National ICT Australia Technical Reports ISSN 1833-9646

Posterior Cramér-Rao Bound for Acoustic Source Tracking in Reverberant Environments

Eric A. Lehmann[†] and *Robert C. Williamson*[‡]

[†]Western Australian Telecommunications Research Institute
35 Stirling Highway, Perth WA 6009, Australia
Phone: +61 8 6488 4642, Fax: +61 8 6488 7254
E-mail: Eric.Lehmann@watri.org.au

[‡]National ICT Australia & Australian National University
Locked Bag 8001, Canberra ACT 2601, Australia
E-mail: Bob.Williamson@nicta.com.au

Abstract

The research presented in this paper is motivated by the desire to determine a lower bound on the estimation error for acoustic source localisation and tracking (ASLT) algorithms operating in reverberant conditions. The present work makes use of the so-called posterior Cramér-Rao bound (PCRB), a generalisation of the well-known Cramér-Rao bound theory that is directly applicable to this kind of target-tracking situation. To this purpose, we propose to represent the considered ASLT problem with a system model that accounts for the effects of the temporal correlation existing between sound intensity values measured in a reverberant sound field by means of a microphone array beamformer. We then investigate the tracking results obtained with an acoustic source localisation algorithm, based on the concept of particle filtering (PF), used in conjunction with the proposed system model. As a result of the PCRB derivations implicitly assuming statistical whiteness when considering the system and observation noise processes, the lower bound derived in this work remains unaffected by the level of correlation existing in the ASLT model. On the other hand, the tracking accuracy of the particle filter clearly degrades as the level of correlation in the sound intensity measurements increases. This result leads to the conclusion that the PCRB does not completely capture the sort of performance one is interested in for practical applications, and highlights the need for a different type of lower bound that would account for potentially non-white noise processes in the system model.

Keywords: Posterior Cramér-Rao bound, acoustic source tracking, microphone array beamforming, reverberant room acoustics, particle filtering.

1 Introduction

Optimal estimators for nonlinear filtering problems are usually impossible to implement in practice. In order to assess the quality of any filtering algorithm in terms of absolute performance, and not simply with respect to some other algorithm, a theoretical efficiency limit can sometimes be derived. This type of bound on the estimation error gives an indication of the best absolute performance achievable with an optimal estimator, and is solely determined by the specific system under consideration. This theoretical performance limit can then be used as a reference to which the implemented estimator can be compared.

Several types of lower bound for nonlinear systems have been presented in the literature, including the well-known Cramér-Rao bound (CRB) for random parameter estimation [1]. Theoretical derivations of such bounds can be found, for example, in [2–4], and the works in [5, 6] present a detailed overview of existing lower bound types. In this paper, the approach described by Tichavský *et al.* in [5] will be used

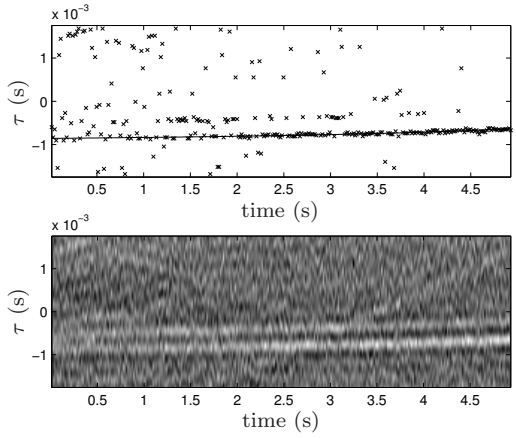

Figure 1. GCC results in reverberant environment, with fast-moving sound source (target speed approximately 1 m/s). *Bottom plot:* GCC output vs. time, dark areas represent low values. *Top plot:* the solid line indicates the target's true time delay, '×' markers denote the lags maximising the GCC output.

to compute the so-called posterior Cramér-Rao bound (PCRB). This approach constitutes a novel and simple method of deriving the estimation error bound for discrete-time nonlinear filtering problems, and it will be applied here specifically to the problem of acoustic source localisation and tracking (ASLT).

The Cramér-Rao bound theory is a popular tool in the target-tracking literature. In [7, 8] for instance, the CRB is used to assess the accuracy of source location estimates delivered using a randomly distributed beamforming sensor array. Ristic and Arulampalam [9–12], in association with various authors, present extensive developments of the CRB for the problem of bearings-only target motion analysis (with an emphasis on defense applications) and ballistic object tracking. The work presented in [13] uses the same principles applied to multiple-target tracking. This type of lower bound analysis has also been used in conjunction with the concept of terrain-aided aircraft navigation [2, 14]. For the problem of acoustic source localisation considered in the present work, the complex process of sound reverberation cannot be easily modeled in a straightforward way. The observations in the case of ASLT typically result from a mixture model comprising a potentially non-white noise process, which creates a challenge for a lower bound analysis of the ASLT problem.

The developments presented in this work are motivated in part by the following speculation. It is known that sound pressure values recorded at two different locations in a reverberant acoustic field present a certain level of correlation that is a function of the distance between the two measuring points (see, e.g., [15]). Likewise, it can be expected that the sound intensity values measured at the same location but at two different times also show some correlation that depends on the displacement of the sound source producing the acoustic field—this reasoning will be elaborated in Section 5. Since a higher level of correlation leads to erroneous observations becoming more consistent temporally, it can be expected that the accuracy of any tracking algorithm for ASLT will decrease accordingly. An example of this effect is illustrated in Figures 1 and 2. The bottom plot in each figure depicts the GCC results (generalised cross-correlation, see [16]) obtained with a two microphone array located in a $3 \times 4 \times 3$ m room with reverberation time $T_{60} \approx 0.1$ s and signal-to-noise ratio SNR ≈ 10 dB. The sensor signals are generated using the image method [17] for a sound source emitting white noise and moving towards the array. The top plots in Figures 1 and 2 show the estimated time delay of arrival of the acoustic source (target), computed as the lag τ maximising the GCC function for each frame of signal data.[1] Figure 1 presents the results obtained with a source moving relatively fast along the pre-defined trajectory. The uncorrelated nature of the erroneous GCC measurements (due to reverberation and additive noise) can be clearly observed in this scenario.[2] On the other hand, Figure 2 depicts the case where the acoustic source moves more slowly along the beginning portion of the same path. In this case, the effects of reverberation lead to spurious measurements which, due to the relatively slow target velocity, appear more correlated

[1] Note that a fast-moving target does not necessarily imply a fast-changing lag τ. In other words, the gradient of the target's time delay vs. time in Figures 1 and 2 is not representative of the target's spatial velocity.

[2] In this figure, the GCC results become more reliable with time as a consequence of the sound source getting closer to the microphone pair (increasing direct-to-reverberant ratio).

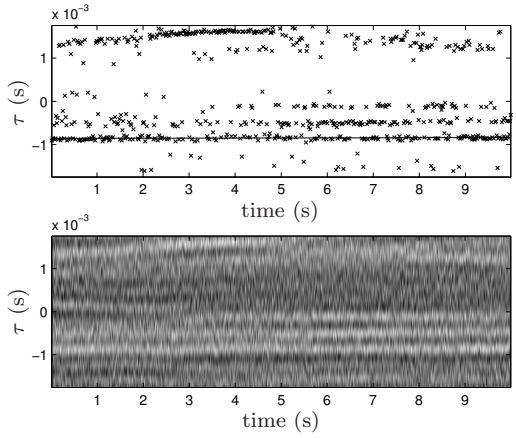

Figure 2. GCC results in reverberant environment, with slow-moving sound source (target speed approximately 0.1 m/s). *Bottom plot:* GCC output vs. time, dark areas represent low values. *Top plot:* the solid line indicates the target's true time delay, '×' markers denote the lags maximising the GCC output.

over time. Such an example is demonstrated in Figure 2 with the strongly correlated erroneous GCC measurement appearing at lag $\tau \approx 1.6 \cdot 10^{-3}\,\mathrm{s}$. The somewhat surprising effect resulting from this is that it can be easier to track a quickly moving source compared to a slowly moving one. The aim of this work is hence to investigate this correlation effect between the measured sound intensity values, and to quantify its influence on the theoretically and practically achievable tracking performance.

