

Multi Target Tracking with Acoustic Power Measurements using Emitted Power Density

Umut Orguner and Fredrik Gustafsson

Division of Automatic Control

Department of Electrical Engineering

Linköping University

581 83 Linköping Sweden

umut.fredrik@isy.liu.se

Abstract – *This paper presents a method to achieve multi target tracking using acoustic power measurements obtained from an acoustic sensor network. We first present a novel concept called emitted power density (EPD) which is an aggregate information state that holds the emitted power distribution of all targets in the scene over the target state space. It is possible to find prediction and measurement update formulas for an EPD which is conceptually similar to a probability hypothesis density (PHD). We propose a Gaussian process based representation for making the related EPD updates using Kalman filter formulas. These updates constitute a recursive EPD-filter which is based on the discretization of the position component of the target state space. The results are illustrated on a real data scenario where experiments are done with two targets constrained to a road segment.*

Keywords: Multiple target tracking, filtering, estimation, acoustic sensor, power, point mass filter, Gaussian process, probability hypothesis density.

1 Introduction

Comparing received signal strength (RSS) measurements in a sensor network enables a unified framework for localization and tracking with a moderate requirement on network synchronization and communication bandwidth. Furthermore, it applies to a variety of signal energy measurements as provided by for instance acoustic, seismic, magnetic, radio, microwave and infrared sensors. Localization from RSS is of course a fairly well studied problem, see the surveys [1–3] and the papers [4, 5]. RSS based localization utilizes the exponential power decay of the involved signals. Dedicated approaches to this problem assume that the path loss exponent is known [4, 5], or include the RSS measurements as a general non-linear relation [6]. Several *ad-hoc* methods to eliminate nuisance parameters have been proposed in this context, including taking pairwise differences or ratios of observations.

Since energy is additive at each sensor, the RSS from different sources cannot be resolved and the framework

is difficult to extend to multiple target tracking. In practice, the exponential signal decay rate implies that the closest target will dominate each sensor observation. One applicable approach is to consider each sensor as a binary proximity sensor as studied in for instance [7]. However, this requires an excessive amount of sensors to get accurate multi target tracking (MTT). The classic MTT approach [8, 9] is based on data association and track handling, where the data association step would again lead to a proximity approximation, where targets that are not the closest one to any sensor are concealed and thus not updated.

MTT ideas with acoustic sensors appeared in the literature extensively for acoustic source tracking and speaker localization with applications like smart video conferencing and human computer interfaces. In part due to requirements of this type of applications, most of the work considers direction of arrival (DOA) [10, 11] and time difference of arrival (TDOA) [12] measurements. The used techniques vary from particle filters [13] to random set based approaches [14]. This type of measurements have also been used in ground target tracking [15, 16] where Kalman filters [15] and joint probabilistic data association (JPDA [8]) based particle filters [16] were used. Although they are more scarce, MTT approaches using power (and/or energy) based measurements also appeared in the literature. When the number of targets is assumed to be known, a maximum likelihood approach is given in [17] for localization. Based on a similar model, but with a time varying number of targets, Bugallo and Djurić have proposed a particle filter with adaptive state dimension in [18].

In this work, a sensor network scenario is considered, where each sensor measures received signal strength (RSS) from one target or the aggregated RSS from several targets. Communication only allows for sending the RSS measurements to the fusion center. Based on this information, we have to find the number of existing targets and given the number, we should provide their state estimates. Mostly inspired by the the recent random set theoretic multiple target tracking methods [14], we here propose a novel density (or intensity), which is

quite similar to a probability hypothesis density (PHD) [19], that we call as *emitted power density* (EPD) (or RSS-density). This density is defined over the state space and unlike a PHD, its Dirac delta functions (impulses) are modulated with the emitted power of the targets. In a way, the integral of (the expectation of) an EPD over a region in state space would give the expected emitted power from that region. Defining the main estimate of the problem as the EPD, we propose a convenient representation for the EPD using Gaussian processes (GPs) (see for example [20]) and derive recursions that resembles a PHD filter. The main characterization of our EPD representation is that it enables almost linear measurement relations between the power measurements and the summary statistics. Note that Mahler has already proposed (C)PHD type filters for superpositional sensors (as in this work) very recently [21] but we believe that by including the emitted powers of the targets into the estimatee (as in the EPD), one can avoid having to estimate them separately. The EPD filter we propose is based on the discretization of the position component of the target state space. We apply our results to tracking two road targets in an acoustic sensor network, and demonstrate that the MTT performance is close to as good as if the two targets were measured separately.

The outline for the remaining parts of the paper is as follows. The MTT problem we are interested in is summarized in Section 2. The main results are given in Section 3 where EPD is defined, recursions for it are given and the target detection and estimation process is detailed. We give the results obtained using the collected real sound data of a field experiment in Section 4. We finalize the paper by drawing some conclusions in Section 5.

2 Problem Formulation

We consider the tracking problem of an unknown number (denoted with N_T) of targets using acoustic power measurements. Suppose we show the target states to be estimated as $\{x_k^j\}_{j=1}^{N_T}$ and suppose also that target states contain the positions of the targets on a Cartesian $x - y$ plane shown as $\{p_k^j\}_{j=1}^{N_T} \triangleq [p_{x,k}^j, p_{y,k}^j]^T$. We assume that the target state dynamics of all targets is characterized by the same known transition density $p(x_k^j | x_{k-1}^j)$ for $j = 1, \dots, N_T$.

We have N_S stationary acoustic sensors placed in the area of interest where the targets can be. We show each sensor with S_m for $m = 1, \dots, N_S$ and their positions are defined as $\mathbf{p}_s^m \triangleq [p_{x,s}^m, p_{y,s}^m]^T$ where the subscript s distinguishes target and sensor position variables. The sensors measure the cumulative acoustic power emitted by the targets modeled as

$$y_k^m = \sum_{j=1}^{N_T} \frac{\Psi^j}{\|\bar{r}_k^{m,j}\|^\alpha} + v_k^m \quad (1)$$

where for simplicity $v_k^m \sim \mathcal{N}(v_k^m; 0, r)$. In the sensor model y_k^m represents the acoustic power measured from

the m th sensor S_m at time k and Ψ^j is the unknown (assumed constant) acoustic power emitted by the j th target and α is the (assumed known) path loss exponent. We define the range vectors $\bar{r}_k^{m,j}$ from the m th sensor to the j th target as

$$\bar{r}_k^{m,j} \triangleq \left[p_{x,k}^j - p_{x,s}^m \quad p_{y,k}^j - p_{y,s}^m \right]^T. \quad (2)$$

We define the set of measurements obtained at the same time as $Y_k \triangleq [y_k^1, y_k^2, \dots, y_k^{N_S}]^T$. The aim of our MTT algorithm is then to estimate recursively the number and the states of the targets given the data $Y_{0:k}$.

