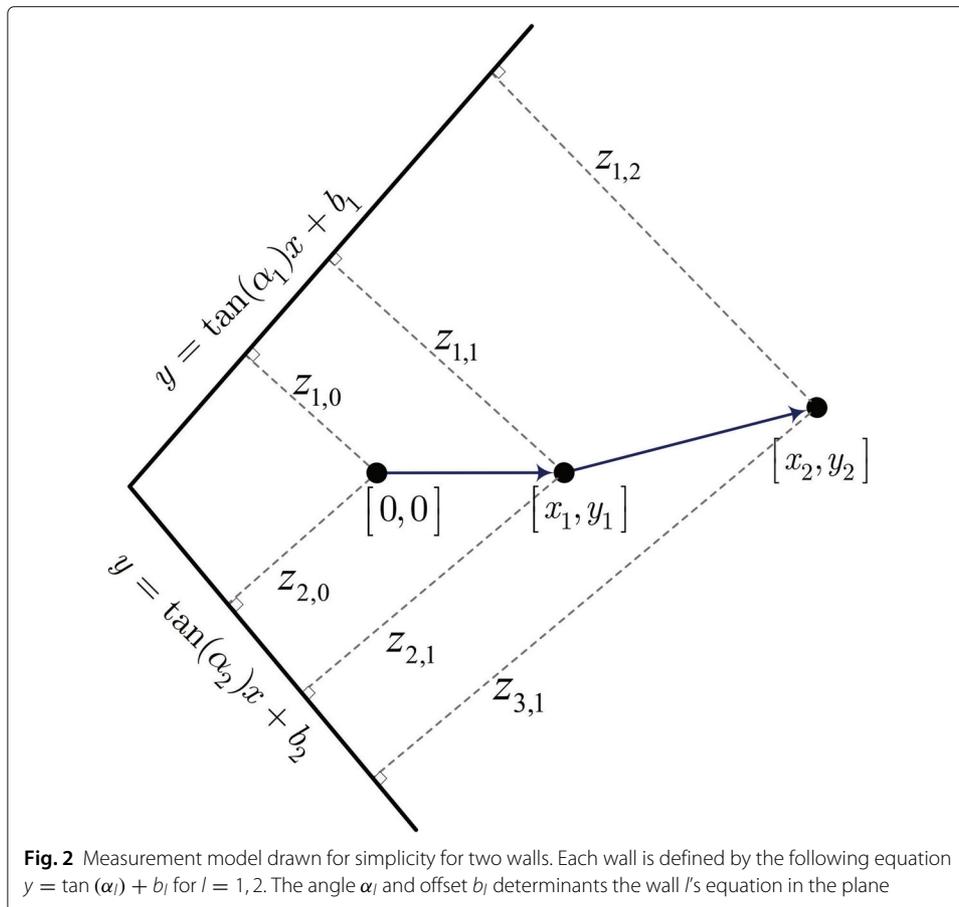
where  $k$  is a time index,  $\mathbf{F}$  is the state-transition matrix, given by

$$\mathbf{F} = \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \rho \in (0, 1) \tag{4}$$

and

$$\mathbf{u}_k = [\mu_k \sin(\phi_k), \mu_k \cos(\phi_k)]^T. \tag{5}$$

The latter denotes the robot's movement command, where  $\mu_k$  is the step size and  $\phi_k$  is the orientation angle with respect to the positive edge of the horizontal axis, i.e., the angle of the robot first step is set as the zero-angle with respect to the  $x$ -axis of the coordinate system. The robot movement is depicted in Fig. 1. The noise  $\mathbf{w}_k$  is zero-mean Gaussian



signal with covariance matrix  $\mathbf{C}_w = \sigma_w^2 \mathbf{I}_{2 \times 2}$ , which expresses inaccuracies in the execution of the robot’s movement commands. As a result, the exact location of the robot is an unknown parameter, to be estimated as part of the estimation problem at hand.

Few comments are in place regarding the matrix  $\mathbf{F}$ . If it had not been for mathematical regularity conditions, the robot location would have better been characterized by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{u}_k + \mathbf{w}_k; \tag{6}$$

In words, the location at time  $k + 1$  would have been determined by the previous location and by the step taken at time  $k$ . This, however, induces an unstable kinematic model, where the noise variance is accumulated and diverges to infinity as  $k$  increases. While such models are generally tractable using the Kalman filter [31, 32], it may be problematic for the EKF [33] due to its embedded linearization. If the robot movement is not regulated, the linear approximation may be too crude and induce an unstable kinematic model that diverges the filter. Therefore, we incorporate the matrix  $\mathbf{F}$ , whose sole purpose is to regulate the robot movement by enforcing stability on the model. It can be shown that the model is stable for  $\rho < 1$ . However, to make this model realistic,  $\rho$  should be close to “1” and accordingly,  $\mathbf{F}$  to be close to the identity matrix.

We conclude with defining the evolution of the unknown deterministic parameter  $\theta$ . Since the room features (wall angles and offsets) are constant, we use the following static propagation model:

$$\theta_{k+1} = \theta_k. \tag{7}$$

### 3.2 Measurement model

The observation at each time instant,  $k$ , is a vector  $\mathbf{z}_k = [z_{1,k} \cdots z_{4,k}]^T$  which measures the distance between the robot and each of the four walls, as shown in Fig. 2. It can be verified by straightforward geometrical considerations that the distance is related to the robot’s location and to the room’s features by the following nonlinear equation:

$$\begin{aligned} \mathbf{z}_k &= \mathbf{h}(\mathbf{x}_k, \theta) + \mathbf{v}_k \\ &= \begin{bmatrix} (b_1 - y_k) \cos \alpha_1 + x_k \sin \alpha_1 \\ (b_2 - y_k) \cos \alpha_2 + x_k \sin \alpha_2 \\ (-b_3 + y_k) \cos \alpha_3 - x_k \sin \alpha_3 \\ (-b_4 + y_k) \cos \alpha_4 - x_k \sin \alpha_4 \end{bmatrix} + \begin{bmatrix} v_{1,k} \\ v_{2,k} \\ v_{3,k} \\ v_{4,k} \end{bmatrix}. \end{aligned} \tag{8}$$

Note that  $z_{i,k} \forall i \in 1, \dots, 4$  is an estimate of the distance between the robot and wall  $i$  at time  $k$ , as we explain in the sequel. Moreover,  $\mathbf{h}(\mathbf{x}_k, \theta)$  is a nonlinear function that expresses the true distance of the robot from each wall in terms of the robot true coordinates. The vector  $\mathbf{v}_k$  is a white Gaussian noise with covariance matrix  $\mathbf{C}_v = \sigma_v^2 \mathbf{I}_{4 \times 4}$ , which depicts the measurement error induced by an imperfect distance estimation. The Gaussian assumption is reasonable in cases where the measurement error is small enough compared to the actual distance [10].

**Definition 1** *The SLAM problem, considered in this paper, is to jointly estimate  $\theta$  and  $\mathbf{x}_k$ , for each  $k$ , based on the state-space models given in (3), (7) and the measurement model (8).*

From a practical point of view, the measurement model in (8) needs some prior processing. The TOA which corresponds to the distance between the robot and a wall [34, 35] must be estimated.

Let  $\tau_{ik}$  be the signal propagation time from the robot to the wall and back to the robot. The latter is related to the distance by

$$d_{ik} = \frac{\tau_{ik}c}{2} \tag{9}$$

where  $c$  is the speed of sound (approximately 340 m/s in air), and  $d_{ik}$  is the distance between the robot and wall  $i$  at time  $k$ . Furthermore, in addition to estimating  $\tau_{i,k}$ , it is necessary to identify which echo corresponds to each wall. This problem is known as echo labeling. The reader is referred to [14, 20, 36, 37] for further insights. Because echo labeling is beyond the scope of this paper we assume, for simplicity, that the echoes are perfectly labeled<sup>2</sup>. Finally, we would like to stress that the loudspeaker and the microphone are co-located only in their 2D-projection into the  $x$ - $y$  plain. In Section 6.2, we place them in different heights, i.e., the source-receiver location differs only in their  $z$ -coordinate, such that they remain co-located in the  $x$ - $y$  plane.