The developments made in this paper will proceed as follows. In the next section, a review of the PCRB theory is given, highlighting the existing concepts and principles of the lower bound computations. The PCRB is then derived for the specific problem of ASLT, starting with a mathematical definition of the considered system in Section 3. Section 4 presents the development of a simplified observation model that is then used as an example for the PCRB derivation. A more complex model, accounting for the correlation effects of diffuse sound fields, is then proposed in Section 5. Finally, Section 6 presents an assessment of the tracking performance of a particle filter algorithm for ASLT, proposed previously in the literature, with respect to the proposed system models and derived PCRB. The paper finally concludes with a discussion of the obtained results.

2 Review of PCRB Theory

This section briefly reviews the basics of the PCRB theory for discrete-time nonlinear filtering. It mainly follows the concepts of [5], and similar developments can also be found in [2, 18]. Let $\mathbf{x}$ be a system state vector comprising d parameters $\theta^{(\cdot)}$ that one would like to estimate:

$$\mathbf{x} = \begin{bmatrix} \theta^{(1)} & \theta^{(2)} & \cdots & \theta^{(d)} \end{bmatrix}^{\mathrm{T}},$$

and let $\mathbf{y}$ denote a vector of measurement data (observation). The aim is to derive a lower bound on the mean squared error (MSE) in estimating parameter $\mathbf{x}$:

$$\mathrm{MSE}_{\mathbf{x}} = \mathbb{E}\big\{ \big(\widehat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}\big)\big(\widehat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}\big)^{\mathrm{T}} \big\}, \tag{1}$$

where $\mathbb{E}\{\cdot\}$ is the statistical expectation operator, and $\widehat{\mathbf{x}}(\mathbf{y})$ represents an estimate of the state variable $\mathbf{x}$ based on the measurement data $\mathbf{y}$. It can be seen from (1) that $\mathrm{MSE}_{\mathbf{x}}$ is defined as the covariance matrix of the state vector $\mathbf{x}$. The lower bound (PCRB) on the estimation error for parameter $\mathbf{x}$ is defined as

$$\mathrm{MSE}_{\mathbf{x}} \geqslant \mathbf{J}_{\mathbf{x}}^{-1} \triangleq \mathrm{PCRB}_{\mathbf{x}}\,. \tag{2}$$

The inequality in (2) means that the difference $\mathrm{MSE}_{\mathbf{x}} - \mathbf{J}_{\mathbf{x}}^{-1}$ is a positive semidefinite matrix, and $\mathbf{J}_{\mathbf{x}}$ is the $d \times d$ Fisher information matrix, whose elements are defined as follows [1]:

$$[\mathbf{J}_{\mathbf{x}}]_{(i,j)} = \mathbb{E}\left\{ -\frac{\partial^2 \log p(\mathbf{x}, \mathbf{y})}{\partial\theta^{(i)}\,\partial\theta^{(j)}} \right\}, \tag{3}$$

with the notation $[\,\cdot\,]_{(i,j)}$ representing the (i,j)-th element of a given matrix, and with $p(\mathbf{x},\mathbf{y})$ denoting the joint probability density function (PDF) of the state and observation parameters. Since $\mathrm{MSE}_{\mathbf{x}}$ corresponds to the covariance matrix of the variable $\mathbf{x}$, the PCRB on the estimation error for each parameter $\theta^{(\cdot)}$ in the vector $\mathbf{x}$ is simply given by the diagonal elements of the inverse of the Fisher information matrix:

$$\mathrm{MSE}_{\theta^{(i)}} \geqslant \left[\mathbf{J}_{\mathbf{x}}^{-1}\right]_{(i,i)} \triangleq \mathrm{PCRB}_{\theta^{(i)}} . \tag{4}$$

Using the conventional notation for the first-order and second-order partial derivative operators:

$$\nabla_{\mathbf{x}} = \left[\frac{\partial}{\partial\theta^{(1)}} \;\; \frac{\partial}{\partial\theta^{(2)}} \;\; \cdots \;\; \frac{\partial}{\partial\theta^{(d)}}\right]^{\mathrm{T}},$$
$$\Delta_{\mathbf{z}}^{\mathbf{x}} = \nabla_{\mathbf{z}}\nabla_{\mathbf{x}}^{\mathrm{T}},$$

for arbitrary vector variables $\mathbf{x}$ and $\mathbf{z}$, (3) can be summarised as

$$\mathbf{J}_{\mathbf{x}} = \mathbb{E}\big\{-\Delta_{\mathbf{x}}^{\mathbf{x}} \log p(\mathbf{x},\mathbf{y})\big\} . \tag{5}$$

The above developments are valid provided the derivatives and expectations in (1) and (3) exist, and subject to some additional constraints on the estimation bias; see [1, 5] for details.

Now let k represent the discrete time variable, $k = 1, 2, \ldots$, and consider the generic filtering problem defined by the following state-space equations:

$$\mathbf{x}_k = g(\mathbf{x}_{k-1}, \mathbf{u}_k) , \tag{6a}$$
$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{v}_k) , \tag{6b}$$

where $g(\cdot)$ and $h(\cdot)$ are the (potentially nonlinear) transition and observation functions respectively, and with $\mathbf{u}_k$ and $\mathbf{v}_k$ two independent and white (potentially non-Gaussian) noise processes. Let $\mathbf{x}_{0:k}$ denote the concatenation of the first $k+1$ state vectors, i.e. $\mathbf{x}_{0:k} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k\}$. Based on (5), the $kd \times kd$ Fisher information matrix $\mathbf{J}_{\mathbf{x}_{0:k}}$ for parameter $\mathbf{x}_{0:k}$ can be decomposed into four submatrices as follows:

$$\begin{aligned}\mathbf{J}_{\mathbf{x}_{0:k}} &= \begin{bmatrix} [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(1,1)} & [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(1,2)} \\ [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(2,1)} & [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(2,2)} \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{E}\big\{-\Delta_{\mathbf{x}_{0:k-1}}^{\mathbf{x}_{0:k-1}} \log p(\mathbf{x}_k,\mathbf{y}_k)\big\} & \mathbb{E}\big\{-\Delta_{\mathbf{x}_{0:k-1}}^{\mathbf{x}_k} \log p(\mathbf{x}_k,\mathbf{y}_k)\big\} \\ \mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{0:k-1}} \log p(\mathbf{x}_k,\mathbf{y}_k)\big\} & \mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k,\mathbf{y}_k)\big\} \end{bmatrix} .\end{aligned} \tag{7}$$

It can be seen that the size of $\mathbf{J}_{\mathbf{x}_{0:k}}$ increases over time. Also, only the estimation error on parameter $\mathbf{x}_k$ is of interest in practice. Hence, the following type of bound is sought:

$$\mathrm{MSE}_{\mathbf{x}_k} = \mathbb{E}\big\{\big(\widehat{\mathbf{x}}_k - \mathbf{x}_k\big)\big(\widehat{\mathbf{x}}_k - \mathbf{x}_k\big)^{\mathrm{T}}\big\} \geqslant \mathbf{J}_{\mathbf{x}_k}^{-1} ,$$

where $\mathbf{J}_{\mathbf{x}_k}$ now represents the $d \times d$ Fisher information submatrix for parameter $\mathbf{x}_k$. For brevity and without loss of generality, the matrix $\mathbf{J}_{\mathbf{x}_k}$ will henceforth be denoted by $\mathbf{J}_k$. The PCRB on parameter $\mathbf{x}_k$ is determined by the right-lower block of the inverse of the Fisher information matrix $\mathbf{J}_{\mathbf{x}_{0:k}}$, i.e.:

$$\mathbf{J}_k = [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(2,2)} - [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(2,1)} \big([\mathbf{J}_{\mathbf{x}_{0:k}}]_{(1,1)}\big)^{-1} [\mathbf{J}_{\mathbf{x}_{0:k}}]_{(1,2)} . \tag{8}$$