The superpositional sensors as is investigated in this work are problematic with especially the existence of multiple targets. The association solutions in conventional MTT [8, 9] are not relevant since a measurement basically contains information from all targets. Inspired by the recent random set theoretical methods in MTT [14], we below propose the main estimatee for our solution.

Definition 1 (Emitted Power Density) *We define the emitted power density (EPD) of the targets as a density over the state space of the targets as follows.*

$$\text{EPD}_k(x_k) \triangleq \sum_{j=1}^{N_T} \Psi^j \delta_{x_k^j}(x_k) \quad (3)$$

where the notation $\delta_x(\cdot)$ denotes a Dirac delta (impulse) function centered at x . ■

Note that an (expectation of) EPD would be quite similar to a probability hypothesis density (PHD) in that they both involve a summation (superposition) over all the targets. However, in the EPD, each target component is modulated by the emitted power of the target.

In this work, we propose estimating the aggregate target information represented by an EPD from the sensor data instead of enumerating each target separately. For this purpose, we would like to write the measurement equation (1) using the EPDs. In order to do this, we define the following function.

$$h_p(x_k) = \left[\frac{1}{\|\bar{r}^1\|^\alpha} \quad \frac{1}{\|\bar{r}^2\|^\alpha} \quad \dots \quad \frac{1}{\|\bar{r}^{N_S}\|^\alpha} \right]^T \quad (4)$$

where the range vectors \bar{r}^m are defined similar to (2) for $m = 1, \dots, N_S$ using the position component of x_k . We can now easily see that

$$Y_k = \int h_p(x_k) \text{EPD}_k(x_k) dx_k + V_k \quad (5)$$

where $V_k \triangleq [v_k^1, \dots, v_k^{N_S}]^T \sim \mathcal{N}(\mathbf{0}, R_V)$ is the white measurement noise. Hence, in terms of an EPD, our sensors are actually integral sensors. Notice that the function $h_p(x_k)$ depends only on the position component of x_k . Hence, if we are to estimate an EPD based

on the sensor information, the velocity related information in the states (if any) should come through correlations between the position and velocity variables. Keeping an accurate track of this correlation information proved extremely difficult in our studies. In order to overcome this problem we here introduce another measurement variable \tilde{Y}_k which is derived from original acoustic power measurements Y_k in order to update the velocity information as follows.

$$\tilde{Y}_k = \frac{Y_{k+1} - Y_k}{T} \quad (6)$$

where T is the sampling period of the measurements. The quantities \tilde{Y}_k are approximate estimates of the time derivative of Y_k . The approximate relationship about the time derivative of acoustic power and the target velocities can be written as

$$\dot{y}^m = \sum_{j=1}^{N_T} -\frac{\alpha \Psi^j}{\|\tilde{r}^{m,j}\|^{\alpha+2}} (\tilde{r}^{m,j})^T \mathbf{v}^j \quad (7)$$

where $\mathbf{v}^j \triangleq [\mathbf{v}_x^j, \mathbf{v}_y^j]^T$ is the velocity component of the j th target state. Now in order to write the relationship between the measurements \tilde{Y}_k and the EPD, we define the following matrix function

$$h_{\mathbf{v}}(x_k) \triangleq - \begin{bmatrix} \frac{\alpha \tilde{r}^1}{\|\tilde{r}^1\|^{\alpha+2}} & \frac{\alpha \tilde{r}^2}{\|\tilde{r}^2\|^{\alpha+2}} & \cdots & \frac{\alpha \tilde{r}^{N_S}}{\|\tilde{r}^{N_S}\|^{\alpha+2}} \end{bmatrix}^T$$

where the quantities \tilde{r}^m are defined similar to their counterparts above at the position \mathbf{p}_k component of x_k . Then,

$$\tilde{Y}_k = \int h_{\mathbf{v}}(x_k) \mathbf{v}_k \text{EPD}_k(x_k) dx_k + \tilde{V}_k \quad (8)$$

where the measurement noise \tilde{V}_k is approximated to be white and distributed as $\mathcal{N}(\mathbf{0}, R_{\tilde{V}})$. Notice that the function $h_{\mathbf{v}}(x_k)$, similar to $h_{\mathbf{p}}(x_k)$, also depends only on the position component of x_k . The multiplication by the velocity \mathbf{v}_k inside the integration would make these measurements directly informative about the velocity components of EPD as well. Notice that the two types of measurements Y_k and \tilde{Y}_k are obviously not independent sources of information. We are going to take care of this fact by using Y_k and \tilde{Y}_k to update separate parts of an EPD. More specifically, we will use Y_k (\tilde{Y}_k) to update only position (velocity) related components of an EPD.

The aim, in the following parts of this paper, is to obtain an as accurate estimate of EPD as possible given all the measurement $Y_{0:k}$ and $\tilde{Y}_{0:k}$. We are going to do this by finding expected (expectation of) EPDs. Obtaining the actual target estimated position and velocities, then should be accomplished by further processing the estimated EPDs. This is done by finding the peaks of the estimated EPD in the state space of the targets. The amplitudes that the peaks reach will give us an information about the target's emitted power and hence one can easily make a thresholding of the peak amplitudes to decide on target existence etc.

3 Estimation Framework

An EPD is composed of Dirac delta functions (impulses), and its estimate would be a smeared version of the original EPD. This case is similar to PHD. In the literature, there are different representations of PHDs like point masses, Gaussian mixtures and particles etc.. For an EPD, similar representations are possible.