#### 4 Simultaneous localization and mapping: algorithm

This section presents a recursive algorithm based on the EKF [5] for the SLAM problem given in Definition 1. Note that this is not a Bayesian estimation problem, but rather a hybrid one. To estimate the deterministic parameter  $\theta_k$  in (7) using the EKF, which is Bayesian, we incorporate randomness into  $\theta_k$ . This, in addition, enables a more “lively” tracking results [38] that cannot be achieved by the algorithm initialization (see, e.g., [5]). Accordingly, we modify (7) as follows

$$\theta_{k+1} = \theta_k + \mathbf{m}_k \tag{10}$$

where  $\mathbf{m}_k$  is white Gaussian noise with covariance matrix  $\mathbf{R}_k$ . The introduction of  $\mathbf{m}_k$  in (10), as well as the initialization  $\hat{\theta}_0$ , incorporate an inaccurate prior probabilistic knowledge about  $\theta$ . To avoid mismodelling, and guarantee that  $\hat{\theta}_0$  is forgotten rapidly, the matrix  $\mathbf{R}_k$  must be “large” at  $k = 0$  and  $\|\mathbf{R}_k\|$  should decrease to zero as  $k$  increases.

$$\mathbf{R}_k = \sigma_0^2 \Lambda^k \tag{11}$$

where  $\sigma_0^2$  is large and  $\Lambda$  is the following diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}. \tag{12}$$

The parameters  $\lambda_1, \lambda_2$  act as decay coefficients that correspond to  $\{\alpha_i\}_{i=1}^4$  and  $\{b_i\}_{i=1}^4$ , respectively, and determine the rate at which the corresponding driving noise decreases. After substituting  $\theta$  with  $\theta_k$  in (7), we are left with the nonlinear state-space model given

<sup>2</sup>A possible scenario where this assumption might hold is if, during the algorithm initialization, the source remains in the same Voronoi region (a region in which the set of distances from the source to the walls maintains its order). Once the algorithm converges, it is possible to identify with a low probability of error, which Voronoi region the source is in, from which it is possible to label the echoes.

by (3), (8), and (10), where the nonlinearity is due to (8). Similar to [5], it is possible to develop the EKF, which linearizes the measurement model and transition matrices [31, 39]. The algorithm consists of two steps for each iteration: prediction and update. In the prediction stage, the robot location is predicted according to (3) as follows

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{u}_k \tag{13}$$

and the room parameter-vector is predicted according to (10) as follows:

$$\hat{\theta}_{k|k-1} = \hat{\theta}_{k-1|k-1}. \tag{14}$$

The joint covariance matrix of the localization and mapping parameters is

$$\mathbf{P}_{k|k-1} = \mathbf{T} \begin{bmatrix} \mathbf{P}_{k-1|k-1}^{\mathbf{xx}} & \mathbf{P}_{k-1|k-1}^{\mathbf{x}\theta} \\ \mathbf{P}_{k-1|k-1}^{\theta\mathbf{x}} & \mathbf{P}_{k-1|k-1}^{\theta\theta} \end{bmatrix} \mathbf{T}^\top + \begin{bmatrix} \mathbf{C}_w & 0 \\ 0 & \mathbf{R}_k \end{bmatrix} \tag{15}$$

where

$$\mathbf{T} = \begin{bmatrix} \mathbf{F} & 0 \\ 0 & \mathbf{I} \end{bmatrix}. \tag{16}$$

As discussed above,  $\mathbf{C}_w$  is the covariance matrix w.r.t. the localization parameters and  $\mathbf{R}_k$  is the randomness inserted into the mapping parameters.

In the update stage, the algorithm estimates the location and the room parameters as

$$\hat{\Phi}_{k|k} = \hat{\Phi}_{k|k-1} + \mathbf{K}_k \left( \mathbf{z}_k - \mathbf{h} \left( \hat{\Phi}_{k|k-1} \right) \right) \tag{17}$$

where

$$\hat{\Phi}_{k|k} = \left[ \hat{\mathbf{x}}_{k|k}, \hat{\theta}_{k|k} \right]^\top, \tag{18}$$

and  $\mathbf{K}_k$  is the Kalman gain, which is given by

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \nabla \mathbf{h}^\top \mathbf{S}_k^{-1}. \tag{19}$$

The matrix  $\mathbf{S}_k$  is calculated as

$$\mathbf{S}_k = \nabla \mathbf{h} \mathbf{P}_{k|k-1} \nabla \mathbf{h}^\top + \mathbf{C}_v \tag{20}$$

where  $\nabla \mathbf{h}$  denotes the Jacobian of  $\mathbf{h}$  at  $\hat{\mathbf{x}}_{k|k-1}$  and  $\hat{\theta}_{k|k-1}$ . In addition, the covariance matrix is updated considering the EKF extension [5]

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^\top. \tag{21}$$

The complete EKF algorithm is summarized in Algorithm 1.

### 5 The hybrid Cramér-Rao lower bound

In this section, we derive the HCRB for the SLAM problem of Definition 1. The objective is to bound the estimation error of  $\mathbf{x}_k$  and  $\theta$  from the measurements  $\mathbf{z}_k$  for each step  $k$ , while treating the previous locations as a nuisance parameter. We denote the overall unknown parameter vector at time  $k$  by  $\psi_k = [\mathbf{x}_0^\top, \dots, \mathbf{x}_k^\top, \theta^\top]^\top$  and collect the set of all robot's locations up to time  $k$  into  $\mathbf{x}^k = [\mathbf{x}_0^\top, \mathbf{x}_1^\top, \dots, \mathbf{x}_k^\top]^\top$ . Similarly,  $\mathbf{z}^k = [\mathbf{z}_0^\top, \mathbf{z}_1^\top, \dots, \mathbf{z}_k^\top]^\top$  comprises all the measurements up to time  $k$ . Since  $\mathbf{x}_k$  is random (3) while  $\theta$  is deterministic, the estimation of  $\psi_k$  is a hybrid estimation problem, i.e., a combination of classic and

---

**Algorithm 1** EKF estimation

---

**Initialization:**

$$\begin{aligned} \hat{\mathbf{x}}_{0|0} &= [0, 0] \\ \hat{\boldsymbol{\theta}}_{0|0} &= \hat{\boldsymbol{\theta}}_0 \\ \mathbf{P}_{0|0}^{\mathbf{xx}} &= \mathbf{P}_0^{\mathbf{xx}} \\ \mathbf{P}_{0|0}^{\boldsymbol{\theta}\boldsymbol{\theta}} &= \mathbf{P}_0^{\boldsymbol{\theta}\boldsymbol{\theta}} \end{aligned}$$

**Predict:**

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \mathbf{F}\hat{\mathbf{x}}_{k-1|k-1} \\ \hat{\boldsymbol{\theta}}_{k|k-1} &= \hat{\boldsymbol{\theta}}_{k-1|k-1} \\ \mathbf{P}_{k|k-1} &= \mathbf{T} \begin{bmatrix} \mathbf{P}_{k-1|k-1}^{\mathbf{xx}} & \mathbf{P}_{k-1|k-1}^{\mathbf{x}\boldsymbol{\theta}} \\ \mathbf{P}_{k-1|k-1}^{\boldsymbol{\theta}\mathbf{x}} & \mathbf{P}_{k-1|k-1}^{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix} \mathbf{T}^\top \\ &\quad + \begin{bmatrix} \mathbf{C}_w & 0 \\ 0 & \mathbf{R}_k \end{bmatrix} \\ \mathbf{T} &= \begin{bmatrix} \mathbf{F} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \end{aligned}$$

**Update:**

$$\begin{aligned} \mathbf{S}_k &= \nabla \mathbf{h} \mathbf{P}_{k|k-1} \nabla \mathbf{h}^\top + \mathbf{C}_v \\ \mathbf{K}_k &= \mathbf{P}_{k|k-1} \nabla \mathbf{h}^\top \mathbf{S}_k^{-1} \\ \hat{\boldsymbol{\Phi}}_{k|k} &= \hat{\boldsymbol{\Phi}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{h}(\hat{\boldsymbol{\Phi}}_{k|k-1})) \\ \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^\top \end{aligned}$$


---

Bayesian estimation philosophies [26, 30, 40]. The HCRB lower bound [24] asserts that for every unbiased estimator  $\hat{\boldsymbol{\psi}}_k(\mathbf{z}_k)$ , i.e.,

$$\mathbb{E} \left\{ \boldsymbol{\psi}_k - \hat{\boldsymbol{\psi}}_k \right\} = 0 \tag{22}$$

the estimation error is lower bounded as follows:

$$\mathbb{E} \left\{ (\hat{\boldsymbol{\psi}}_k - \boldsymbol{\psi}_k) (\hat{\boldsymbol{\psi}}_k - \boldsymbol{\psi}_k)^\top \right\} \geq \mathbf{J}_k^{-1}, \tag{23}$$

where the expectation  $\mathbb{E}$  is taken w.r.t. the joint probability density function (p.d.f.) of  $(\mathbf{z}^k, \mathbf{x}^k)$ , i.e.  $p(\mathbf{z}^k, \mathbf{x}^k; \boldsymbol{\theta})$ , where we used semicolon to distinguish between the deterministic vector  $\boldsymbol{\theta}$ , and the random vectors  $\mathbf{z}^k$  and  $\mathbf{x}^k$ . Note that  $\hat{\boldsymbol{\psi}}_k(\mathbf{z}_k)$  is a

function of the measurements  $\mathbf{z}_k$ . We omit the explicit dependency for brevity. The matrix

$$\mathbf{J}_k = -\mathbb{E} \left\{ \frac{\partial^2 \ln p(\mathbf{z}^k, \mathbf{x}^k; \boldsymbol{\theta})}{\partial (\mathbf{x}^k; \boldsymbol{\theta}) \partial (\mathbf{x}^k; \boldsymbol{\theta})^\top} \right\} \quad (24)$$

is known as the hybrid Fisher information matrix (HFIM). In our case, the dimension of the full HFIM,  $\mathbf{J}_k$ , is  $\mathbb{R}^{k+8 \times k+8}$ . Since we focus on performance evaluation of online algorithms, for each current time step  $k$ , we are interested in  $[\mathbf{J}_k^{-1}]^{\mathbf{x}_k, \boldsymbol{\theta}}$ , which is the lower-right  $10 \times 10$  block of  $\mathbf{J}_k^{-1}$  that corresponds to the sub-vector  $[\mathbf{x}_k^\top, \boldsymbol{\theta}^\top]^\top$  of  $\boldsymbol{\psi}_k$ . The following proposition presents a recursive expression for  $[\mathbf{J}_k^{-1}]^{\mathbf{x}_k, \boldsymbol{\theta}}$  as a function of  $[\mathbf{J}_{k-1}^{-1}]^{\mathbf{x}_{k-1}, \boldsymbol{\theta}}$ .