Equation (8) represents the general formula to compute the PCRB for the parameter of interest. However, it involves the inversion of a large matrix and is hence computationally demanding. Tichavský *et al.* [5] proposed an efficient way of computing $\mathbf{J}_k$ recursively which avoids the manipulation of large matrices. This recursion is given by:

$$\mathbf{J}_k = \mathbf{D}_{k-1}^{(4)} - \mathbf{D}_{k-1}^{(3)}\big(\mathbf{J}_{k-1} + \mathbf{D}_{k-1}^{(1)}\big)^{-1}\mathbf{D}_{k-1}^{(2)} , \tag{9}$$

with the following definitions:

$$\mathbf{D}_{k-1}^{(1)} = \mathbb{E}\big\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_{k-1}} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\big\} , \tag{10a}$$
$$\mathbf{D}_{k-1}^{(2)} = \mathbb{E}\big\{-\Delta_{\mathbf{x}_{k-1}}^{\mathbf{x}_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\big\} , \tag{10b}$$
$$\mathbf{D}_{k-1}^{(3)} = \mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_{k-1}} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\big\} , \tag{10c}$$
$$\mathbf{D}_{k-1}^{(4)} = \mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1})\big\} + \mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{y}_k|\mathbf{x}_k)\big\} , \tag{10d}$$

where $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p(\mathbf{y}_k|\mathbf{x}_k)$ are the transition and observation PDFs respectively, which follow directly from (6). See [5] for a formal proof of (9) and (10). Assuming that the transition and observation equations $g(\cdot)$ and $h(\cdot)$ are time-invariant, the various matrices defined by (10) remain constant over time. Then the recursion for $\mathbf{J}_k$ can be shown to converge to a finite matrix $\mathbf{J}_\infty$ as $k \to \infty$ [5]. This steady-state solution determines the PCRB value of interest, which can be obtained numerically or by solving (9) with $\mathbf{J}_k = \mathbf{J}_{k-1} \triangleq \mathbf{J}_\infty$. The expectations involved in (10) can be practically computed either analytically or as Monte Carlo average over a series of realisations of the different variables inside the expectation operator. This latter process will be explained in more detail in Section 4.2.

It can be seen from the developments given in this section that the lower bound on estimation error only depends on the fundamental properties of the given filtering problem. The error bound is dependent upon the type and level of noise introduced into the system by the variables $\mathbf{u}_k$ and $\mathbf{v}_k$ in (6). It is also influenced by the amount of information regarding $\mathbf{x}_k$ provided by the observation variable $\mathbf{y}_k$, which is ultimately also determined by (6). The PCRB results obtained in Sections 4 and 5 are hence only as relevant in practice as the assumed theoretical model is close to the practical system it attempts to reproduce; this general caveat should be kept in mind throughout the rest of this work. The main task with respect to the above PCRB computations is now to determine the exact transition and observation PDFs for the ASLT problem definition.

3 System Transition Definition

The variables and specific assumptions used for the PCRB computations first need to be defined for the particular context of acoustic source tracking. On the basis of these definitions, the PCRB matrices of (10) can be derived. The developments given in this section only cover the state variable $\mathbf{x}$ and the system transition equation, whereas the definition of variables and concepts related to the measurement process $\mathbf{y}$ (observations) is dealt with in Sections 4 and 5.

3.1 System State Variable

The type of ASLT scenario considered in this work is the same as that used in previous literature, in which the specific position of an acoustic source is to be estimated on the basis of data recorded with an array of acoustic sensors (see, e.g., [19, 20]). For simplicity, and due to the fact that the considered transition model is assumed independent and identical in all Cartesian coordinates x, y and z, the problem will be considered in one single dimension only. Hence, the state variable $\mathbf{x}_k$ that will be assumed throughout the rest of this work is defined as

$$\mathbf{x}_k = [x_k \;\; \dot{x}_k]^{\mathrm{T}},$$

where $\dot{x}$ represents the velocity variable. This situation can be viewed for instance as a case where the sound source is moving along a straight line in a reverberant room. Only the position x_k of the target along this "virtual" line is required, and the observation variable $\mathbf{y}_k$ also results from some measurement recorded along this one-dimensional path.

Since usually only the PCRB on the target position is of practical interest (the velocity component in the state vector is a requirement from the considered system transition model, see next subsection), the lower bound on the location estimate x_k is computed via (4):

$$\mathrm{MSE}_{x_k} \geqslant \left[\mathbf{J}_{\mathbf{x}_k}^{-1}\right]_{(1,1)} \triangleq \mathrm{PCRB}_{x_k} .$$

Of course the same developments are valid for any coordinate, and if desired, a global PCRB value can be determined, e.g., in three dimensions, via

$$\mathrm{PCRB}_{\mathbf{x}_k} \triangleq \mathrm{PCRB}_{x_k} + \mathrm{PCRB}_{y_k} + \mathrm{PCRB}_{z_k} ,$$

which corresponds to the trace of the inverted Fisher information matrix $\mathbf{J}_{\mathbf{x}_k}$ (with the unnecessary target velocity information discarded).

3.2 System Equation

The transition equation used in the sequel to model the target dynamics in (6a) is

$$\mathbf{x}_k = \underbrace{\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}}_{\mathbf{G}} \mathbf{x}_{k-1} + \mathbf{u}_k \,, \tag{11}$$

where T denotes the time interval separating two consecutive values of the discrete time variable k, and with the zero-mean Gaussian noise process $\mathbf{u}_k \sim \mathcal{N}\big([0 \;\; 0]^{\mathrm{T}}, \mathbf{Q}\big)$. The system noise covariance matrix is

$$\mathbf{Q} = q \begin{bmatrix} \frac{1}{2}T^2 & T \\ T & 1 \end{bmatrix}. \tag{12}$$

The type of model defined by (11) and (12) is commonly used in the target-tracking literature (see, e.g., [10–13]) and allows one to control the amount of process noise in the target dynamics through a single parameter q. The transition PDF simply follows from (11) as

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}\left(\mathbf{x}_k; \mathbf{G}\mathbf{x}_{k-1}, \mathbf{Q}\right), \tag{13}$$

where $\mathcal{N}(\boldsymbol{\xi}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ corresponds to the density of a multi-dimensional normally distributed random variable with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, evaluated at $\boldsymbol{\xi}$.

3.3 System-Related PCRB Computations

Given the result of (13), most of the matrices in (10) can be explicitly computed. We have

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \frac{1}{2\pi\sqrt{|\mathbf{Q}|}} \exp\Big(-\frac{1}{2}\left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right)^{\mathrm{T}} \mathbf{Q}^{-1} \left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right)\Big),$$

where $|\mathbf{Q}| = \det(\mathbf{Q})$. The generic term $\log p(\mathbf{x}_k|\mathbf{x}_{k-1})$ in (10) then becomes

$$\log p(\mathbf{x}_k|\mathbf{x}_{k-1}) = -\log\big(2\pi\sqrt{|\mathbf{Q}|}\big) - \frac{1}{2}\left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right)^{\mathrm{T}} \mathbf{Q}^{-1} \left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right). \tag{14}$$

The first term in the right-hand side of (14) is a constant and therefore disappears upon derivation. Using standard algebraic developments, the negative second-order partials of (14) with respect to variable $\mathbf{x}_k$, which is involved in (10d), can then be shown to be

$$\begin{aligned} -\Delta^{\mathbf{x}_k}_{\mathbf{x}_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1}) &= \frac{1}{2}\, \Delta^{\mathbf{x}_k}_{\mathbf{x}_k} \left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right)^{\mathrm{T}} \mathbf{Q}^{-1} \left(\mathbf{x}_k - \mathbf{G}\mathbf{x}_{k-1}\right) \\ &= \mathbf{Q}^{-1}. \end{aligned} \tag{15}$$

Since the matrix $\mathbf{Q}$ is a constant, the first term of matrix $\mathbf{D}^{(4)}_{k-1}$ is hence given as

$$\mathbb{E}\big\{ -\Delta^{\mathbf{x}_k}_{\mathbf{x}_k} \log p(\mathbf{x}_k|\mathbf{x}_{k-1}) \big\} = \mathbf{Q}^{-1}. \tag{16a}$$

Similar developments lead to the following results:

$$\mathbf{D}^{(1)}_{k-1} = \mathbf{G}^{\mathrm{T}} \mathbf{Q}^{-1} \mathbf{G}\,, \tag{16b}$$

$$\mathbf{D}^{(2)}_{k-1} = -\mathbf{G}^{\mathrm{T}} \mathbf{Q}^{-1}\,, \tag{16c}$$

$$\mathbf{D}^{(3)}_{k-1} = -\mathbf{Q}^{-1} \mathbf{G}\,. \tag{16d}$$

The results of (16) define most of the variables required for a computation of the Fisher information submatrix $\mathbf{J}_k$ given in (9). The last term to be computed, i.e. the second term in (10d), is based on the observation density $p(\mathbf{y}_k|\mathbf{x}_k)$, which in turn requires the formulation of an observation model.

4 Simple Observation Model

A simplified observation model is developed in this section and subsequently used as an example in the PCRB computations. This model is based on some basic principles of statistical room acoustics, but does not include the effects of the correlation between sound intensity measurements. This will be treated specifically in Section 5.