In this paper, we are going to propose a novel EPD representation based on GPs [20]. For this purpose, we are going to assume that our targets are road constrained i.e., has one dimensional position space. Generalizations to 2-D and higher dimensional position spaces are theoretically straightforward and not considered here in this work. Suppose that we have a road segment defined on $x - y$ coordinate axes that we would like to track the targets on. We are only interested in the longitudinal position and velocities of the targets along the road and hence do not care about the lateral positions and velocity of the targets on the road. We define a one dimensional coordinate axis on the road which is denoted by $\eta \in [0, L]$. Here, without losing generality, 0 denotes the origin of the onroad coordinate axis which represents the start of the road segment and the length of the road segment is denoted by $L \in \mathbb{R}$ which represents the end of road segment. We define the onroad state vectors $x_{\eta,k}^j$ of the targets as

$$x_{\eta,k}^j \triangleq \begin{bmatrix} \mathbf{p}_{\eta,k}^j & \mathbf{v}_{\eta,k}^j \end{bmatrix}^T \quad (9)$$

for $j = 1, \dots, N_T$ where $\mathbf{p}_{\eta,k}^j$ and $\mathbf{v}_{\eta,k}^j$ denote the position and the speed of the j th target on the η -coordinates at time k . The global state vectors of the targets are shown by $x_{x-y,k}^j$ which are defined as

$$x_{x-y,k}^j \triangleq \begin{bmatrix} \mathbf{p}_{x,k}^j & \mathbf{p}_{y,k}^j & \mathbf{v}_{x,k}^j & \mathbf{v}_{y,k}^j \end{bmatrix}^T \quad (10)$$

We assume that there exists a suitable coordinate transformation $\mathcal{T}_{x-y,\eta}$ which can be used to transform the onroad target states $x_{\eta,k}^j$ into global target states $x_{x-y,k}^j$. Hence, we have $x_{x-y,k}^j = \mathcal{T}_{x-y,\eta}(x_{\eta,k}^j)$. We assume that targets move according to the nearly constant velocity model given below.

$$x_{\eta,k}^j = \underbrace{\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}}_{\triangleq A} x_{\eta,k-1}^j + \underbrace{\begin{bmatrix} T^2/2 \\ T \end{bmatrix}}_{\triangleq B} w_k \quad (11)$$

where $w_k \sim \mathcal{N}(w_k; 0, q)$ is the process noise.

We define the representation for an EPD as

$$\text{EPD}_k(x_{\eta}) \triangleq \mathbf{w}_k(\mathbf{p}_{\eta}) \delta_{\nu_k(\mathbf{p}_{\eta})}(\mathbf{v}_{\eta}). \quad (12)$$

where $\mathbf{w}_k(\cdot)$ and $\nu_k(\mathbf{p}_{\eta})$ are the following GPs [20] defined on the onroad position coordinates.

$$\mathbf{w}_k(\cdot) \sim \mathcal{GP}(\hat{\mathbf{w}}_k(\cdot), k_w(\cdot, \cdot)) \quad \nu_k(\cdot) \sim \mathcal{GP}(\hat{\nu}_k(\cdot), k_{\nu}(\cdot, \cdot))$$

The quantities $\hat{\mathbf{w}}_k(\cdot)$ and $\hat{\nu}_k(\cdot)$ are the corresponding mean functions. We denote the corresponding covariance kernels as $k_w(\cdot, \cdot)$ and $k_{\nu}(\cdot, \cdot)$. Note that (12) is

just a more convenient representation for an EPD defined in (3) and it is already an approximation by itself. In the representation (12), we have the magnitude of the EPD represented by the GP $\mathbf{w}_k(\cdot)$ as a function of the position coordinate. The velocity related component of the EPD is represented by the GP $\nu_k(\cdot)$ as a function of position as well. Hence for each position value, in loose terms, we keep an amplitude and velocity function determining the characteristics of the EPD. In our problem we would like to find the expected EPDs as below.

$$\text{EPD}_{k|k}(x_\eta) \triangleq E \left[\text{EPD}_k(x_\eta) \middle| Y_{0:k}, \tilde{Y}_{0:k} \right] \quad (13)$$

$$= E \left[\mathbf{w}_k(\mathbf{p}_\eta) \delta_{\nu_k(\mathbf{p}_\eta)}(\mathbf{v}_\eta) \middle| Y_{0:k}, \tilde{Y}_{0:k} \right] \quad (14)$$

$$= \underbrace{E \left[\mathbf{w}_k(\mathbf{p}_\eta) \middle| Y_{0:k}, \tilde{Y}_{0:k} \right]}_{=\hat{\mathbf{w}}_k(\mathbf{p}_\eta)} \underbrace{E \left[\delta_{\nu_k(\mathbf{p}_\eta)}(\mathbf{v}_\eta) \middle| Y_{0:k}, \tilde{Y}_{0:k} \right]}_{=\mathcal{N}(\mathbf{v}_\eta; \hat{\nu}_k(\mathbf{p}_\eta), P_{\hat{\nu}_k(\mathbf{p}_\eta)})} \quad (15)$$

$$= \hat{\mathbf{w}}_k(\mathbf{p}_\eta) \mathcal{N}(\mathbf{v}_\eta; \hat{\nu}_k(\mathbf{p}_\eta), P_{\hat{\nu}_k(\mathbf{p}_\eta)}).$$

Notice that the estimated functions $\hat{\mathbf{w}}_k(\mathbf{p}_\eta)$ and $\hat{\nu}_k(\mathbf{p}_\eta)$ forming our summary statistics for the estimated EPD are continuous. Hence we have to choose another representation for them to be able to store them in the computer. For this purpose (and for the purpose of approximating integrals that will appear later), we deterministically discretize the position component of the η -space as $\{\mathbf{p}_\eta^{(i)}\}_{i=1}^{N_p}$. We assume that the discretization is uniform and we define

$$L_p \triangleq \mathbf{p}_\eta^{(i+1)} - \mathbf{p}_\eta^{(i)} \quad \text{for } i = 1, \dots, N_p - 1. \quad (16)$$

We then define the quantity \mathbf{W}_k as

$$\mathbf{W}_k \triangleq \left[\mathbf{w}_k(\mathbf{p}_\eta^{(1)}) \quad \mathbf{w}_k(\mathbf{p}_\eta^{(2)}) \quad \dots \quad \mathbf{w}_k(\mathbf{p}_\eta^{(N_p)}) \right]^T \quad (17)$$

Then our summary statistics for $\hat{\mathbf{w}}_k(\mathbf{p}_\eta)$ would be $\hat{\mathbf{W}}_k$ (the mean of \mathbf{W}_k) and $P_{\hat{\mathbf{W}}_k}$ (the covariance of \mathbf{W}_k).