The following comment about assumption (22) is in place. Consider first the deterministic parameters; we note that in the finite sample regime, there are useful estimators that do not satisfy (22), e.g., the maximum likelihood (ML) in the pure classic case and the joint ML/MAP in the hybrid case (see [25]). Nevertheless, in the large-sample regime, an estimator that does not satisfy (22) concerning the deterministic static parameters is not consistent. Since inconsistent estimators are useless in the large-sample regime for these parameters, the assumption in (22) is not restricting.<sup>3</sup> Now to the random parameters. If there is a bias, it must be a known-function of the deterministic parameters. Thus, if the estimate of the deterministic parameters is consistent, the bias becomes known in the large-sample regime, and therefore, could and should be subtracted from the estimator. This operation yields an asymptotically unbiased estimate.

Before presenting the bound, the following definitions are required. Let  $[\mathbf{J}_k^{-1}]^{\mathbf{x}_k, \boldsymbol{\theta}}$  be the lower-right  $10 \times 10$  block of  $\mathbf{J}_k^{-1}$  defined in (24) and let

$$\tilde{\mathbf{J}}_k = \begin{bmatrix} \tilde{\mathbf{J}}_k^{\mathbf{x}_k \mathbf{x}_k} & \tilde{\mathbf{J}}_k^{\mathbf{x}_k \boldsymbol{\theta}} \\ \tilde{\mathbf{J}}_k^{\boldsymbol{\theta} \mathbf{x}_k} & \tilde{\mathbf{J}}_k^{\boldsymbol{\theta} \boldsymbol{\theta}} \end{bmatrix} \triangleq \left[ [\mathbf{J}_k^{-1}]^{\mathbf{x}_k, \boldsymbol{\theta}} \right]^{-1}. \quad (25)$$

To calculate the HCRB, we will use the following recursive formula for  $\tilde{\mathbf{J}}_k^{\mathbf{x}_k, \boldsymbol{\theta}}$ , that was derived [41] for a general state-space model with additive Gaussian noise

$$\begin{aligned} \tilde{\mathbf{J}}_k^{\mathbf{x}_k \mathbf{x}_k} &= \mathbf{D}_{k-1}^{22} - (\mathbf{D}_{k-1}^{12})^\top \\ &\quad \times \left( \mathbf{D}_{k-1}^{11} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \mathbf{x}_{k-1}} \right)^{-1} \mathbf{D}_{k-1}^{12} \end{aligned} \quad (26a)$$

$$\begin{aligned} \tilde{\mathbf{J}}_k^{\mathbf{x}_k \boldsymbol{\theta}} &= \mathbf{D}_{k-1}^{23} - (\mathbf{D}_{k-1}^{12})^\top \left( \mathbf{D}_{k-1}^{11} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \mathbf{x}_{k-1}} \right)^{-1} \\ &\quad \times \left( \mathbf{D}_{k-1}^{13} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \boldsymbol{\theta}} \right) \end{aligned} \quad (26b)$$

$$\begin{aligned} \tilde{\mathbf{J}}_k^{\boldsymbol{\theta} \boldsymbol{\theta}} &= \mathbf{D}_{k-1}^{33} + \tilde{\mathbf{J}}_{k-1}^{\boldsymbol{\theta} \boldsymbol{\theta}} - \left( \mathbf{D}_{k-1}^{13} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \boldsymbol{\theta}} \right)^\top \\ &\quad \times \left( \mathbf{D}_{k-1}^{11} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \mathbf{x}_{k-1}} \right)^{-1} \\ &\quad \times \left( \mathbf{D}_{k-1}^{13} + \tilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1} \boldsymbol{\theta}} \right) \end{aligned} \quad (26c)$$

$$\tilde{\mathbf{J}}_k^{\boldsymbol{\theta} \mathbf{x}_k} = \left( \tilde{\mathbf{J}}_k^{\mathbf{x}_k \boldsymbol{\theta}} \right)^\top \quad (26d)$$

<sup>3</sup>For further reading on different types of unbiased conditions, the reader is referred to [25].

where  $\{\mathbf{D}_{k-1}^{ij}\}_{i,j=1}^3$  are defined by

$$\mathbf{D}_{k-1}^{11} = -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \mathbf{x}_{k-1} \partial \mathbf{x}_{k-1}^\top} \right] \tag{27a}$$

$$\mathbf{D}_{k-1}^{12} = -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \mathbf{x}_{k-1} \partial \mathbf{x}_k^\top} \right] \tag{27b}$$

$$\mathbf{D}_{k-1}^{13} = -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \mathbf{x}_{k-1} \partial \boldsymbol{\theta}^\top} \right] \tag{27c}$$

$$\begin{aligned} \mathbf{D}_{k-1}^{22} &= -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{z}_k | \mathbf{x}_k; \boldsymbol{\theta})}{\partial \mathbf{x}_k \partial \mathbf{x}_k^\top} \right] \\ &\quad - \mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \mathbf{x}_k \partial \mathbf{x}_k^\top} \right] \end{aligned} \tag{27d}$$

$$\begin{aligned} \mathbf{D}_{k-1}^{23} &= -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{z}_k | \mathbf{x}_k; \boldsymbol{\theta})}{\partial \mathbf{x}_k \partial \boldsymbol{\theta}^\top} \right] \\ &\quad - \mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \mathbf{x}_k \partial \boldsymbol{\theta}^\top} \right] \end{aligned} \tag{27e}$$

$$\begin{aligned} \mathbf{D}_{k-1}^{33} &= -\mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{z}_k | \mathbf{x}_k; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right] \\ &\quad - \mathbb{E} \left[ \frac{\partial^2 \ln p(\mathbf{x}_k | \mathbf{x}_{k-1})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right]. \end{aligned} \tag{27f}$$

Before calculating the HCRB, we note that  $\tilde{\mathbf{J}}_k$  defined in (25), henceforth dubbed the sub-hybrid Fisher information matrix (SHFIM), is not the actual HFIM. The full HFIM is a matrix which inverse is the complete HCRB on the entire parameter vector, while the inverse of  $\tilde{\mathbf{J}}_k$  in (25) is the sub-block of the HCRB that corresponds to  $[\mathbf{x}_k^\top, \boldsymbol{\theta}^\top]^\top$ . Also note that the expression in (26) is recursive, where  $\tilde{\mathbf{J}}_k$  is expressed in terms of  $\tilde{\mathbf{J}}_{k-1}$ . Thus, given  $\tilde{\mathbf{J}}_{k-1}$ , it is sufficient to calculate the innovations  $\{\mathbf{D}_{k-1}^{ij}\}_{i,j=1}^3$ , according to (27) in order to obtain the updated value of  $\tilde{\mathbf{J}}_k$ . A road map and the proof summary is given in Table 1.

**Table 1** Proposition and proof road map

**1. Proposition 1**

2. In Proposition 1, we present the expressions of the innovations  $\{\mathbf{D}_{k-1}^{ij}\}_{i,j=1}^3$  for our models presented in (3), (7), and (8).

3. The innovations are inserted into (26) for estimating the SHFIM in (25).

**4. Guidelines**

5. We begin by calculating  $\mathbf{D}_{k-1}^{11}$ ,  $\mathbf{D}_{k-1}^{12}$  and  $\mathbf{D}_{k-1}^{13}$ .

6. We then present the general expressions for  $\mathbf{D}_{k-1}^{22}$ ,  $\mathbf{D}_{k-1}^{23}$  and  $\mathbf{D}_{k-1}^{33}$ . which require the Jacobian of  $\mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})$ .

7. The Jacobian of  $\mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})$  is calculated.

8. The derivative of  $\mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})$  w.r.t.  $\mathbf{x}_k$  is calculated. It depends only on the wall angles and denoted as  $\mathbf{G}(\boldsymbol{\alpha})$ .

9. This is used to calculate  $\mathbf{D}_{k-1}^{22}$ .

10. Next the derivative of  $\mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})$  w.r.t.  $\boldsymbol{\theta}$  is calculated.