4.1 Model Derivation

The purpose of the observation equation in the filtering problem of (6) is to provide an approximate model of the practical measurement process. Within the framework of acoustic source localisation, several processes have an influence on the measurement value ultimately used by the tracking algorithm. These mainly include acoustic reverberation and the processing of the sensor signals with a specific localisation method, based e.g. on steered beamforming (SBF) or time-delay estimation (TDE). In the following, the principle of steered beamforming will be used as example to derive the PCRB for ASLT. This method was previously shown to achieve a better tracking performance compared to TDE-based approaches [19, 21]. Consequently, the SBF principle is used for the purpose of analysing how well this preferred method performs compared to the PCRB. Furthermore, the source signal will be assumed to be band-limited white Gaussian noise in the range $f \in [300\,\mathrm{Hz}, 3000\,\mathrm{Hz}]$.

Assuming that a steered beamformer is implemented to measure the acoustic intensity $\mathcal{P}(x)$ at some focus location x, the measurement $\mathbf{y}_k$ derived from maximising the SBF output corresponds to a direct observation of a potential source position in the state space. The measurement process implicitly used with SBF is hence

$$y \triangleq \mathbf{y} = \arg\max_x \{\mathcal{P}(x)\}\,. \tag{17}$$

Since the observation value defined by the right-hand side of (17) corresponds to a scalar, the notation y_k will be used from now on instead of $\mathbf{y}_k$. The aim of this section is to model (17) with an equation of the form $y_k = h(\mathbf{x}_k, \mathbf{v}_k)$, from which the observation PDF $p(y_k|\mathbf{x}_k)$ can then be derived.

The main idea behind the derivation of the following observation model can be explained as follows. The outcome of the SBF measurement can be one of the following two simple events:

$$\begin{aligned}&\mathcal{S}: \text{target position correctly detected,}\\ &\mathcal{F}: \text{target position incorrectly detected (false detection),}\end{aligned}$$

with respective probabilities $P_\mathcal{S} = \Pr\{\mathcal{S}\}$ and $P_\mathcal{F} = \Pr\{\mathcal{F}\} = 1 - P_\mathcal{S}$. In environments with high SNR levels and high direct-to-reverberant ratios, the SBF output can be expected to provide nearly perfect measurements of the source position, yielding $P_\mathcal{S} \to 1$. As the reverberation level increases in the enclosure, the statistical probability of false detections $P_\mathcal{F}$ will increase accordingly.

In a diffuse sound field comprising many different frequency components, the energy density is assumed uniform throughout the considered enclosure [22]. This consequently leads to the natural assumption that in the case of a false detection, the SBF measurement results in a random location uniformly distributed along the state space. The observation equation then easily follows from this simple analysis as

$$y_k = \begin{cases} [1\ \ 0] \cdot \mathbf{x}_k + v_k & \text{with probability } P_\mathcal{S},\\ \varepsilon_k & \text{with probability } (1 - P_\mathcal{S}),\end{cases} \tag{18}$$

where ε_k is a random variable with uniform distribution in the state space, i.e. along the x-axis between the state-space limits x_a and x_b: $\varepsilon_k \sim \mathcal{U}(x_\mathrm{a}, x_\mathrm{b})$. The scalar noise process v_k accounts for slight errors in the source position estimates for hypothesis $\mathcal{S}$, and is assumed to be a zero-mean Gaussian random variable: $v_k \sim \mathcal{N}(0, \sigma_v^2)$. The standard deviation σ_v of this observation noise process can be roughly related to the width of the main lobe in the SBF beampattern. Given the observation model of (18) and using the vector notation $\mathbf{h} = [1\ \ 0]^\mathrm{T}$, the observation density now trivially follows as a mixture model:

$$p(y_k|\mathbf{x}_k) = P_\mathcal{S} \cdot \mathcal{N}\big(y_k; \mathbf{h}^\mathrm{T}\mathbf{x}_k, \sigma_v^2\big) + (1 - P_\mathcal{S}) \cdot \mathcal{U}\big(y_k; x_\mathrm{a}, x_\mathrm{b}\big)\,. \tag{19}$$

The probability $P_\mathcal{S}$ of detecting the correct source location can be derived using principles of statistical room acoustics [23, 24]. In practice, $P_\mathcal{S}$ will typically depend on the various structural parameters of the

considered problem such as reverberation time, enclosure volume, the source power and its frequency content. In an attempt to reduce the complexity of the model derivations however, the number of free parameters in the analysis will be cut down by only considering the probability $P_{\mathcal{S}}$ from now on. The numerical values of other model variables then follow either directly from $P_{\mathcal{S}}$ or as a result of assigning specific values to other parameters of lesser importance (such as room volume or state-space limits).

4.2 Observation-Related PCRB Computations

Based on the observation PDF defined in (19), the PCRB computations can now be finalised for the simplified observation model. In (10d), the second term involved in the computation of $\mathbf{D}_{k-1}^{(4)}$ requires the derivation of the term

$$\log p(y_k|\mathbf{x}_k) = \log\big(P_{\mathcal{S}} \cdot \mathcal{N}(y_k; x_k, \sigma_v^2) + (1 - P_{\mathcal{S}}) \cdot \mathcal{U}(y_k; x_{\mathrm{a}}, x_{\mathrm{b}})\big)\,, \tag{20}$$

where the simplification $\mathbf{h}^{\mathrm{T}}\mathbf{x}_k = x_k$ has been made. The negative second-order partial derivatives of (20) can be derived as

$$\begin{aligned} -\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(y_k|\mathbf{x}_k) &= -\,\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}\, \log\bigg(\frac{P_{\mathcal{S}}}{\sigma_v\sqrt{2\pi}} \exp\Big(-\frac{(y_k - x_k)^2}{2\sigma_v^2}\Big) - \frac{1 - P_{\mathcal{S}}}{x_{\mathrm{b}} - x_{\mathrm{a}}}\bigg) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \Lambda(\alpha_k)\,, \end{aligned} \tag{21}$$

where the term $\Lambda(\alpha_k)$ is defined as follows:

$$\Lambda(\alpha_k) = \frac{P_{\mathcal{S}}}{\sigma_v^4 \beta^2 \sqrt{2\pi}} \cdot \frac{\big(\sigma_v - \frac{\alpha_k}{\sigma_v}\big)\beta \exp\big(\frac{\alpha_k}{2\sigma_v^2}\big) + \frac{P_{\mathcal{S}}}{\sqrt{2\pi}}}{\Big(\exp\big(\frac{\alpha_k}{2\sigma_v^2}\big) + \frac{P_{\mathcal{S}}}{\sigma_v \beta \sqrt{2\pi}}\Big)^2}\,, \tag{22}$$

with the constant

$$\beta = \frac{1 - P_{\mathcal{S}}}{x_{\mathrm{b}} - x_{\mathrm{a}}}\,.$$

In (22), the squared observation error α_k at time k is defined as

$$\alpha_k = (y_k - x_k)^2\,.$$

The second term of matrix $\mathbf{D}_{k-1}^{(4)}$ defined in (10d) is hence given by

$$\mathbb{E}\big\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(y_k|\mathbf{x}_k)\big\} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \mathbb{E}\big\{\Lambda(\alpha_k)\big\}, \tag{23}$$

which provides the last formula required for the PCRB computations. Together with the results previously given in (16) (which are reproduced here for convenience), the final computations of the PCRB recursion matrices can be summarised as follows:

$$\mathbf{D}_{k-1}^{(1)} = \mathbf{G}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{G}\,, \tag{24a}$$

$$\mathbf{D}_{k-1}^{(2)} = -\mathbf{G}^{\mathrm{T}}\mathbf{Q}^{-1}\,, \tag{24b}$$

$$\mathbf{D}_{k-1}^{(3)} = -\mathbf{Q}^{-1}\mathbf{G}\,, \tag{24c}$$

$$\mathbf{D}_{k-1}^{(4)} = \mathbf{Q}^{-1} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \mathbb{E}\big\{\Lambda(\alpha_k)\big\}. \tag{24d}$$