The related quantities for $\hat{\nu}_k(\mathbf{p}_\eta)$ would be $\hat{\mathbf{V}}_k$ and $P_{\hat{\mathbf{V}}_k}$ that can be defined for the function $\hat{\nu}_k(\mathbf{p}_\eta)$. Similar to $\text{EPD}_{k|k}(x_\eta)$ one can define the predicted EPD as

$$\text{EPD}_{k|k-1}(x_\eta) \triangleq E \left[\text{EPD}_k(x_\eta) \middle| Y_{0:k-1}, \tilde{Y}_{0:k-1} \right] \quad (18)$$

$$= \hat{\mathbf{w}}_{k|k-1}(\mathbf{p}_\eta) \mathcal{N}(\mathbf{v}_\eta; \hat{\nu}_{k|k-1}(\mathbf{p}_\eta), P_{\hat{\nu}_{k|k-1}(\mathbf{p}_\eta)}). \quad (19)$$

with summary statistics $\hat{\mathbf{W}}_{k|k-1}$, $P_{\hat{\mathbf{W}}_{k|k-1}}$ and $\hat{\mathbf{V}}_{k|k-1}$, $P_{\hat{\mathbf{V}}_{k|k-1}}$.

Suppose that we are given the previous estimated EPD as

$$\begin{aligned} \text{EPD}_{k-1|k-1}(x_\eta) \\ = \hat{\mathbf{w}}_{k-1}(\mathbf{p}_\eta) \mathcal{N}(\mathbf{v}_\eta; \hat{\nu}_{k-1}(\mathbf{p}_\eta), P_{\hat{\nu}_{k-1}(\mathbf{p}_\eta)}). \end{aligned} \quad (20)$$

In the following, we are first going to examine how to do a prediction update, i.e., to obtain (summary statistics of) $\text{EPD}_{k|k-1}$ given (the summary statistics

of) $\text{EPD}_{k-1|k-1}$ in Section 3.1. Then, in Section 3.2 we are going to show our proposed update to be used for obtaining (the summary statistics of) $\text{EPD}_{k|k}$ given (the summary statistics of) $\text{EPD}_{k|k-1}$, which completes a single loop of our EPD filter. The third and last subsection Section 3.3 discusses how to detect targets and calculate their estimates.

3.1 Prediction Update

An important task for a well defined estimation procedure is to define a prediction equation to be used while calculating $\text{EPD}_{k|k-1}$ from $\text{EPD}_{k-1|k-1}$. Assuming that the same targets exist at time $k-1$ and k and there are no new targets at time k , we can derive the following prediction update formula to obtain $\text{EPD}_{k|k-1}(x_\eta)$.

$$\begin{aligned} \text{EPD}_{k|k-1}(x_\eta) = \int p(x_\eta | x_{\eta, k-1}) \\ \times \text{EPD}_{k-1|k-1}(x_{\eta, k-1}) dx_{\eta, k-1} \end{aligned} \quad (21)$$

For space considerations, we had to omit the derivation of (21). See [22] for a proof. Now, in order to obtain $\text{EPD}_{k|k-1}$, one has to substitute the previous EPD of (20) into (21) and then take the integral. We achieve taking this integral using the discretization $\{\mathbf{p}_\eta^{(i)}\}_{i=1}^{N_p}$ of the position component of the η -space. The result of the integral must then be approximated in the same form as (19). This involved derivation and approximation are skipped for space considerations (See [22] for the derivation.) and below we only summarize the result.

$$\begin{aligned} \hat{\mathbf{W}}_{k|k-1} &= K_{\hat{\mathbf{W}}\hat{\mathbf{W}}} \left(K_{\bar{\mathbf{W}}\bar{\mathbf{W}}} + P_{\hat{\mathbf{W}}_{k-1}} \right)^{-1} \hat{\mathbf{W}}_{k-1} \\ P_{\hat{\mathbf{W}}_{k|k-1}} &= K_{\hat{\mathbf{W}}\hat{\mathbf{W}}} - K_{\hat{\mathbf{W}}\bar{\mathbf{W}}} \left(K_{\bar{\mathbf{W}}\bar{\mathbf{W}}} + P_{\hat{\mathbf{W}}_{k-1}} \right)^{-1} K_{\bar{\mathbf{W}}\hat{\mathbf{W}}} \\ \hat{\mathbf{V}}_{k|k-1} &= K_{\hat{\mathbf{V}}\hat{\mathbf{V}}} \left(K_{\bar{\mathbf{V}}\bar{\mathbf{V}}} + P_{\hat{\mathbf{V}}_{k-1}} + qT^2 I_{N_p} \right)^{-1} \hat{\mathbf{V}}_{k-1} \\ P_{\hat{\mathbf{V}}_{k|k-1}} &= K_{\hat{\mathbf{V}}\hat{\mathbf{V}}} - K_{\hat{\mathbf{V}}\bar{\mathbf{V}}} \left(K_{\bar{\mathbf{V}}\bar{\mathbf{V}}} + P_{\hat{\mathbf{V}}_{k-1}} + qT^2 I_{N_p} \right)^{-1} K_{\bar{\mathbf{V}}\hat{\mathbf{V}}} \end{aligned}$$

where the covariance matrices $K_{\bar{\mathbf{W}}\bar{\mathbf{W}}}$, $K_{\hat{\mathbf{W}}\bar{\mathbf{W}}}$, $K_{\bar{\mathbf{W}}\hat{\mathbf{W}}}$, $K_{\hat{\mathbf{W}}\hat{\mathbf{W}}}$ and $K_{\bar{\mathbf{V}}\bar{\mathbf{V}}}$, $K_{\hat{\mathbf{V}}\bar{\mathbf{V}}}$, $K_{\bar{\mathbf{V}}\hat{\mathbf{V}}}$, $K_{\hat{\mathbf{V}}\hat{\mathbf{V}}}$ are generated from the covariance kernels $k_w(\cdot, \cdot)$ and $k_\nu(\cdot, \cdot)$ of the two related GPs respectively. The vectors $\bar{\mathbf{W}}_{k|k-1}$ and $\bar{\mathbf{V}}_{k|k-1}$ are composed of values of the functions $\hat{\mathbf{w}}_{k|k-1}(\cdot)$, $\hat{\nu}_{k|k-1}(\cdot)$ evaluated at the predicted positions $\hat{\mathbf{p}}_{\eta, k|k-1}^{(i)} \triangleq \mathbf{p}_\eta^{(i)} + T\hat{\nu}_{k-1}^{(i)}$ for $i = 1, \dots, N_p$. So the elements of the covariance matrices $K_{\bar{\mathbf{W}}\bar{\mathbf{W}}}$, $K_{\hat{\mathbf{W}}\bar{\mathbf{W}}}$, $K_{\bar{\mathbf{W}}\hat{\mathbf{W}}}$, $K_{\hat{\mathbf{W}}\hat{\mathbf{W}}}$ can be constructed as

$$[K_{\bar{\mathbf{W}}\bar{\mathbf{W}}}]_{i,j} = k_w(\mathbf{p}_\eta^{(i)} + T\hat{\nu}_{k-1}^{(i)}, \mathbf{p}_\eta^{(j)} + T\hat{\nu}_{k-1}^{(j)}) \quad (22)$$

$$[K_{\hat{\mathbf{W}}\bar{\mathbf{W}}}]_{i,j} = k_w(\mathbf{p}_\eta^{(i)}, \mathbf{p}_\eta^{(j)} + T\hat{\nu}_{k-1}^{(j)}) \quad (23)$$

$$[K_{\bar{\mathbf{W}}\hat{\mathbf{W}}}]_{i,j} = k_w(\mathbf{p}_\eta^{(i)} + T\hat{\nu}_{k-1}^{(i)}, \mathbf{p}_\eta^{(j)}) \quad (24)$$

$$[K_{\hat{\mathbf{W}}\hat{\mathbf{W}}}]_{i,j} = k_w(\mathbf{p}_\eta^{(i)}, \mathbf{p}_\eta^{(j)}) \quad (25)$$