11. Expressions  $\mathbf{D}_{k-1}^{23}$  and  $\mathbf{D}_{k-1}^{33}$  are calculated.

**Proposition 1** The components of  $\tilde{\mathbf{J}}_k$  (cf. (25)-(26b)), are given by

$$\mathbf{D}_{k-1}^{11} = \mathbf{F}^\top \mathbf{C}_w^{-1} \mathbf{F} \tag{28a}$$

$$\mathbf{D}_{k-1}^{12} = -\mathbf{F}^\top \mathbf{C}_w^{-1} \tag{28b}$$

$$\mathbf{D}_{k-1}^{13} = 0 \tag{28c}$$

where the matrix  $\mathbf{F}$  is defined in (4) and  $\mathbf{D}_{k-1}^{11}, \mathbf{D}_{k-1}^{12} \in \mathbb{R}^{2 \times 2}$ . Furthermore,

$$\begin{aligned} \mathbf{D}_{k-1}^{22} &= \mathbf{G}^\top (\boldsymbol{\alpha}) \mathbf{C}_v^{-1} \mathbf{G} (\boldsymbol{\alpha}) + \mathbf{C}_w \\ &= \frac{1}{\sigma_v} \begin{bmatrix} \sum_{i=1}^4 \sin^2 (\alpha_i) & -\sum_{i=1}^4 \sin (\alpha_i) \cos (\alpha_i) \\ -\sum_{i=1}^4 \cos (\alpha_i) \sin (\alpha_i) & \sum_{i=1}^4 \cos^2 (\alpha_i) \end{bmatrix} + \mathbf{C}_w. \end{aligned} \tag{29}$$

where

$$\mathbf{G} (\boldsymbol{\alpha}) = \begin{bmatrix} \sin (\alpha_1) & -\cos (\alpha_1) \\ \sin (\alpha_2) & -\cos (\alpha_2) \\ -\sin (\alpha_3) & \cos (\alpha_3) \\ -\sin (\alpha_4) & \cos (\alpha_4) \end{bmatrix}. \tag{30}$$

The matrix  $\mathbf{D}_{k-1}^{23} \in \mathbb{R}^{8 \times 2}$  is given as follows: its first column is obtained by stacking

$$[\mathbf{D}_{k-1}^{23}]_{i1} = \begin{bmatrix} \text{sign}(2-i) (\eta_{x_k} \cos \alpha_i \sin \alpha_i - (b_i - \eta_{y_k}) \sin^2 \alpha_i) \\ \sin \alpha_i \cos \alpha_i \end{bmatrix} \tag{31}$$

for  $i = 1, \dots, 4$  on top of each other where  $\text{sign}(x) = 1$  if  $x \geq 0$ . The second column is similarly obtained by replacing  $[\mathbf{D}_{k-1}^{23}]_{i1}$  with

$$[\mathbf{D}_{k-1}^{23}]_{i2} = \begin{bmatrix} \text{sign}(i-2) ((b_i - \eta_{y_k}) \sin^2 \alpha_i - \eta_{x_k} \cos^2 \alpha_i) \\ -\cos^2 \alpha_i \end{bmatrix} \tag{32}$$

where

$$\eta_{x_k} \triangleq \mathbb{E} [x_k] = \sum_{i=0}^{k-2} \rho^i \mu \sin (\phi_{k-1-i}) \tag{33a}$$

$$\eta_{y_k} \triangleq \mathbb{E} [y_k] = \sum_{i=0}^{k-2} \rho^i \mu \cos (\phi_{k-1-i}), \tag{33b}$$

Finally,  $\mathbf{D}_{k-1}^{33}$  is the following block diagonal matrix

$$\mathbf{D}_{k-1}^{33} = \text{blockdiag} (\boldsymbol{\Delta}_1, \dots, \boldsymbol{\Delta}_4), \tag{34}$$

where

$$\boldsymbol{\Delta}_i = \frac{1}{\sigma_v^2} \begin{bmatrix} \zeta_i & \chi_i \\ \chi_i & \cos^2 \alpha_i \end{bmatrix} \tag{35}$$

$i = 1, 2, 3, 4$ , and

$$\begin{aligned} \zeta_i &= \sin^2 (\alpha_i) (-2b_i \eta_{y_k} + b_i^2 + m_{y_k}) + \sin (2\alpha_i) \\ &\quad (\eta_{x_k} \eta_{y_k} - b_i \eta_{x_k}) + \cos^2 (\alpha_i) m_{x_k} \end{aligned} \tag{36}$$

$$\chi_i = -(b_i - \eta_{y_k}) \sin \alpha_i \cos \alpha_i + \eta_{x_k} \cos^2 \alpha_i \tag{37}$$

with

$$m_{x_k} \triangleq \mathbb{E}[x_k^2] = \sigma_{x_k}^2 + \eta_{x_k}^2 = \sum_{i=0}^{k-2} \rho^{2i} \left( \sigma_v^2 + (\mu \sin(\phi_{k-1-i}))^2 \right) \quad (38)$$

$$m_{y_k} \triangleq \mathbb{E}[y_k^2] = \sigma_{y_k}^2 + \eta_{y_k}^2 = \sum_{i=0}^{k-2} \rho^{2i} \left( \sigma_v^2 + (\mu \cos(\phi_{k-1-i}))^2 \right). \quad (39)$$

*Proof* We would like to show that for the problem at hand (cf. Definition 1), (27a)-(27f) lead to  $\mathbf{D}_{k-1}^{ij}$ ,  $i, j \in \{1, 2, 3\}$  given in Proposition 1. We begin by noting that our problem fits the general structure of the filtering problem [29] in the sense that both are hybrid with a linear state-space kinematic model and a nonlinear measurement model with additive Gaussian noise. This establishes (28c) and, in addition, it implies

$$\mathbf{D}_{k-1}^{22} = \mathbb{E} \left[ \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right)^\top \mathbf{C}_v^{-1} \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right) \right] + \mathbf{C}_w^{-1} \quad (40a)$$

$$\mathbf{D}_{k-1}^{23} = \mathbb{E} \left[ \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} \right)^\top \mathbf{C}_v^{-1} \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right] \quad (40b)$$

$$\mathbf{D}_{k-1}^{33} = \mathbb{E} \left[ \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top \mathbf{C}_v^{-1} \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right] \quad (40c)$$

Next, from (40a)–(40b), it follows that to derive  $\mathbf{D}_{k-1}^{ij}$  for  $i, j \in \{2, 3\}$ , it is necessary to derive the Jacobian of  $\mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})$ . For our kinematic and measurement model (3),(8), the Jacobian component with respect to  $\mathbf{x}_k$  is given by

$$\frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \mathbf{x}_k} = \mathbf{G}(\boldsymbol{\alpha}) \quad (41)$$

where  $\mathbf{G}(\boldsymbol{\alpha})$  is defined in (30). Thus, by substituting (41) and (30) into (40a), we establish (29). To show (31) and (32), we note that

$$\frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \text{blockdiag}(\Gamma_1, \Gamma_2, -\Gamma_3, -\Gamma_4) \quad (42)$$

where

$$\Gamma_i = [\gamma_i, \cos \alpha_i] \quad (43)$$

and

$$\gamma_i = \cos(\alpha_i) x_k - \sin \alpha_i (b_i - y_k). \quad (44)$$

From (30) and (41),  $\mathbf{D}_{k-1}^{23}$  can be written as

$$\mathbf{D}_{k-1}^{23} = \frac{1}{\sigma_v} \mathbb{E} \left[ \left( \frac{\partial \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^\top \mathbf{G}(\boldsymbol{\alpha}) \right], \in \mathbb{R}^{8 \times 2}. \quad (45)$$

Then, a brute-force derivation of (45) establishes (31) and (32).

To complete the proof, it remains to show (34). It can be shown that

$$\boldsymbol{\Delta}_i = \frac{1}{\sigma_v} \mathbb{E} \left[ \begin{array}{cc} \gamma_i^2 & \gamma_i \cos(\alpha_i) \\ \gamma_i \cos(\alpha_i) & \cos^2(\alpha_i) \end{array} \right] \quad (46)$$

for  $i = 1, 2, 3, 4$  where  $\gamma_i$  is defined in (44). Taking the expectation while noting that  $x_k$  and  $y_k$  are statistically independent, it follows that

$$\mathbb{E} \{ \gamma_i^2 \} = \zeta_i \quad (47)$$

$$\mathbb{E} \{ \gamma_i \} \cos(\alpha_i) = \chi_i \quad (48)$$

where  $\zeta_i$  and  $\chi_i$  are defined in (36) and (37), respectively. It follows that (46), (47), and (48) establish (35) thus establishing (34).  $\square$

## 6 Simulation results and discussion

In this section, we present numerical results of two simulation setups. In the first setup (Section 6.1), the signal is generated via pseudo-random sequences according to (3), (4), and (8), and the proposed EKF algorithm is compared to the HCRB. It will be shown that the proposed algorithm approaches the bound. Therefore, it is not compared to any additional estimation method. In the second setup (Section 6.2), the EKF algorithm is evaluated using audio chirp signals that were convoluted with a simulated room impulse response (RIR) using [42]. The result obtained in the second setup is not compared to the HCRB since the measurements do not suit the model under which the HCRB was developed. Explicitly, the HCRB was derived for noisy observation of the distances between the robot and the walls, whereas the observation in the second setup is the raw acoustic signal. To extract the distance from that measurement, it is necessary to preprocess it by an additional estimation pipeline. The reason we include the second setup is to show that the proposed algorithm is potentially applicable to actual problems. Nevertheless, we note that in real-life scenarios, this requires additional components, e.g., echo labeling block.