The expectation in (24d) can be computed using Monte Carlo averaging over a large number of realisations of $\Lambda(\alpha_k)$. In other words, the sequence $\big\{\Lambda(\alpha_k)\big\}_{k=1}^{K}$ is numerically computed according to (22) for a large number N of trajectories $\big\{x_k^{(n)}\big\}_{k=1}^{K}, n = 1, \ldots, N$, and, for each trajectory n, a large number M of observations $\big\{y_k^{(n,m)}\big\}_{k=1}^{K}, m = 1, \ldots, M$. The value $\mathbb{E}\{\Lambda(\alpha_k)\}$ then results as the average taken over all the $\big\{\Lambda(\cdot)\big\}_{k=1}^{K}$ realisations:

$$\mathbb{E}\big\{\Lambda(\alpha_k)\big\} \approx \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} \Lambda\big(\alpha_k^{(n,m)}\big)\,,$$

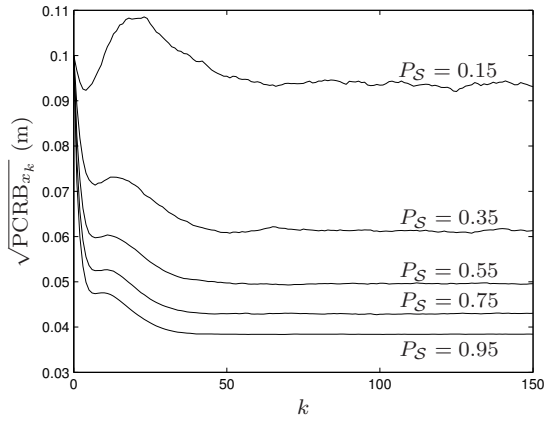

Figure 3. Example of square-root PCRB on estimation error for x_k, using the simplified observation model. Results obtained for $q = 5.1 \cdot 10^{-4}\,\mathrm{m}^2/\mathrm{s}^2$ and with a variable detection probability $P_{\mathcal{S}}$.

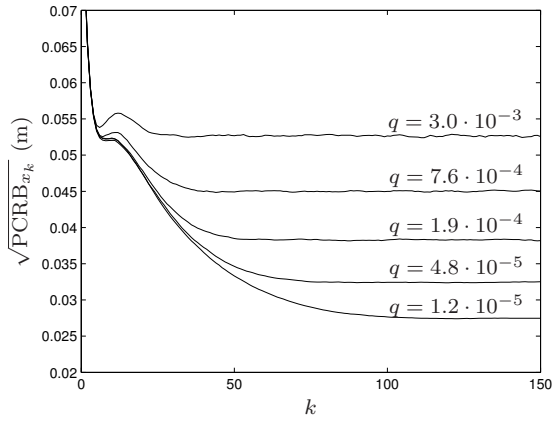

Figure 4. Example of square-root PCRB on estimation error for x_k, using the simplified observation model. Results obtained with a variable process noise parameter q (in $\mathrm{m}^2/\mathrm{s}^2$) and $P_{\mathcal{S}} = 0.75$.

with

$$\alpha_k^{(n,m)} = \left(y_k^{(n,m)} - x_k^{(n)}\right)^2.$$

The large number of trajectories and observations required for this process are typically simulated using (11) and (18), respectively.

4.3 PCRB Results

Some practical PCRB results are now presented based on the theory developed so far. This is done by considering a numerical example with the following simulation parameters. The target is assumed to be in motion within the state-space limits $x_{\mathrm{a}} = 0\,\mathrm{m}$ and $x_{\mathrm{b}} = 10\,\mathrm{m}$. The distribution of the initial target state is assumed Gaussian, i.e. $\mathbf{x}_0 \sim \mathcal{N}(\overline{\mathbf{x}}_0, \mathbf{P}_0)$. The original target position is set to $x_0 = 2\,\mathrm{m}$ with an initial target velocity $\dot{x}_0 = 0.3\,\mathrm{m/s}$, i.e. $\overline{\mathbf{x}}_0 = [2 \;\; 0.3]^{\mathrm{T}}$, and the covariance matrix of the original source distribution is defined with the following numerical values:

$$\mathbf{P}_0 = \begin{bmatrix} \sigma_{x_0}^2 & 0 \\ 0 & \sigma_{\dot{x}_0}^2 \end{bmatrix} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.04 \end{bmatrix}.$$

Following a standard practice, the initial Fisher information matrix at time $k = 0$ in (9) is defined by the original state covariance matrix, i.e. $\mathbf{J}_0 = \mathbf{P}_0^{-1}$. The standard deviation σ_v of the observation noise process is set to 0.1 m.

When using Monte Carlo averaging to compute the expectations of (24d), a total of 2500 simulation runs were used, corresponding to the average over 50 simulations of $\{\mathbf{x}_k\}_{k=1}^K$ using (11), and for each of

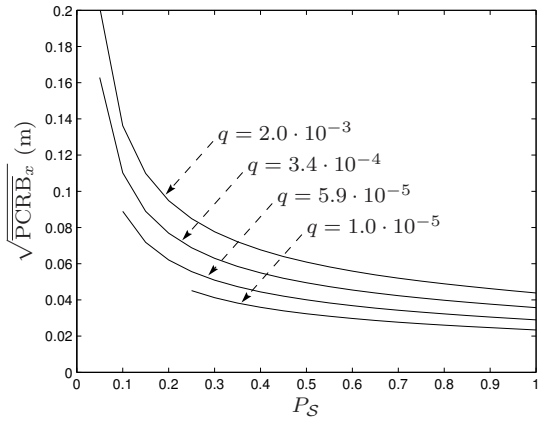

Figure 5. Steady-state square-root PCRB for x_k vs. target detection probability P_S, with variable process noise parameter q (in $\mathrm{m}^2/\mathrm{s}^2$).

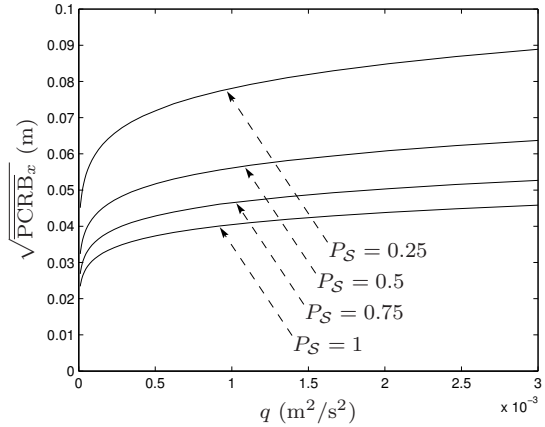

Figure 6. Steady-state square-root PCRB for x_k vs. process noise q, with variable probability parameter P_S.

these trajectories, 50 simulations of the observation sequence $\{y_k\}_{k=1}^{K}$ computed according to (18). The total number of simulation steps K was arbitrarily set to 150, and the update interval was finally defined as $T = 0.05\,\mathrm{s}$.

Figures 3 and 4 present some typical examples of the obtained results. Figure 3 depicts $\sqrt{\mathrm{PCRB}_{x_k}}$ as a function of the time index k and for a variable probability P_S. The different curves in this plot were all obtained with a fixed noise parameter $q = 5.1 \cdot 10^{-4}\,\mathrm{m}^2/\mathrm{s}^2$. Figure 4 shows the same PCRB results obtained with a variable process noise parameter q, with each curve computed using a detection probability P_S set to 0.75.

Due to the relative complexity of the PCRB analysis presented in this work, it is generally difficult to compute the limit $\mathbf{J}_\infty$ analytically as $k \to \infty$. However, as clearly depicted in Figures 3 and 4, the approximate steady-state value $\overline{\mathrm{PCRB}}_x \triangleq \lim_{k\to\infty} \mathrm{PCRB}_{x_k}$ can be easily estimated from the plots. Figures 5 and 6 present the results obtained for $\sqrt{\overline{\mathrm{PCRB}}_x}$, i.e. the steady-state square-root PCRB value for parameter x_k. Figure 5 shows the dependence of the steady-state lower bound as a function of the detection probability P_S, for a few values of the free parameter q. Figure 6 shows the same PCRB results with the role of P_S and q inverted.

The results presented in this section confirm the obvious expectation that higher values of system noise and lower detection probabilities yield a higher theoretical bound on estimation error. Both these conditions explicitly correspond to an increase of the overall disturbance level in the system, meaning that the target position estimates delivered by any tracking algorithm can only become less accurate.

5 Correlated Observations Model

In the following developments, we propose to extend the previous analysis of Section 4 to yield an observation model that is a better representation of what happens in practice. In particular, the effects of the correlation between sound intensity measurements in a reverberant sound field is introduced into this framework. The following subsection explains the reasons motivating this model update, and then presents the theoretical concepts leading to the new observation equation.