The matrices $K_{\bar{\mathbf{V}}\bar{\mathbf{V}}}$, $K_{\hat{\mathbf{V}}\bar{\mathbf{V}}}$, $K_{\bar{\mathbf{V}}\hat{\mathbf{V}}}$, $K_{\hat{\mathbf{V}}\hat{\mathbf{V}}}$ can be similarly calculated using the velocity covariance kernel function $k_\nu(\cdot, \cdot)$ instead.

3.2 Measurement Update

For finding a suitable measurement update, we substitute the representation (12) into (5) to find the relationship of Y_k with the summary statistics.

$$Y_k \triangleq \int h_{\mathbf{p}}(x_\eta) \text{EPD}_k(x_\eta) dx_\eta + V_k \quad (26)$$

$$= \int h_{\mathbf{p}}(\mathbf{p}_\eta) \mathbf{w}_k(\mathbf{p}_\eta) \underbrace{\int \delta_{\nu_k(\mathbf{p}_\eta)}(\mathbf{v}_\eta) d\mathbf{v}_\eta}_{=1} d\mathbf{p}_\eta + V_k$$

$$= \int h_{\mathbf{p}}(\mathbf{p}_\eta) \mathbf{w}_k(\mathbf{p}_\eta) d\mathbf{p}_\eta + V_k \quad (27)$$

Using again the discretization $\{\mathbf{p}_\eta^{(i)}\}_{i=1}^{N_p}$ of the position component of the η -space to take the integral, we can write

$$Y_k = \sum_{i=1}^{N_p} L_p h_{\mathbf{p}}(\mathbf{p}_\eta^{(i)}) \mathbf{w}_k(\mathbf{p}_\eta^{(i)}) + V_k \quad (28)$$

Vectorizing this equation gives the final measurement equation as

$$Y_k = L_p H_p \mathbf{W}_k + V_k \quad (29)$$

where we defined matrix H_p as

$$H_p \triangleq \begin{bmatrix} h_{\mathbf{p}}(\mathbf{p}_\eta^{(1)}) & h_{\mathbf{p}}(\mathbf{p}_\eta^{(2)}) & \dots & h_{\mathbf{p}}(\mathbf{p}_\eta^{(N_p)}) \end{bmatrix}. \quad (30)$$

Hence the measurement Y_k is linearly related to the vector \mathbf{W}_k . One can therefore use Y_k to update the summary statistics $\hat{\mathbf{W}}_{k|k-1}$ and $P_{\hat{\mathbf{W}}_{k|k-1}}$ using a single Kalman filter measurement update. This would unfortunately not affect the summary statistics $\hat{V}_{k|k-1}$ and $P_{\hat{V}_{k|k-1}}$. In order to update those, we will do a similar analysis for the (derived) measurement \tilde{Y}_k below. Note first that an equivalent of (8) can be written in η -coordinates by setting

$$h_{\mathbf{v}}(x_\eta) = - \left[\frac{\alpha \cos(\theta^1)}{\|\tilde{\mathbf{r}}^1\|^{\alpha+1}} \quad \frac{\alpha \cos(\theta^2)}{\|\tilde{\mathbf{r}}^2\|^{\alpha+1}} \quad \dots \quad \frac{\alpha \cos(\theta^{N_S})}{\|\tilde{\mathbf{r}}^{N_S}\|^{\alpha+1}} \right]^T \quad (31)$$

where the quantity θ^m denotes the angle that $\tilde{\mathbf{r}}^m$ makes with the tangent to the road at \mathbf{p}_η . Then,

$$\tilde{Y}_k \triangleq \int h_{\mathbf{v}}(x_\eta) \mathbf{v}_\eta \text{EPD}_k(x_\eta) dx_\eta + \tilde{V}_k \quad (32)$$

$$= \int h_{\mathbf{v}}(\mathbf{p}_\eta) \mathbf{w}_k(\mathbf{p}_\eta) \underbrace{\int \mathbf{v}_\eta \delta_{\nu_k(\mathbf{p}_\eta)}(\mathbf{v}_\eta) d\mathbf{v}_\eta}_{=\nu_k(\mathbf{p}_\eta)} d\mathbf{p}_\eta + \tilde{V}_k$$

$$= \int h_{\mathbf{v}}(\mathbf{p}_\eta) \mathbf{w}_k(\mathbf{p}_\eta) \nu_k(\mathbf{p}_\eta) d\mathbf{p}_\eta + \tilde{V}_k \quad (33)$$

$$= \sum_{i=1}^{N_p} L_p h_{\mathbf{v}}(\mathbf{p}_\eta^{(i)}) \mathbf{w}_k(\mathbf{p}_\eta^{(i)}) \nu_k(\mathbf{p}_\eta^{(i)}) + \tilde{V}_k \quad (34)$$

Now vectorization gives

$$\tilde{Y}_k = L_p H_v [\mathbf{W}_k \otimes \mathbf{V}_k] + \tilde{V}_k \quad (35)$$

where we define the Hadamard (elementwise) product of vectors with the sign \otimes . The matrix H_v is defined similarly to H_p above (with $h_{\mathbf{v}}(x_\eta)$ of (31)). Notice that this second equation is nonlinear in terms of the variables \mathbf{W}_k and \mathbf{V}_k . We are still going use this equation linearly by substituting \mathbf{W}_k by its last estimate $\hat{\mathbf{W}}_k$ to update $\hat{\mathbf{V}}_{k|k-1}$ with a Kalman filter measurement update.