### 6.1 Evaluation of the EKF algorithm with respect to the HCRB

We consider a scenario in which the robot travels in an enclosed four-wall room, where each wall is characterized by an offset and angle as depicted in Fig 2. For each Monte-Carlo trial, the robot movement path was generated according to (3). For each  $k$ , the movement command  $\mathbf{u}_k$  was defined by a constant step size  $\mu = 0.5$  m and a uniformly distributed angle  $\phi_k$ . In determining the robot step-size, we had the following consideration. Recall that for each  $k$ , the EKF uses a first-order series approximation of the distance function  $\mathbf{h}$  around  $\Phi_{k|k-1}$ , and note that a large step-size leads to a large variation in  $\Phi$ . Thus, the linear approximation becomes crude if the step size is too large, causing the EKF to diverge. On the other hand, if the step-size is too small, the robot will not travel enough distance required for mapping. In addition, the parameter  $\rho$  in (4) was set to  $\rho = 0.97$ . The variance of the robot step-error,  $\mathbf{w}_k$  (cf. (3), and the subsequent paragraph), was taken as  $\sigma_w = 0.02$  m, which is 4% of the robot step-size. The measurements  $\mathbf{z}_k$  were generated as follows. For each  $\mathbf{x}_k$ , we calculated  $\mathbf{h}(\mathbf{x}_k, \theta)$  using (8), and then, to reflect the inaccuracy in measuring the distance, we added the Gaussian noise  $\mathbf{v}_k$  with  $\sigma_v = 0.02$  m, which is also 4% of the step size. Note that these values, i.e.,  $\sigma_w$ , and  $\sigma_v$  being 4%, are significant since it is accumulated as the algorithm progresses.

The experiment was repeated and averaged over 40 different rooms. All rooms are 4-wall rooms that differ in the angle between each adjacent wall. The angle was randomized from a uniform distribution in the range  $[-5, 5]$  degrees. Figures 3 and 4 depict the mean

square error (MSE) of the estimation of the room geometry parameters,  $\{\alpha_i, b_i\}_{i=1}^4$ , using the proposed EKF algorithm. The MSE is defined as

$$e_{\bar{\alpha}} = \frac{1}{4} \sum_{i=1}^4 (\hat{\alpha}_i - \alpha_i)^2 \tag{49a}$$

$$e_{\bar{b}} = \frac{1}{4} \sum_{i=1}^4 (\hat{b}_i - b_i)^2 \tag{49b}$$

and calculated via Monte-Carlo. The figures also depict the corresponding HCRBs

$$\text{HCRB}_{\bar{\alpha}} = \frac{1}{4} \sum_{i=1}^4 \text{HCRB}_{\alpha_i} \tag{50a}$$

$$\text{HCRB}_{\bar{b}} = \frac{1}{4} \sum_{i=1}^4 \text{HCRB}_{b_i}. \tag{50b}$$

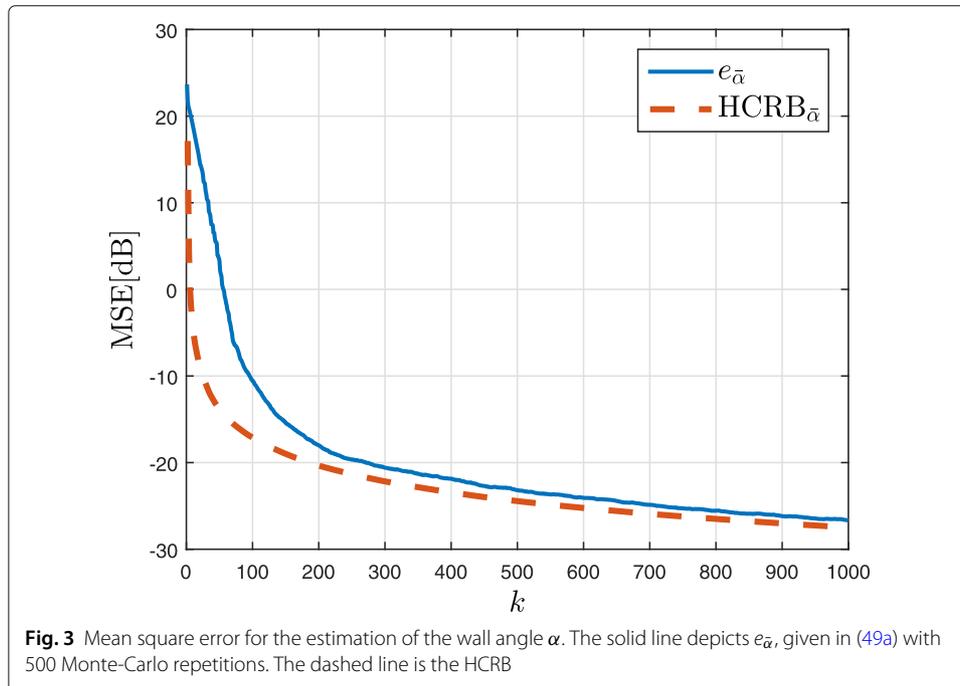
Figure 5 depicts the estimation error of the robot's location

$$e_{\bar{x}} = (\hat{x}_k - x_k)^2 + (\hat{y}_k - y_k)^2 \tag{51}$$

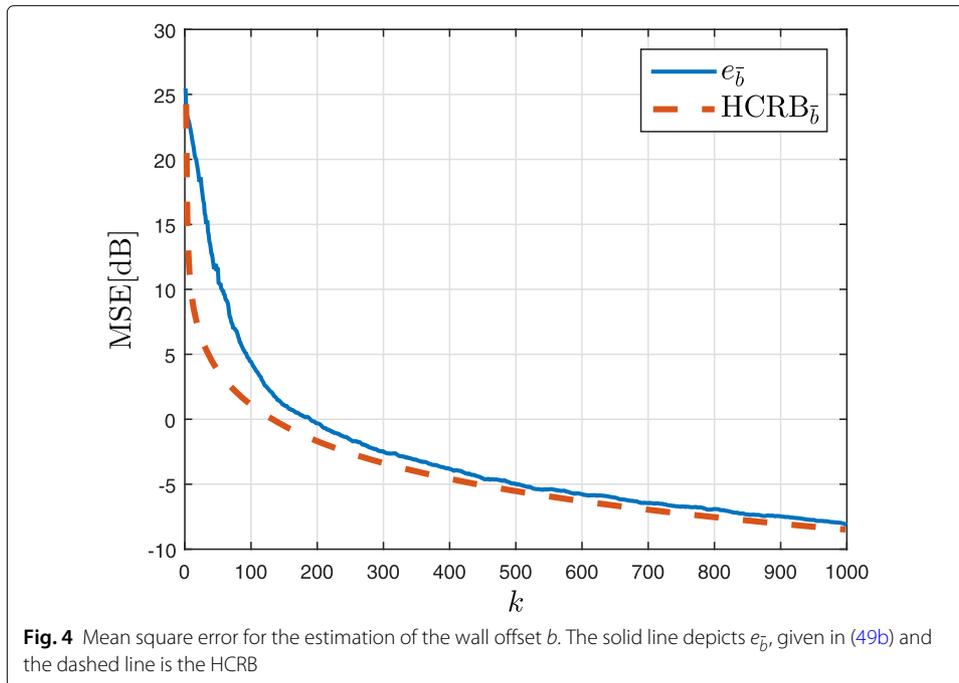
and compares it to the HCRB

$$\text{HCRB}_{\bar{x}} = \text{HCRB}_x + \text{HCRB}_y. \tag{52}$$

Examining the result in Fig. 5, the estimation error of the EKF algorithm during the first steps, i.e., approximately until  $k = 50$ , is lower than the HCRB. This low estimation error in the robot location follows from the fact that the initial robot location determines the origin of the coordinate system. Thus, with respect to this coordinate system, the robot location is known. As a result, we see a discrepancy between the localization results and the bound in the first 50 steps. Accordingly, this part of the graph should be ignored.