5.1 Theoretical Developments

A well-known property of diffuse sound fields is the spatial correlation existing between the sound pressure values $p(t)$ measured at two different locations in the state space, a distance r apart [15, 23]. For the case of a diffuse sound field consisting of one single tone frequency f, the correlation coefficient $\rho_p(r)$ for sound pressure results in

$$\rho_p(r) = \frac{\sin(\kappa r)}{\kappa r}\,,$$

where the wave number is defined as $\kappa = 2\pi f/c$, with c the propagation speed of acoustic waves. For observations based on the sound intensity $I(t) \propto p^2(t)$, such as those derived from the SBF principle, the correlation coefficient $\rho_I(r)$ can be shown to be

$$\rho_I(r) = \frac{\sin(2\kappa r)}{2\kappa r}\,. \tag{25}$$

The sound field in an enclosure can be analysed using the principles of geometrical room acoustics, i.e. with the concept of sound *waves* replaced by the concept of sound *rays* propagating in straight lines [17, 23]. Thus, the intensity measured at a specific point in the enclosure can be viewed as the result of a multitude of sound rays originating from the acoustic source, bouncing off the walls a number of times, and then concentrating again at the receiver position. The correlation existing between intensity values at two slightly different locations can be seen consequently as coming from the "same" rays propagating through space along slightly different trajectories to reach one or the other receiver position. From this point of view, it does not matter whether the difference in the paths followed by the sound rays is introduced through two different positions at the receiver or at the source. By reciprocity, the same correlation therefore exists when the sound intensity is measured at the same location but with a slightly different source position.[3] Now assuming that a period of time has elapsed between the two consecutive intensity measurements (corresponding typically to the time required for the source to move from one location to the other), the velocity of the source can be finally linked to a *temporal* correlation existing between measurements.

The temporal correlation $\rho_I(\tau) \triangleq \rho_I\big(\tau(r)\big)$ between consecutive sound intensity values directly follows from (25) with the variable r now corresponding to the source (target) displacement: $r = \dot{x}_\mathrm{s}\,\tau$, where $\dot{x}_\mathrm{s}$ denotes the source velocity. In the case of band-limited noise, the frequency-averaged correlation coefficient $\bar{\rho}_I(\tau)$ simply results from averaging $\rho_I(\tau)$ over the frequency band of interest:

$$\bar{\rho}_I(\tau) = \frac{1}{\kappa_\mathrm{b} - \kappa_\mathrm{a}} \int_{\kappa_\mathrm{a}}^{\kappa_\mathrm{b}} \frac{\sin(2\kappa\dot{x}_\mathrm{s}\tau)}{2\kappa\dot{x}_\mathrm{s}\tau}\,\mathrm{d}\kappa\,. \tag{26}$$

Figure 7 shows, for a few values of $\dot{x}_\mathrm{s}$, examples of temporal correlation coefficient for sound intensity computed as in (26) with frequency averaging over $f \in [300\,\mathrm{Hz}, 3000\,\mathrm{Hz}]$.

Note that in practice, the correlation between sound intensity values is directly affected by the impulse responses in the enclosure. In the case of a heavily reverberant room, the acoustic impulse response is likely to be non-minimum phase, and as a result, much more likely to give erroneous measurements via the SBF process.

The original assumptions given previously in Section 4.1 remain valid when introducing the concept of temporal correlation into the observation model of (18). In particular, the observation process is

[3] This result can also be explained as follows. The spatial correlation defined in (25) results from the similarity between the transfer functions from the source position to each of the receiver locations. The reciprocity hence follows from the fact that the same impulse response is obtained from one room location to another regardless of which position is occupied by the source.

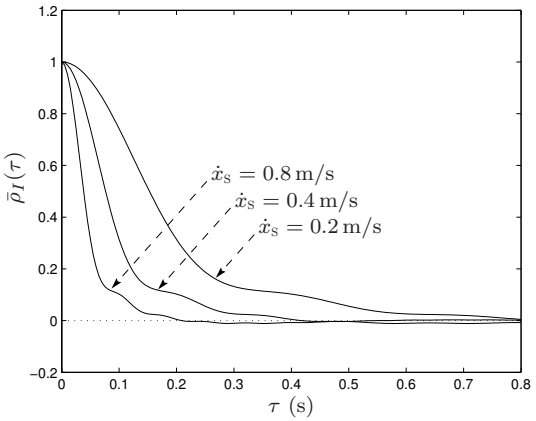

Figure 7. Temporal correlation coefficient for sound intensity, with variable target velocity $\dot{x}_\mathrm{S}$.

still governed by the detection probability $P_\mathcal{S}$, and the resulting observation y_k in a misdetected target situation is here also assumed to be uniformly distributed across the state space. Consequently, the general observation model and observation PDF in (18) and (19) also remain identical for the current developments. The major difference however appears in the sequential occurrences of the hypotheses $\mathcal{S}$ and $\mathcal{F}$. Whereas this process was defined as purely random so far (under the constraint that $\mathcal{S}$ would occur with probability $P_\mathcal{S}$ on average), it is now to be related to the temporal correlation $\bar{\rho}_I(\tau)$ of the sound intensity values measured across the state space. A practical example of this principle is as follows. Assume that at time k, a spurious peak in the SBF output generates an erroneous observation $y_k \neq x_{\mathrm{s},k}$, i.e., the observation fails to detect the real target. This practically means that at least one sound intensity measurement in the reverberant part of the sound field registers a value larger than the output of a beamformer aimed at the true source position. In a situation where a large temporal correlation exists between successive measurements, the sound intensity recorded at the spurious peak location will remain large for a few successive frames, making it likely for hypothesis $\mathcal{F}$ to occur consistently for several frames at a time, rather than in a purely random fashion. The updated observation model attempts to take this concept into account when generating a series of observations. In the following simulations, realisations of the correlated observations process are generated by filtering a white noise sequence with an FIR filter implemented to yield a temporal correlation as defined by (26).

Figure 8 shows three realisations of the observation process derived in this section. It graphically shows the specific measurement values y_k as the source moves along the state space. The '$\circ$' markers correspond to the specific values of y_k for a situation where the target is correctly detected (hypothesis $\mathcal{S}$), whereas '$\times$' markers are observations resulting from a misdetection. The numerical settings leading to this simulation result were similar to previous examples, with the state-space limits set to $x_\mathrm{a} = 0$ and $x_\mathrm{b} = 10\,\mathrm{m}$, and the observation noise level defined as $\sigma_v = 0.1\,\mathrm{m}$. All three plots in Figure 8 were obtained with a fixed detection probability $P_\mathcal{S} = 0.5$. This figure clearly demonstrates the increasing level of correlation in the measurements as the velocity $\dot{x}_\mathrm{S}$ of the source decreases. Also, note the similarity of these results compared to those obtained in Figures 1 and 2.[4]

Finally, note that in practical situations, the effects of the temporal correlation derived in this section can be expected to vary depending on the target's location in the state space. As demonstrated in Figure 1 for instance, the measurements will typically become increasingly reliable as the source becomes very close to the sensor array. For simplicity however, this distinction will not be made for the experimental simulations, and in that sense, the following developments can be considered as a worst case analysis.

5.2 Simulation Results

In Section 6, the performance of a tracking algorithm for ASLT will be investigated with respect to the level of correlation in the sound intensity measurements, i.e., with respect to the target's speed in the state space. In order to relate these tracking results to the presented lower bound derivations, the PCRB

[4]The GCC computations effectively correspond to the output of a two microphone steered beamformer.

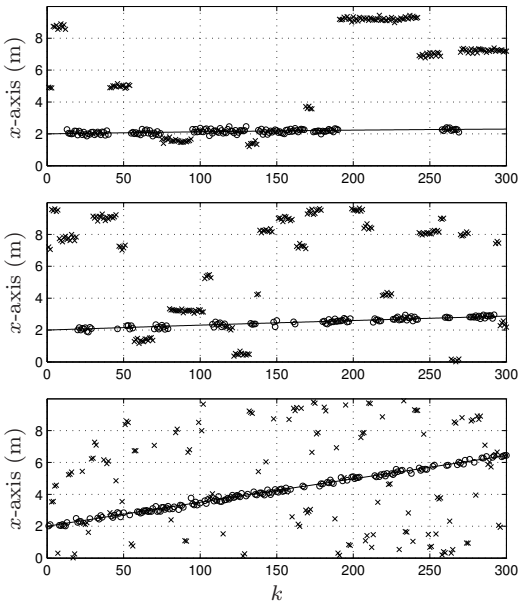

Figure 8. Realisations of the correlated observations process. Solid lines represent the target trajectory along the state space, '$\circ$' markers denote observations for a correctly detected target, '$\times$' markers are erroneous observations. The source velocity $\dot{x}_\mathrm{s}$ is, from top to bottom, 0.05, 0.1 and 0.5 m/s. In all three cases, the detection probability is $P_\mathcal{S} = 0.5$.

parameter is now computed with respect to the same velocity parameter. To this purpose, we use the correlated observations model derived in the previous subsection. It must be noted however that the same results would also be obtained when using the simplified model of Section 4.1, as a result of the whiteness assumption made by the PCRB theory regarding the system noise processes in (6).