In summary, we propose the following sequential updates

1. We first update $\hat{\mathbf{W}}_{k|k-1}$ and $P_{\hat{\mathbf{W}}_{k|k-1}}$ by only Y_k using (29) as follows.

$$S_{\hat{\mathbf{W}}} \triangleq H_p P_{\hat{\mathbf{W}}_{k|k-1}} H_p^T + R_V \quad (36)$$

$$K_{\hat{\mathbf{W}}} \triangleq P_{\hat{\mathbf{W}}_{k|k-1}} H_p^T S_{\hat{\mathbf{W}}}^{-1} \quad (37)$$

$$\hat{\mathbf{W}}_k = \hat{\mathbf{W}}_{k|k-1} + K_{\hat{\mathbf{W}}}(Y_k - H_p \hat{\mathbf{W}}_{k|k-1}) \quad (38)$$

$$P_{\hat{\mathbf{W}}_k} = P_{\hat{\mathbf{W}}_{k|k-1}} - K_{\hat{\mathbf{W}}} S_{\hat{\mathbf{W}}} K_{\hat{\mathbf{W}}}^T \quad (39)$$

2. We then update $\hat{\mathbf{V}}_{k|k-1}$ and $P_{\hat{\mathbf{V}}_{k|k-1}}$ by only \tilde{Y}_k using (35) after substituting $\hat{\mathbf{W}}_k$ (obtained from the first update above) as the true value of \mathbf{W}_k .

$$S_{\hat{\mathbf{V}}} \triangleq H_v \text{diag}(\hat{\mathbf{W}}_k) P_{\hat{\mathbf{V}}_{k|k-1}} \text{diag}(\hat{\mathbf{W}}_k) H_v^T + R_{\tilde{V}}$$

$$K_{\hat{\mathbf{V}}} \triangleq P_{\hat{\mathbf{V}}_{k|k-1}} \text{diag}(\hat{\mathbf{W}}_k) H_v^T S_{\hat{\mathbf{V}}}^{-1}$$

$$\hat{\mathbf{V}}_k = \hat{\mathbf{V}}_{k|k-1} + K_{\hat{\mathbf{V}}}(\tilde{Y}_k - H_v \text{diag}(\hat{\mathbf{W}}_k) \hat{\mathbf{V}}_{k|k-1})$$

$$P_{\hat{\mathbf{V}}_k} = P_{\hat{\mathbf{V}}_{k|k-1}} - K_{\hat{\mathbf{V}}} S_{\hat{\mathbf{V}}} K_{\hat{\mathbf{V}}}^T$$

where $\text{diag}(\cdot)$ operator forms a diagonal matrix whose diagonal elements are given in the vector argument.

3.3 Target Detection and Estimates

The EPD estimates give a distribution of acoustic power over the state space. It is also evident from the definition of the EPD that the peaks of the EPD will be around the true target states. The easiest proposal for the target detection is then to look for the peaks of the EPD (in the estimated EPD magnitudes in $\hat{\mathbf{W}}_k$ each element of which represents a different position value $\mathbf{p}_\eta^{(i)}$) and declare a target if the value reached at the peak is above a threshold $\gamma_{\mathcal{W}}$. The kernel functions (of the GPs) used in the prediction update forces our EPD estimates to be smooth. Hence, our GP representation generally avoids finding too many peaks. The threshold $\gamma_{\mathcal{W}}$ to detect the targets must generally be selected experimentally considering the possible target sound power levels.

The implementation that we proposed in the previous subsections above has used discretization of the position coordinates. The discretization size must be adjusted according to the computational resources. Searching for the EPD's local maxima over these discrete position values (i.e., in the elements of $\hat{\mathbf{W}}_k$ for different position

values $\mathbf{p}_\eta^{(i)}$) is quite easy. However the discretization $\{\mathbf{p}_\eta^{(i)}\}_{i=1}^{N_p}$ can actually be too coarse an approximation for the true target position states. Hence it is reasonable to actually apply an interpolation and then find the peaks. This would give more smooth state histories for the targets and the use of an interpolation filter would also further smooth the values of \hat{W}_k to avoid too many peaks in a small neighborhood. We here propose to use a Gaussian interpolation window that has zero mean and standard deviation σ_W . Suppose \mathbf{p}_η is any position value. Then the interpolated function $\hat{w}_k(\cdot)$ would be

$$\hat{w}_k(\mathbf{p}_\eta) = \sum_{i=1}^{N_p} [\hat{W}_k]_i \mathcal{N}(\mathbf{p}_\eta, \mathbf{p}_\eta^{(i)}, \sigma_W^2) \quad (40)$$

where the notation $[\hat{W}_k]_i$ denotes the i th element of the vector \hat{W}_k . With this equation, one then can evaluate $\hat{w}_k(\mathbf{p}_\eta)$ on a finer grid and find the peaks easily. The related threshold for the peaks can be selected as $\frac{\gamma W}{\sigma_W}$.

Another idea for obtaining $\hat{w}_k(\cdot)$ on a finer grid is to use the GP property of it. This would also give similar results to above but will lack the extra design parameter σ_W which can be adjusted by the user to suit his/her needs.

4 Example

In this section, we are going to run the proposed EPD filter on some real data collected in an area close to town Skövde, Sweden. The map of the area is shown in Figure 1 along with the road segment information and microphone positions. The road segment information is composed of connected linear smaller road segments. The onroad position coordinates \mathbf{p}_η are marked in the figure at each 50 meters. We have 10 microphones collecting data at 4kHz frequency placed around the interval $300 < \mathbf{p}_\eta < 400\text{m}$. Each microphone position is illustrated with a cross sign in Figure 1.

In this example, we use the synchronized recordings of a motorcycle and a car whose correct positions are measured using GPS sensors. The correct positions of the targets projected onto the road coordinates are shown in Figure 2. The microphone network is also illustrated in Figure 2 with cross signs at $t = 0$ denoting the closest onroad point to each microphone. The recordings for the motorcycle and the car were obtained separately and we obtain our two-target data by adding the sound waveforms for the two cases. It is important to emphasize here that we do not add the two acoustic power waveforms which would make our data biased towards our measurement model. We do the addition with the raw data and then the single power waveform for the two-target case is obtained from the summed raw sound waveform. Hence our measurement model is still objective and it might be invalid if the separate target signatures has common frequency harmonics.