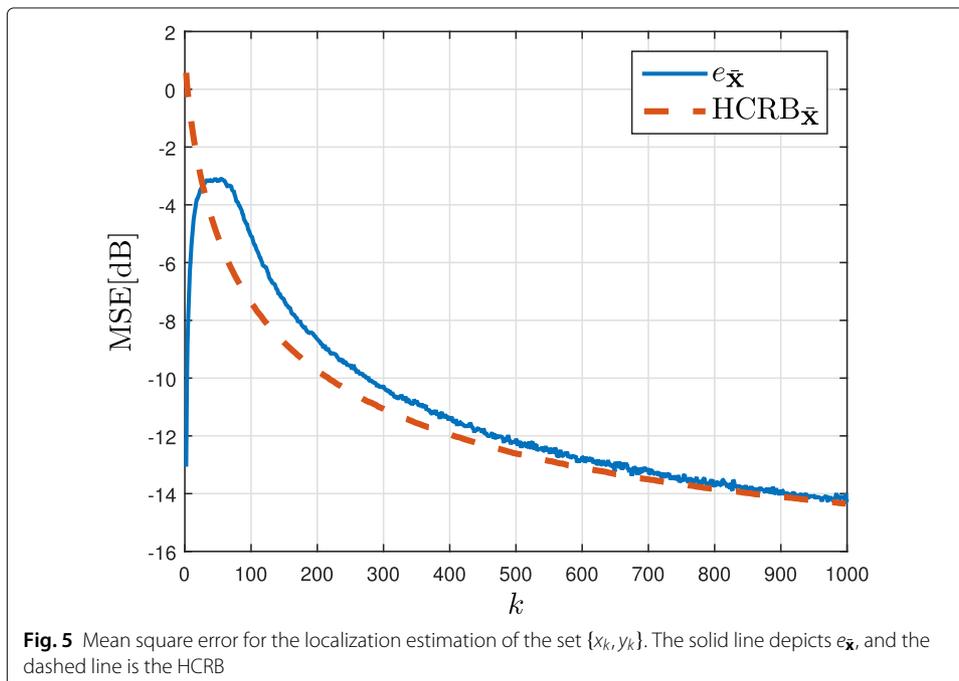


**Fig. 3** Mean square error for the estimation of the wall angle  $\alpha$ . The solid line depicts  $e_{\bar{\alpha}}$ , given in (49a) with 500 Monte-Carlo repetitions. The dashed line is the HCRB



This is not the case for the randomly initialized room geometry parameters as depicted in Figs. 3 and 4.

The results show that the HCRB is large for small values of  $k$ , which is expected in the case of a few observations. Moreover, the bound drops significantly as  $k$  increases up to  $k = 20$ . This can be attributed to the significant contribution of each new measurement when the overall number of measurements is small. We note that at approximately  $k = 80$ , the slope becomes very moderate. We also note that the estimation error approaches the

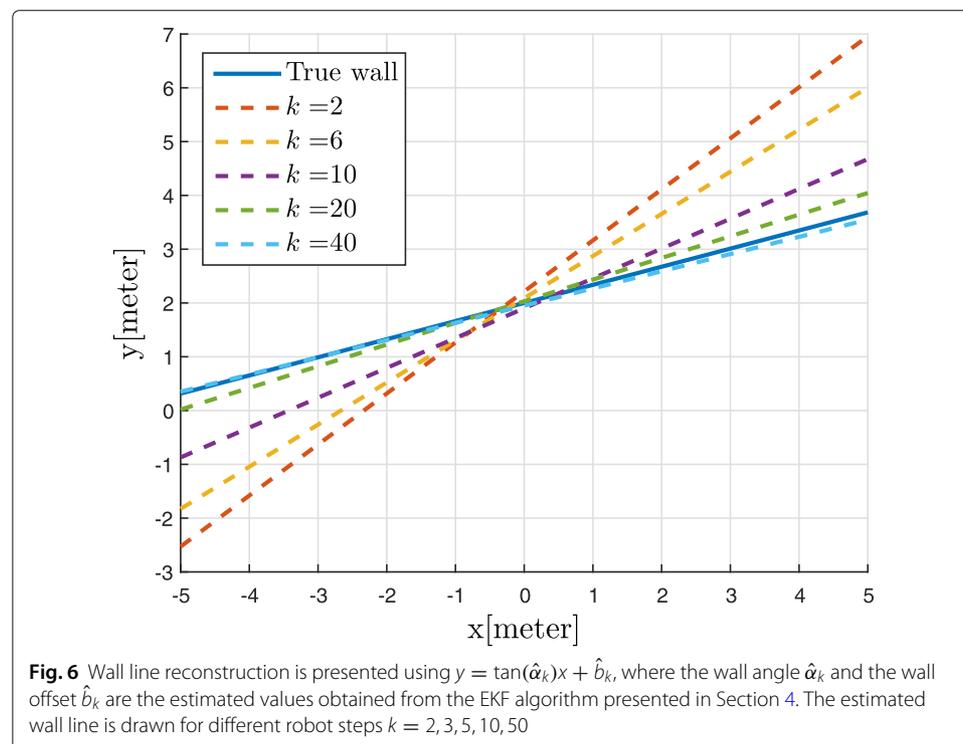


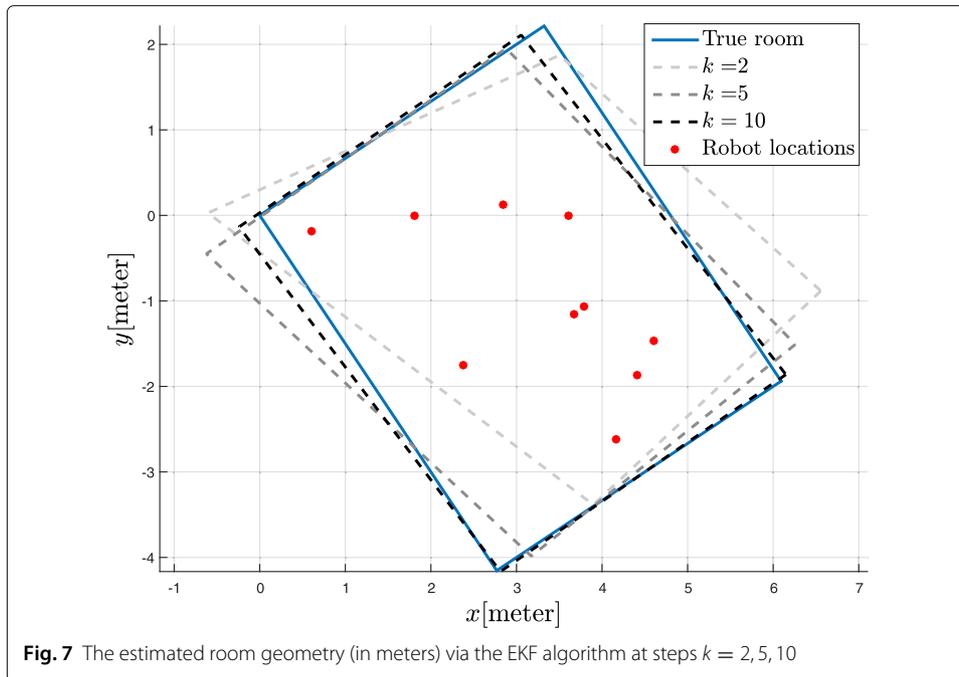
bound as  $k$  is increased as depicted in Figs. 3, 4, and 5, and the estimation error approaches the HCRB, implying that the EKF is an asymptotically optimal estimator and that the HCRB is asymptotically tight.

Figure 6 depicts the estimated line-equation of one of the room's walls, i.e., the line in the plane which is determined by the estimated angle and offset. Each line corresponds to a different value of  $k$ . The results show that the estimated line coincides with the true line as  $k$  increases.

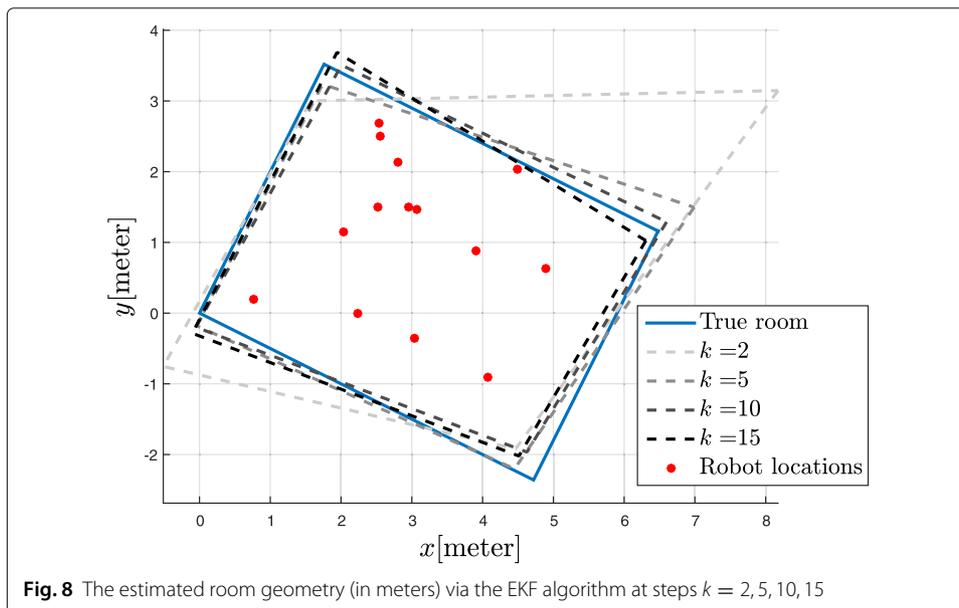
## 6.2 Evaluation based on simulated room impulse response