In order to compute the PCRB as a function of the target velocity parameter $\dot{x}_\mathrm{s}$, a constant-velocity source motion is necessary. Since the process noise might slowly change the speed of the target over time (see the system transition defined by (11) and (12)), this would in turn imply setting $q = 0$. However, a non-zero system noise is necessary for the PCRB to reach a steady-state value different from zero. A compromise is reached by setting q to a small constant, $q = 10^{-5}\,\mathrm{m}^2/\mathrm{s}^2$, and the target velocity defined in the initial state $\overline{\mathbf{x}}_0 = [x_0 \;\; \dot{x}_\mathrm{s}]^\mathrm{T}$ is hence only marginally altered over time.

Figure 9 depicts the steady-state square-root PCRB value $\sqrt{\overline{\mathrm{PCRB}}_x}$ for the position estimates as a function of $\dot{x}_\mathrm{s}$ for the correlated observations model. These simulation results were obtained with model parameters defined as previously described in Section 4.3, except for the upper state-space limit set here to $x_\mathrm{b} = 14\,\mathrm{m}$, and the update interval now defined as $T = 0.03\,\mathrm{s}$. Monte Carlo averaging was carried out over 400 different simulation runs, corresponding to 20 source trajectories with 20 realisations of the correlated observations sequence each. The results in Figure 9 serve as a confirmation of the fact that the PCRB criterion used in this work does not account for the level of temporal correlation in the realisations of the observation noise process. In other words, a lower bound analysis based on the PCRB parameter states that the correlation between measurements should theoretically not affect the performance of a tracking algorithm.

6 Comparison with Tracking Algorithm

Despite the PCRB results obtained in the previous section, it should be clear that on the basis of a single realisation of the target trajectory, a tracking algorithm is obviously very likely to be mislead by a long series of coherent false detections (see, e.g., top plots in Figures 8 and 2). Thus, for practically relevant situations, the tracking accuracy can be expected to depart from the results depicted in Figure 9. From this point of view, there exists a relevant difference between the PCRB and the type of performance assessment which is of real interest in practice. To support this claim, the performance of an acoustic source tracking algorithm is assessed using the specific ASLT models presented in this work. The consid-

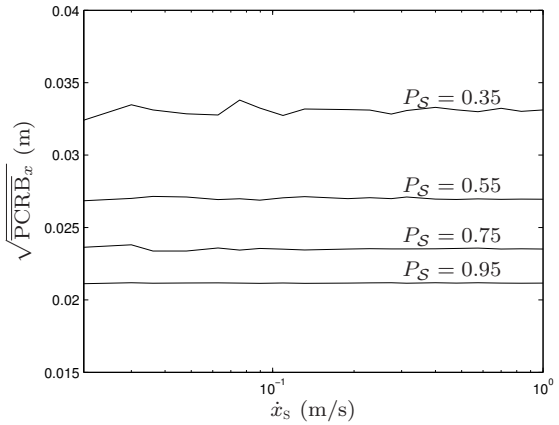

Figure 9. Steady-state square-root PCRB vs. target velocity $\dot{x}_\mathrm{S}$ computed for correlated observations, with variable detection probability $P_\mathcal{S}$.

ered tracking method is based on a particle filtering (PF) approach and is described briefly in the next subsection.

6.1 PF Algorithm

The particle filter to be investigated is based on the SBF-GL method (steered beamforming with Gaussian likelihood) proposed by Ward *et al.* in [19], and the reader is referred to this literature work for a detailed algorithm description. More information on particle filtering can also be found, e.g., in [25, 26]. The PF approach requires the definition of two main concepts. The first one is the transition equation, which is here defined as

$$\mathbf{x}_k = \begin{bmatrix} 1 & aT \\ 0 & a \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} bT \\ b \end{bmatrix} u_k \,, \tag{27a}$$

$$u_k \sim \mathcal{N}(0,1), \tag{27b}$$

with the model parameters defined as

$$a = \exp(-\beta\, T)\,,$$
$$b = \bar{v}\,\sqrt{1-a^2}\,,$$

where $\bar{v}$ is the steady-state velocity parameter, and β the rate constant. This type of dynamics model is commonly found in the acoustic source tracking literature [19, 20]. The second concept involved in the PF algorithm is the likelihood function (measurement PDF) $p(y_k|\mathbf{x}_k)$. The current PF implementation makes use of a Gaussian likelihood with background probability, defined as follows:

$$p(y_k|\mathbf{x}_k) = (1-\psi)\cdot\mathcal{N}(y_k; x_k, \sigma_y^2) + \psi\cdot\mathcal{U}(y_k; x_\mathrm{a}, x_\mathrm{b})\,, \tag{28}$$

where the small probability ψ accounts for the fact that the measurements might be erroneous due to disturbances.

It must be noted that the transition and observation model definitions in (27) and (28) are not directly related to the definitions given in (11) and (18) and used as a basis for the PCRB derivations. In the context of particle filtering, the system equations constitute a generic model of the considered problem, and typically have to be broad enough to include all the potential movements of the considered target (including those resulting from (11)) in order to achieve successful tracking results.

In the following simulations, the various PF parameters were optimised by hand and set as follows: $\sigma_y = 0.2\,\mathrm{m}$, $\psi = 0.2$, and the number of particles is set to 100. These settings were found to yield good results for the considered example. The transition model parameters are defined as $\bar{v} = 0.5\,\mathrm{m/s}$ and $\beta = 10\,\mathrm{Hz}$. Finally, as commonly done when assessing the tracking performance of PF algorithms, the particle set is consistently initialised at the location corresponding to the initial target state x_0 prior to each simulation, in order to avoid undesired track acquisition effects.

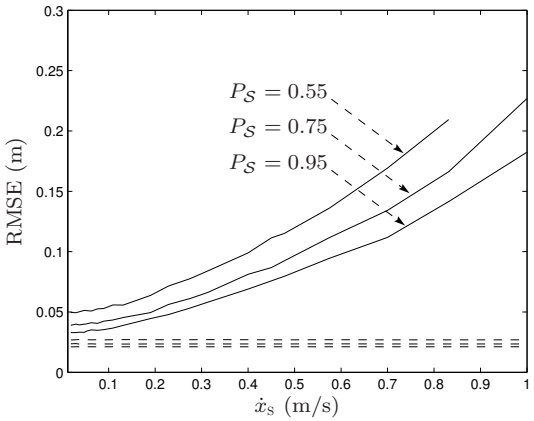

Figure 10. RMSE performance results for PF algorithm with simplified observation model. Dashed lines: PCRB results reproduced from Figure 9 (for corresponding values of $P_\mathcal{S}$).

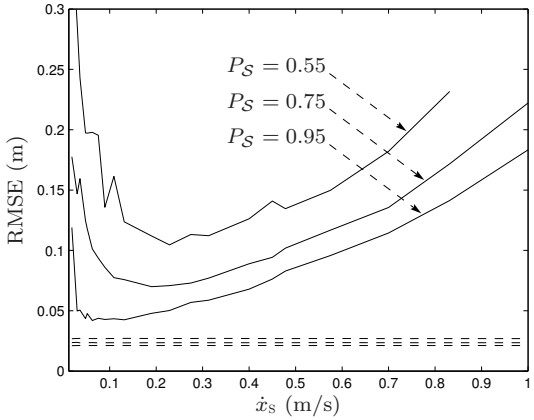

Figure 11. RMSE performance results for PF algorithm with correlated observations model. Dashed lines: PCRB results reproduced from Figure 9 (for corresponding values of $P_\mathcal{S}$).