In acoustic power measurement generation, we take the square of the sound waveforms from each micro-

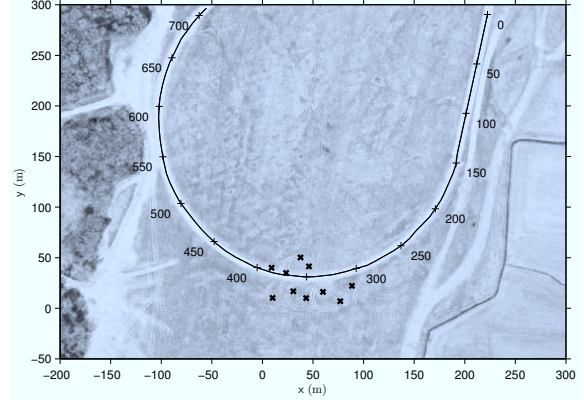


Figure 1: The map of the area, road segment, microphones and coordinates used in the example. The distance markings on the road segment denote the onroad position coordinates \mathbf{p}_η . Microphone positions are illustrated with cross signs.

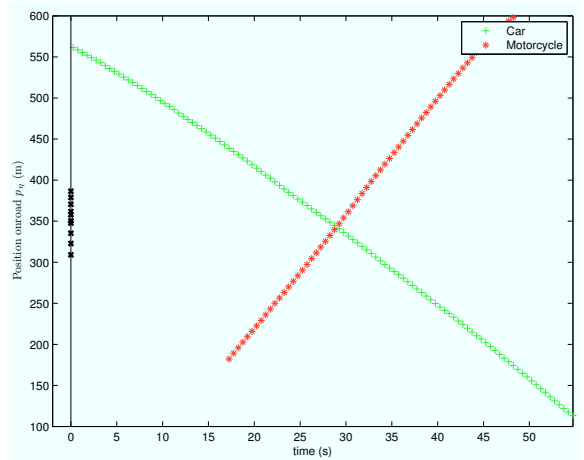


Figure 2: The correct onroad positions of the two targets. The closest onroad point to each microphone is also illustrated with cross signs at $t = 0$.

phone and we obtain the average of the squared samples at each $T = 0.25$ seconds. As an example, we illustrate the raw sound data and the acoustic power measurements generated from it for the microphone located at $[46\text{m}, 41.3\text{m}]$ in Figure 3. In the EPD filter implementation, we have discretized the onroad position coordinates \mathbf{p}_η uniformly at $N_p = 140$ points with 5m distance between the points. We have used the kernel functions

$$k_w(\mathbf{p}_1, \mathbf{p}_2) = 30^2 \exp(-|\mathbf{p}_1 - \mathbf{p}_2|/100) \quad (41)$$

$$k_\nu(\mathbf{p}_1, \mathbf{p}_2) = 5^2 \exp(-|\mathbf{p}_1 - \mathbf{p}_2|/100) \quad (42)$$

which are standard type of kernel functions in GPs [20]. We set the measurement covariances $R_V = 5^2 I_{N_s}$ and $R_{\hat{V}} = \left(\frac{5\sqrt{2}}{T}\right)^2 I_{N_s}$. The state process noise variance was taken as $q = 0.1^2 \text{ (m/s)}^2$. It has been seen that the position grid spacing of 5m's is too coarse and the peaks of the the functions $\hat{w}(\cdot)$ was found on a 1m spaced uni-

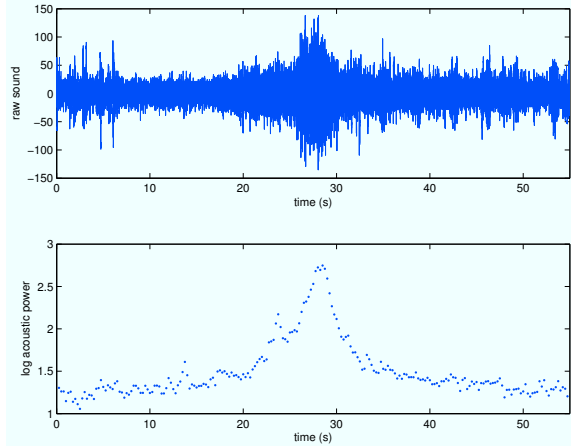


Figure 3: The raw sound data (upper plot) and the acoustic power measurements (lower plot) generated from it for the microphone located at [46m, 41.3m].

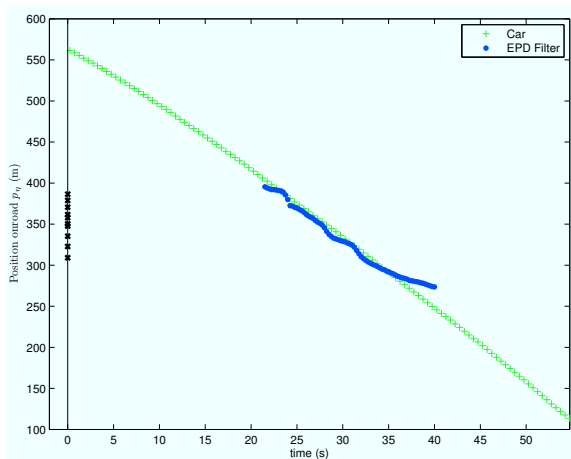


Figure 4: The position estimates of the EPD filter with the single target (car) data.

form grid with a Gaussian interpolation filter of standard deviation $\sigma_w = 10\text{m}$'s. The threshold $\gamma_w = 10$ is selected for target detection. The threshold for the interpolated grid data is then taken as $\frac{\gamma_w}{\sigma_w} = 1$. We have run the EPD with such parameters on the acoustic power measurements for three cases.

1. For car's sound data only;
2. For motorcycle's sound data only;
3. For the superposed sound data of the car and the motorcycle.

The resulting position estimates obtained are illustrated in Figures 4, 5 and 6 respectively. The EPD-filter can easily handle all the cases. One target detection and tracking seems to be quite good except for some occasional missing detections in the motorcycle only case. In the two target case, the target initiation delays a little and target loss happens a little earlier. However, both targets are tracked quite similar to the

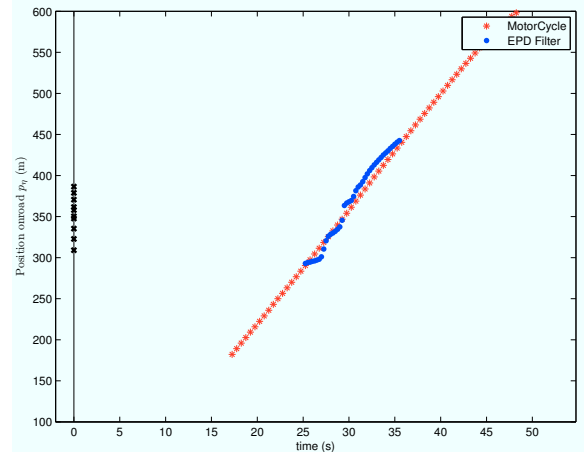


Figure 5: The position estimates of the EPD filter with the single target (motorcycle) data.