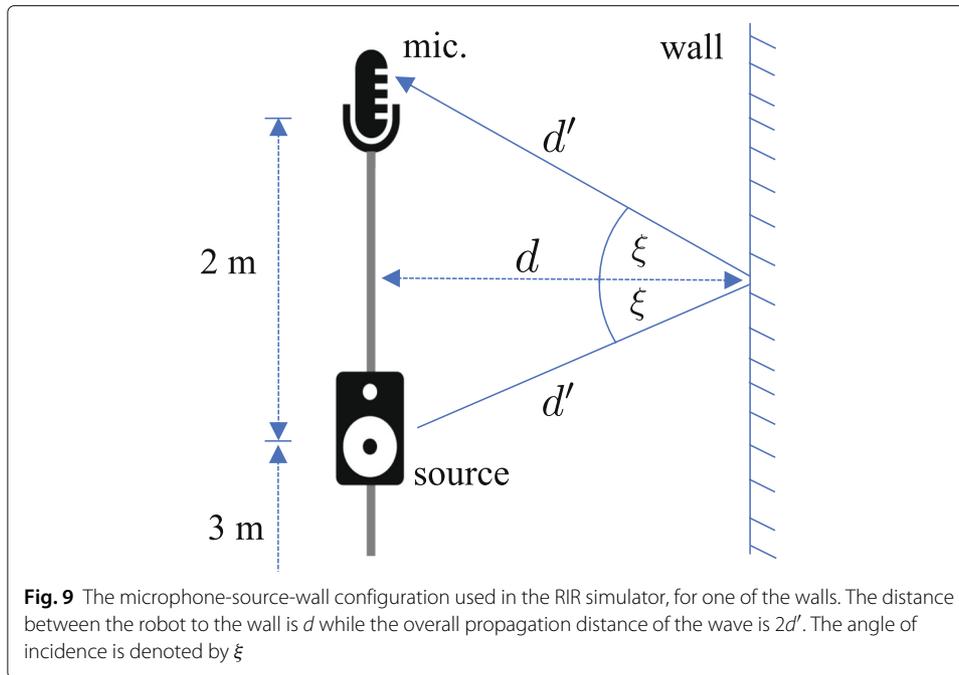
In this simulation, the robot moves a small number of steps chosen randomly (up to 15 steps for each run) and estimates the mapping parameters of rectangular rooms. In Figs. 7 and 8, we present two examples of robot trajectories and the evolution of the room shape estimates, as the robot moves in the room, in comparison to the true room geometry. There is, however, a technical difficulty in carrying out the simulation using the RIR generator [42]. In the latter, it is impossible to place the emitting source and the microphone at the same point, an assumption that underlies our 2D model (c.f. Section 3). To overcome that, we fed the simulator with a three-dimensional room with dimensions  $4 \times 5 \times 7$  m. The idea is to only have horizontal reflections, as in a 2D room. To achieve that, a low reflection coefficient, of value 0.1, was set to the ceiling and floor, while the other four walls were attributed with a high reflection coefficient of 0.95, as mentioned. We used a feature of the RIR generator [42] that allows us to construct an RIR comprising only the direct-path and the first-order reflections (with pre-defined reflection coefficients). The reverberation-time, which indicates the decay of the reverberation tail is therefore meaningless. In addition, the robot walked horizontally, where the emitting source and the microphone were placed on the robot such that their heights (with





respect to the room’s floor) were 3 m and 5 m, respectively, as depicted in Fig. 9. This way, the source and microphone were not at the same point, but their projections, as well as the projection of the distances to the walls, induced a 2D room in which the source and microphone were collocated. To project these distances, one has to discard the first peak, which corresponds to the direct-path between the microphone and source, and to take into account the incidence angle to each wall at every step. Explicitly, taking the midpoint between the microphone and the source as the robot position, we calculate the distance to each wall by compensating and incident angle, as depicted in Fig. 9.





**Fig. 9** The microphone-source-wall configuration used in the RIR simulator, for one of the walls. The distance between the robot to the wall is  $d$  while the overall propagation distance of the wave is  $2d'$ . The angle of incidence is denoted by  $\xi$

In more explicit terms, let  $d'_{i,k}$ ,  $1 \leq i \leq 4$  be the distance between wall- $i$  and the robot after step  $k$  and  $\hat{d}'_{i,k}$  be its corresponding estimate, extracted from the TOA. Then,  $\mathbf{z}_k$  is set as  $z_{ik} = \hat{d}_{ik}$ , where  $\hat{d}_{ik}$  is the two dimensional distance, given by  $\hat{d}_{ik} = \hat{d}'_{ik} \cos(\xi_{ik})$ , and  $\xi_{ik}$  is the incidence-angle<sup>4</sup>. The result demonstrates that the estimates improve as the number of steps increases. Additional scenarios were simulated for several trajectories demonstrating similar results. Note that the orientation of the coordinate system is determined by the first robot step and is different in each scenario.

The run time of each EKF estimation step on i7 Intel processor, implemented using Matlab®, took 260 mSec. Convergence time depends on the model and measurement errors described in (3), (10), and (8), as well as the decay coefficients in (12). As mentioned before, this simulation is not compared to the HCRB since it involves a prior stage that is not taken into consideration while calculating the bound. This simulation study is merely added to demonstrate that the proposed algorithm may be applied to real-life scenarios, provided that additional algorithmic blocks are implemented.

Since a small number of steps are taken, we set  $\rho = 1$ , where  $\rho$  is given in (4), and  $\sigma_w = 0.05$  m (c.f. Section 3.1). The measurements  $\mathbf{z}_k$  were generated as follows. At each step,  $k$ , we extracted the corresponding RIR using [42], with only the first-order reflections simulated (although unrealistic, it still provides an interesting scenario for the current study. Full reverberation regime is left for future research.). The RIR was convoluted with a linear chirp signal of 128 mSec long, with start and stop frequencies 0.5 KHz and 5 KHz, respectively. Sampling rate was set to 16 KHz. The output was then contaminated with a white, additive Gaussian noise. The signal-to-noise ratio (SNR) at the microphone is set to 5.6 dB. The received signal was correlated with the known chirp signal, and from the

<sup>4</sup>Note that  $\xi_{ik}$ , is not an unknown parameter in the original problem (cf. Definition 1), but rather an artifact of the setup in Fig. 9, designated to overcome the constraint that the source and microphone cannot be collocated with the RIR simulator implementation that we have used. Thus, we used the robot's true location at time  $k$  to calculate  $\xi_{ik}$ , since otherwise we would have incorporated an additional nuisance parameter.

result, we picked the TOA of all peaks that were higher than one ninth of the strongest peak. In the rare cases where less than four peaks satisfied this condition, we took the four strongest. The simulation reflection coefficient was set to 0.95, which is relatively high. This was done to ensure that all echoes corresponding to the walls' direct paths would be discovered. If no matching echo is found for a specific wall, the mapping parameters (of that wall) cannot be estimated at the current robot step. We then solved the echo labeling<sup>5</sup> problem and set  $\hat{\tau}_{ik}$ , the estimate of the TOA between the robot and wall- $i$  (cf. Section 3.2), as half the TOA of the corresponding echo. Finally, we calculated the distance  $\hat{d}_{ik}$  according to (9) and concatenated all distances in the measurement vector  $\mathbf{z}_k$ , i.e.,  $z_{ik} = \hat{d}_{ik}$ ,  $i = 1, \dots, 4$ .

## 7 Conclusions

In this paper, we considered the self-localization of a source that travels in an unknown environment, while simultaneously estimating its geometry from noisy time-of-arrival observations. We derived a recursive algorithm, based on the EKF, for this problem as well as the corresponding HCRB. Simulation results showed that the algorithm achieves the HCRB as the number of observations increases. Although this is not a formal proof, it implies that asymptotically, the EKF is an optimal estimator and that the HCRB a tight bound.

As a concrete application, we considered the case where the time-of-arrival is extracted from acoustic echoes and demonstrated the applicability of the proposed EKF algorithm.

An important direction for future research is to test theoretically if the tightness of the HCRB, as well as the optimality of the EKF that were observed in the simulation, is valid in general. The main challenge here is to extend the results about the tightness of HCRB [25] to cases where it is calculated recursively. Another important extension is to incorporate model order selection techniques to determine the number of walls, prior to applying the proposed EKF algorithm. In addition, other type of mismodelling could be considered [43], e.g., if the assumed kinematic model is not accurate.

### Abbreviations

SLAM: Simultaneous localization and mapping; RIR: Room impulse response; SNR: Signal-to-noise ratio; EKF: Extended Kalman filter; CRB: Cramér-Rao bound; FIM: Fisher information matrix; HFIM: Hybrid Fisher information matrix; MDS: Markovian dynamic system; TOA: Time of arrival; HCRB: Hybrid CRB; BCRB: Bayesian CRB; MSE: Mean square error; SHCRB: Sub-hybrid Cramér-Rao bound; SHFIM: Sub-hybrid Fisher information matrix; TDOA: Time difference of arrival; p.d.f.: Probability density function

### Acknowledgements

Not applicable.

### Authors' contributions

MV, YN, and SG developed the model. Experimental testing was carried out by MV and YN. The paper was written by MV, YS, and SG, with MV the first author. The authors read and approved the final manuscript.