6.2 Simulation Results

The simulations in this section were carried out using the same system parameters as defined in Section 5.2, with $q = 10^{-5}\,\mathrm{m}^2/\mathrm{s}^2$, $x_\mathrm{b} = 14\,\mathrm{m}$ and $T = 0.03\,\mathrm{s}$. The tracking performance of the PF algorithm was determined for both the simplified and the correlated observations models, allowing one to clearly distinguish the influence of the temporal correlation between measurements from the influence of other parameters. The PF simulations were carried out in the same way as for the PCRB computations, i.e. the tracking accuracy of the algorithm results as the average error over a series of 400 Monte Carlo simulations over $K = 150$ time steps, using either of the two observation models under scrutiny.

Figures 10 and 11 present the tracking results for the PF method for the simplified and correlated observations models respectively. These plots show the average tracking results in terms of the RMSE parameter (root mean squared error). For each case, the results have been computed with the detection probability $P_\mathcal{S}$ as variable parameter. The PCRB results of Figure 9 are also reproduced in the RMSE plots (dashed lines) to allow for a direct comparison with the PF performance results.

6.3 Result Analysis

The tracking performance results obtained with non-correlated observations (Figure 10) conform with the way the PF algorithm operates. This method is based on a fixed-parameter dynamics model, involving notably the steady-state velocity constant $\bar{v}$, which determines the mobility of the particles in the state space. As the velocity of the source increases, it becomes more and more difficult for the tracker to "keep up" with the target and the overall performance of the filter decreases, as reflected by the RMSE

parameter.

For the PF simulation results using the correlated observations model (Figure 11), the overall tracking performance clearly decreases as the target velocity reaches low values, which confirms the expectation that the temporal correlation between sound intensity measurements is detrimental to the tracking accuracy in practice, even though the tracker is initialised at the correct target position. As mentioned at the beginning of this section, long series of coherent false detections increase the chance of the tracking algorithm being mislead, which might ultimately result in a total track loss. Also as expected, it can be seen from a comparison of Figures 10 and 11 that the effects of the correlation slowly disappear as $\dot{x}_\mathrm{S}$ increases, as the PF's average RMSE performance tends towards the results obtained for the non-correlated scenario for large $\dot{x}_\mathrm{S}$.

In respect to the previously computed PCRB, the parameter $\sqrt{\overline{\mathrm{PCRB}}_x}$ theoretically sets a lower limit on the RMSE values resulting from the tracking algorithm. Comparing the PCRB and RMSE results in Figure 10, for corresponding values of $P_\mathcal{S}$, indeed shows that the inequality $\mathrm{RMSE} \geqslant \sqrt{\overline{\mathrm{PCRB}}_x}$ is verified. This comparison also illustrates the fact that for most values of $P_\mathcal{S}$, there exists a significant gap between the tracking performance of the current PF implementation and the theoretically achievable minimum estimation error. In Figure 11, the discrepancy between RMSE and PCRB results is obviously also due to the fact that the considered PCRB criterion is unable to account for the effects of a non-white measurement sequence.

7 Conclusion

The analysis presented in this paper originates from a desire to theoretically gauge the performance of acoustic source tracking methods in reverberant environments with respect to a minimum error bound on the target position estimates. To this aim, a model of the considered ASLT problem was presented to simulate the effects of reverberation on acoustic measurements obtained with a steered beamformer. Experimental simulations using this model have demonstrated that the temporal correlation existing between consecutive sound intensity measurements in a reverberant sound field is detrimental to the accuracy of a practical tracking algorithm. In contrast, this correlation effect was shown not to influence the error bound derived on the basis of the posterior Cramér-Rao bound theory, which raises the concern that this parameter is not completely appropriate for the purpose of a practical assessment of the ASLT problem. This fact hence highlights the need, in order to yield a complete understanding of the considered problem, for a lower bound theory that would account for systems with non-white noise processes, which represents a potentially complex task. The different concepts presented in this paper (such as, for instance, the reverberant observations model) can be used in future research as a basis for the derivation of such a lower bound theory in the context of ASLT.

Acknowledgement

National ICT Australia is funded by the Australian Government's Backing Australia's Ability initiative, in part through the Australian Research Council.

References

[1] H. L. Van Trees. *Detection, estimation, and modulation theory (part I).* John Wiley & Sons, Inc., New York, 1968.

[2] N. Bergman. Posterior Cramér-Rao bounds for sequential estimation. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo methods in practice*, pages 321–338. Springer-Verlag, New York, 2000.

[3] J. H. Taylor. The Cramér-Rao estimation error lower bound computation for deterministic nonlinear systems. *IEEE Transactions on Automatic Control*, AC-24(2):343–344, April 1979.

[4] J. S. Abel. A bound on mean-square-estimate error. *IEEE Transactions on Information Theory*, 39(5):1675–1680, September 1993.

[5] P. Tichavský, C. H. Muravchik, and A. Nehorai. Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46(5):1386–1396, May 1998.

[6] T. H. Kerr. Status of CR-like lower bounds for nonlinear filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 25(5):590–601, September 1989.

[7] J. C. Chen, K. Yao, T.-L. Tung, C. W. Reed, and D. Chen. Source localization and tracking of a wideband source using a randomly distributed beamforming sensor array. *International Journal of High Performance Computing Applications*, 16(3):52–65, August 2002.

[8] J. C. Chen, R. E. Hudson, and K. Yao. Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. *IEEE Transactions on Signal Processing*, 50(8):1843–1854, August 2002.

[9] B. Ristic, M. S. Arulampalam, and C. Musso. The influence of communication bandwidth on target tracking with angle only measurements from two platforms. *Signal Processing*, 81(9):1801–1811, September 2001.

[10] B. Ristic, S. Arulampalam, and C. Musso. On Cramér-Rao bounds for sequential angle-only target motion analysis. In *Proceedings of the Third Australian Workshop on Signal Processing Applications*, Brisbane, QLD, Australia, December 2000.

[11] B. Ristic and M. S. Arulampalam. Tracking a manoeuvring target using angle-only measurements: algorithms and performance. *Signal Processing*, 83(6):1223–1238, June 2003.

[12] B. Ristic, A. Farina, D. Benvenuti, and M. S. Arulampalam. Performance bounds and comparison of nonlinear filters for tracking a ballistic object on re-entry. *IEE Proceedings – Radar, Sonar and Navigation*, 150(2):65–70, April 2003.

[13] C. Hue, J.-P. Le Cadre, and P. Pérez. Performance analysis of two sequential Monte Carlo methods and posterior Cramér-Rao bounds for multi-target tracking. In *Proceedings of the Fifth International Conference on Information Fusion*, volume 1, pages 464–473, Annapolis, MD, USA, July 2002.

[14] N. Bergman, L. Ljung, and F. Gustafsson. Point-mass filter and Cramer-Rao bound for terrain-aided navigation. In *Proceedings of the 36th IEEE Conference on Decision and Control*, volume 1, pages 565–570, San Diego, CA, USA, December 1997.

[15] R. K. Cook, R. V. Waterhouse, R. D. Berendt, S. Edelman, and M. C. Thompson. Measurement of correlation coefficients in reverberant sound fields. *Journal of the Acoustical Society of America*, 27(6):1072–1077, November 1955.

[16] C. H. Knapp and G. C. Carter. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, August 1976.

[17] J. B. Allen and D. A. Berkeley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943–950, April 1979.

[18] B. Ristic, M. S. Arulampalam, and N. Gordon. *Beyond the Kalman filter: particle filters for tracking applications*. Artech House, Boston – London, 2004.

[19] D. B. Ward, E. A. Lehmann, and R. C. Williamson. Particle filtering algorithms for tracking an acoustic source in a reverberant environment. *IEEE Transactions on Speech and Audio Processing*, 11(6):826–836, November 2003.

[20] J. Vermaak and A. Blake. Nonlinear filtering for speaker tracking in noisy and reverberant environments. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 3021–3024, Salt Lake City, UT, USA, May 2001.

[21] E. A. Lehmann, D. B. Ward, and R. C. Williamson. Experimental comparison of particle filtering algorithms for acoustic source localization in a reverberant room. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 177–180, Hong Kong, China, April 2003.

[22] R. V. Waterhouse. Statistical properties of reverberant sound fields. *Journal of the Acoustical Society of America*, 43(6):1436–1444, 1968.

[23] H. Kuttruff. *Room acoustics*. Spon Press, London – New York, 2000.

[24] L. L. Beranek. *Acoustics*. McGraw-Hill, New York, 1954.

[25] N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings on Radar and Signal Processing*, 140(2):107–113, April 1993.

[26] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, February 2002.