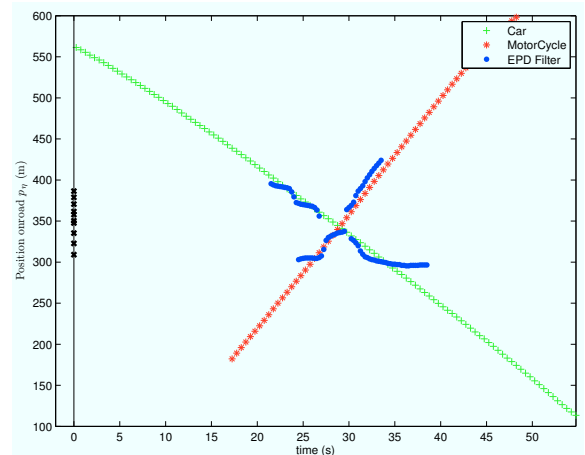


Figure 6: The position estimates of the EPD filter with the two target (car and motorcycle) data.

single target cases. During the crossing, since the target peaks in the estimated EPD form a single peak, only a single target is detected as expected but as soon as targets are separated, two targets are distinguished similar to the case before crossing.

5 Conclusions

A novel estimatee, EPD, inspired by the random set approaches in MTT, has been proposed and used successfully to track multiple targets from acoustic power measurements. For the EPD, approximate recursions was given which forms an EPD-filter. The GP representation for an EPD has been quite useful in obtaining practical Kalman filter type update formulas for the summary statistics. Possible use of such GP representations with PHDs and possibly proper densities must be investigated. Especially ill-conditioned point mass implementations can benefit from such an approach. As a future work, authors would like to investigate connections to the random set theory in more detail.

Acknowledgments

The authors gratefully acknowledge fundings from the Swedish Research Council VR in the Linnaeus Center CADICS, and Swedish Foundation for Strategic Research in the center MOVIII. The strategic motivation and practical relevance for this contribution stem from the Vinnova, SSF (Swedish Foundation for Strategic Research) and KKS *Institute Excellence Centre* for Advanced Sensors, Multi sensors and Sensor Networks (FOCUS). The authors would also like to thank FOI (Swedish Defense Research Agency) for the data used in the example. The first author would like to specifically thank Fredrik Gunnarsson of LiU and David Lindgren of FOI for helps with the database; and Henrik Ohlsson of LiU for his comments on the manuscript and fruitful discussions.

References

- [1] N. Patwari, A. Hero III, M. Perkins, N. Correal, and R. O’Dea, “Relative location estimation in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 51, no. 8, August 2003.
- [2] S. Gezici, Z. Tian, B. Giannakis, H. Kobayashi, and A. Molisch, “Localization via ultra-wideband radios,” *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.
- [3] F. Gustafsson and F. Gunnarsson, *Localization Algorithms and Strategies for Wireless Sensor networks*. IGI, 2009, ch. Measurements used in Wireless Sensor Networks Localization.
- [4] D. Li and Y. Hu, “Energy-based collaborative source localization using acoustic microsensor array,” *Journal of Applied Signal Processing*, pp. 321–337, 2003.
- [5] Y. Huang, J. Benesty, and B. Elko, “Passive acoustic source localization for video camera steering,” in *IEEE Conference on Acoustics, Speech and Signal Processing*, 2000.
- [6] F. Gustafsson and F. Gunnarsson, “Possibilities and fundamental limitations of positioning using wireless communication networks measurements,” *IEEE Signal Processing Magazine*, vol. 22, pp. 41–53, 2005.
- [7] Y. Boers, H. Driessen, and L. Schipper, “Particle filter based sensor selection in binary sensor networks,” in *Proceedings of the 11th International Conference on Information Fusion*, 2008.
- [8] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles, Techniques*. Storrs, CT: YBS Publishing, 1995.
- [9] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.
- [10] M. Fallon and S. Godsill, “Multi target acoustic source tracking with an unknown and time varying number of targets,” in *Proceeding of the Conference on Hands-Free Speech Communication and Microphone Arrays (HSCMA 2008)*, May 2008, pp. 77–80.
- [11] V. Cevher, R. Velmurugan, and J. McClellan, “Acoustic multitarget tracking using direction-of-arrival batches,” *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2810–2825, Jun. 2007.
- [12] W.-K. Ma, B.-N. Vo, S. Singh, and A. Baddeley, “Tracking an unknown time-varying number of speakers using TDOA measurements: A random finite set approach,” *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3291–3304, Sep. 2006.
- [13] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. Springer Verlag, 2001.
- [14] R. Mahler, *Statistical Multisource Multitarget Information Fusion*. Norwood, MA: Artech House, 2007.
- [15] D. Lindgren, H. Habberstad, M. Holmberg, and A. Lauberts, “Robust fusion of multiple microphone and geophone arrays in a ground sensor network,” Meeting Proceedings of Battlefield Acoustic Sensing for ISR Applications. RTO-MP-SET-107, 2006. [Online]. Available: <http://handle.dtic.mil/100.2/ADA478760>
- [16] M. Ekman, K. Davstad, and L. Sjoberg, “Ground target tracking using acoustic sensors,” in *Proceedings of Information, Decision and Control (IDC ’07)*, Feb. 2007, pp. 182–187.
- [17] X. Sheng and Y.-H. Hu, “Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 44 – 53, Jan. 2005.
- [18] M. Bugallo and P. Djuric, “Tracking of time-varying number of moving targets in wireless sensor fields by particle filtering,” in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, vol. 4, May 2006, pp. 865–868.
- [19] R. Mahler, “Multitarget Bayes filtering via first-order multitarget moments,” vol. 39, no. 4, Oct. 2003, pp. 1152–1178.
- [20] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: The MIT Press, 2006.
- [21] R. Mahler, “CPHD filters for superpositional sensors,” in *Signal and Data Processing of Small Targets 2009*, O. E. Drummond and R. D. Teichgraber, Eds., vol. 7445, no. 1. SPIE, 2009.
- [22] U. Orguner and F. Gustafsson, “Multi target tracking with acoustic power measurements using emitted power density,” Department of Electrical Engineering, Linköping University, Tech. Rep. LiTH-ISY-R-2947, Apr. 2010. [Online]. Available: www.control.isy.liu.se/research/reports/2010/2947.pdf