### Funding

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 871245.

### Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

### Ethics approval and consent to participate

Not applicable.

<sup>5</sup>Since echo-labeling is beyond the scope of this paper, we assume throughout that the labeling is given. In this simulation, we used the exact robot location to label the echoes, by picking the label combination that minimizes the square error between the true TOA and the estimated TOA.

**Consent for publication**

All authors agree to publish the submitted paper in this journal.

**Competing interests**

The authors declare that they have no competing interests.

Received: 30 April 2020 Accepted: 2 October 2020

Published online: 07 January 2021

**References**

1. X. Chang, C. Yang, J. Wu, X. Shi, Z. Shi, in *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, A surveillance system for drone localization and tracking using acoustic arrays (IEEE, Sheffield, 2018), pp. 573–577
2. A. Collin, A. T. Espinoza, in *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, SLAM-based performance quantification of sensing architectures for autonomous vehicles (IEEE, Madrid, 2018), pp. 1–6
3. C. Bibby, I. Reid, in *IEEE International Conference on Robotics and Automation*, A hybrid slam representation for dynamic marine environments (IEEE, Anchorage, 2010), pp. 257–264
4. F. Demim, A. Nemra, H. Abdelkadri, A. Bazoula, K. Louadj, M. Hamerlain, in *The IEEE 25th International Conference on Systems, Signals and Image Processing (IWSSIP)*, Slam problem for autonomous underwater vehicle using SVSF filter (IEEE, Maribor, 2018), pp. 1–5
5. H. Durrant-Whyte, T. Bailey, Simultaneous localization and mapping: part I. *IEEE Robot. Autom. Mag.* **13**(2), 99–110 (2006)
6. X. Sheng, Y.-H. Hu, in *Information Processing in Sensor Networks*. ed. by F. Zhao, L. Guibas, Energy based acoustic source localization (Springer, Berlin, Heidelberg, 2003), pp. 285–300
7. S. T. Birchfield, R. Gangishetty, in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005, vol. 4*, Acoustic localization by interaural level difference (IEEE, Philadelphia, 2005), pp. 1109–11124
8. N. D. Gaubitch, W. B. Kleijn, R. Heusdens, in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Auto-localization in ad-hoc microphone arrays (IEEE, Vancouver, 2013), pp. 106–110
9. I. Dokmanić, L. Daudet, M. Vetterli, in *22nd European Signal Processing Conference (EUSIPCO)*, How to localize ten microphones in one finger snap (IEEE, Lisbon, 2014), pp. 2275–2279
10. S. Gannot, T. G. Dvorkind, Microphone array speaker localizers using spatial-temporal information. *EURASIP J. Adv. Sig. Proc.* **2006**(12), 1–17 (2006)
11. T. Kundu, Acoustic source localization. *Ultrasonics.* **54**(1), 25–38 (2014). <https://doi.org/10.1016/j.ultras.2013.06.009>
12. S. Tevo, T. Tossavainen, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 3D room geometry estimation from measured impulse responses (IEEE, Kyoto, 2012), pp. 513–516
13. I. Dokmanić, Y. M. Lu, M. Vetterli, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Can one hear the shape of a room: the 2-D polygonal case (IEEE, Prague, 2011), pp. 321–324
14. I. Jager, R. Heusdens, N. D. Gaubitch, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Room geometry estimation from acoustic echoes using graph-based echo labeling (IEEE, Shanghai, 2016)
15. T. S. Ho, Y. C. Fai, E. S. L. Ming, in *10th Asian Control Conference (ASCC)*, Simultaneous localization and mapping survey based on filtering techniques (IEEE, Sabah, 2015)
16. C. Evers, A. H. Moore, P. A. Naylor, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Acoustic simultaneous localization and mapping (A-SLAM) of a moving microphone array and its surrounding speakers (IEEE, Shanghai, 2016), pp. 6–10
17. M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, M. Csorba, A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Trans. Robot. Autom.* **17**(3), 229–241 (2001)
18. M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, in *Eighteenth National Conference on Artificial Intelligence*, FastSLAM: a factored solution to the simultaneous localization and mapping problem (American Association for Artificial Intelligence, Menlo Park, CA, USA, 2002), pp. 593–598
19. I. Dokmanić, L. Daudet, M. Vetterli, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, From acoustic room reconstruction to SLAM, (Shanghai, 2016), pp. 6345–6349
20. M. Krekovic, I. Dokmanić, M. Vetterli, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, EchoSLAM: simultaneous localization and mapping with acoustic echoes, (Shanghai, 2016), pp. 11–15
21. C. Gentner, T. Jost, W. Wang, S. Zhang, A. Dammann, U. C. Fiebig, Multipath assisted positioning with simultaneous localization and mapping. *IEEE Trans. Wirel. Commun.* **15**(9), 6104–6117 (2016)
22. M. Kreković, I. Dokmanić, M. Vetterli, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Omnidirectional bats, point-to-plane distances, and the price of uniqueness, (New Orleans, 2017), pp. 3261–3265
23. M. Krekovic, I. Dokmanić, M. Vetterli, Shapes from echoes: uniqueness from point-to-plane distance matrices. *IEEE Trans. Sig. Proc.* **68**, 2480–2498 (2020)
24. Y. Rockah, P. Schultheiss, Array shape calibration using sources in unknown locations—part I: far-field sources. *IEEE Trans. Acoust. Sig. Process.* **35**(3), 286–299 (1987)
25. Y. Noam, H. Messer, Notes on the tightness of the hybrid Cramér Rao lower bound. *IEEE Trans. Sig. Process.* **57**(6), 2074–2084 (2009)
26. H. Messer, in *Fourth IEEE Workshop on Sensor Array and Multichannel Processing (SAM)*, The hybrid Cramér-Rao lower bound - from practice to theory, (Waltham, 2006), pp. 304–307
27. S. Bay, B. Geller, A. Renaux, J. Barbot, J. Brossier, On the hybrid Cramér Rao bound and its application to dynamical phase estimation. *IEEE Sig. Process. Lett.* **15**, 453–456 (2008)
28. X. Zhang, P. Willett, Y. Bar-Shalom, Dynamic Cramér-Rao bound for target tracking in clutter. *IEEE Trans. Aerosp. Electron. Syst.* **41**(4), 1154–1167 (2005)
29. L. Xu, X. R. Li, in *12th International Conference on Information Fusion*, Hybrid Cramér-Rao lower bound on tracking ground moving extended target (IEEE, Seattle, 2009), pp. 1037–1044

30. H. L. Van Trees, K. L. Bell, Bayesian bounds for parameter estimation and nonlinear filtering/tracking. *AMC*. **10**, 12 (2007)
31. J. Benesty, M. M. Sondhi, Y. Huang, *Springer handbook of speech processing*. (Springer, Berlin, Heidelberg, 2008)
32. R. G. Brown, P. Y. C. Hwang, *Introduction to random signals and applied kalman filtering: with matlab exercises and solutions*, 3rd edn. (John Wiley & Sons, Reading, Massachusetts, 1997)
33. E. A. Wan, R. Van Der Merwe, in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, The unscented Kalman filter for nonlinear estimation (IEEE, Lake Louise, 2000), pp. 153–158
34. D. Hertz, Time delay estimation by combining efficient algorithms and generalized cross-correlation methods. *IEEE Trans. Acoust. Speech Sig. Process.* **34**(1), 1–7 (1986)
35. J. Chen, Y. Huang, J. Benesty, *Time delay estimation*. (Y. Huang, J. Benesty, eds.) (Springer, Boston, MA, 2004), pp. 197–227
36. M. Crocco, A. Trucco, A. Del Bue, in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Room reflectors estimation from sound by greedy iterative approach (IEEE, Calgary, 2018), pp. 6877–6881
37. I. Dokmanić, R. Parhizkar, A. Walther, Y. M. Lu, M. Vetterli, Acoustic echoes reveal room shape. *Proc. Natl. Acad. Sci.* **110**(30), 12186–12191 (2013)
38. S. Gannot, M. Moonen, in *The International Workshop on Acoustic Echo and Noise Control (IWAENC)*, On the application of the unscented Kalman filter to speech processing, (Kyoto, Japan, 2003), pp. 8–11
39. S. Haykin, *Kalman filtering and neural networks, vol. 47*. (John Wiley & Sons, New York, 2004)
40. F. Gini, R. Reggiannini, On the use of Cramér-Rao-like bounds in the presence of random nuisance parameters. *IEEE Trans. Commun.* **48**(12), 2120–2126 (2000)
41. C. Ren, J. Galy, E. Chaumette, F. Vincent, P. Larzabal, A. Renaux, Recursive hybrid Cramér Rao bound for discrete-time Markovian dynamic systems. *IEEE Sig. Process. Lett.* **22**(10), 1543–1547 (2015)
42. E. A. P. Habets, Room impulse response generator (2016). <https://github.com/ehabets/RIR-Generator>. Accessed 2 Jan 2014
43. Y. Noam, J. Tabrikian, Marginal likelihood for estimation and detection theory. *IEEE Trans. Sig. Process.* **55**(8), 3963–3974 (2007)

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---