Shooter Localization in Wireless Microphone Networks[*]

David Lindgren[1][†]   **Olof Wilsson**   **Fredrik Gustafsson**[2]   **Hans Habberstad**[1]

[1]Swedish Defence Research Agency, Linköping, Sweden, firstname.lastname@foi.se
[2]Linköping University, Linköping, Sweden, fredrik@isy.liu.se

Abstract − *Shooter localization in a wireless network of microphones is studied. Both the acoustic muzzle blast (MB) from the gunfire and the ballistic shock wave (SW) from the bullet can be detected by the microphones and considered as measurements. The MB measurements give rise to a standard sensor network problem, similar to time difference of arrivals in cellular phone networks, and the localization accuracy is good, provided that the sensors are well synchronized compared to the MB detection accuracy. The detection times of the SW depend on both shooter position and aiming angle, and may provide additional information beside the shooter location, but again, this requires good synchronization. We analyze the approach to base the estimation on the time difference of MB and SW at each sensor, which becomes insensitive to synchronization inaccuracies. Cramér-Rao lower bound analysis indicates how a lower bound of the root mean square error depends on the synchronization error for the MB and the MB–SW difference, respectively. The estimation problem is formulated in a separable nonlinear least squares framework. Results from field trials with different types of ammunition show excellent accuracy using the MB–SW difference for both the position and the aiming angle of the shooter.*

Keywords: Sniper Localization, Sensor Network, Acoustics, Synchronization, Cramér-Rao Analysis

1   Introduction

Several acoustic shooter localization systems are today commercially available, see for instance [1–4]. Typically, one or more microphone arrays are used, each synchronously sampling acoustic phenomena associated with gunfire. An overview is found in [5]. Some of these systems are mobile, and in [6] it is even described how soldiers can carry the microphone arrays on their helmets. One interesting attempt to find direction of sound from one microphone only is described in [7]. It is based on direction dependent spatial filters (mimicking the human outer ear) and prior knowledge of the sound waveform, but this approach has not yet been applied to gun shots.

Indeed, less common are shooter localization systems based on singleton microphones geographically distributed in a wireless sensor network. An obvious issue in wireless networks is the sensor synchronization. For localization algorithms that rely on accurate timing like the ones based on *time difference of arrival* (TDOA), it is of major importance that synchronization errors are carefully controlled. Regardless if the synchronization is solved by using GPS or other techniques, see for instance [8–10], the synchronization procedures are associated with costs in battery life or communication resources that usually must be kept at a minimum.

In [11] the synchronization error impact on the sniper localization ability of an urban network is studied by using Monte Carlo simulations. One of the results is that the inaccuracy increased significantly (> 2 m) for synchronization errors exceeding approximately 4 ms. 56 small wireless sensor nodes were modeled. Another closely related work that deals with mobile asynchronous sensors is [12], where the estimation bounds with respect to both sensor synchronization and position errors are developed and validated by Monte-Carlo simulations. Also [13] should be mentioned, where combinations of directional and omnidirectional acoustic sensors for sniper localization are evaluated by perturbation analysis. In [14], estimation bounds for multiple acoustic arrays are developed and validated by Monte-Carlo simulations.

In this paper we derive fundamental estimation bounds for shooter localization systems based on wireless sensor networks, with the synchronization errors in focus. An accurate method independent of the synchro-

nization errors will be analyzed (the *MB–SW model*), as well as a useful bullet deceleration model. The algorithms are tested on data from a field trial with 10 microphones spread over an area of 100 m and with gunfire at distances up to 400 m. Partial results of this investigation appeared in [15] and almost simultaneously in [12].

The outline is as follows. Section 2 sketches the localization principle and describes the acoustical phenomena that are used. Section 3 gives the estimation framework. Section 4 derives the signal models for the muzzle blast (MB), shock wave (SW), combined MB;SW and difference MB–SW, respectively. Section 5 derives expressions for the root mean square error (RMSE) Cramér-Rao lower bound (CRLB) for the described models, and provides numerical results from a realistic scenario. Section 6 presents the results from field trials, and Section 7 gives the conclusions.

2 Localization Principle

Two acoustical phenomena associated with gunfire will be exploited to determine the shooter's position: the *muzzle blast* and the *shock wave*. The principle is to detect and time stamp the phenomena as they reach microphones distributed over an area, and let the shooter's position be estimated by, in a sense, the most likely point, considering the microphone locations and detection times.

The muzzle blast (MB) is the sound that probably most of us associate with a gun shot; the "bang". The MB is generated by the pressure depletion in effect of the bullet leaving the gun barrel. The sound of the MB travels at the speed of sound in all directions from the shooter. Provided that a sufficient number of microphones detect the MB, the shooters position can be more or less accurately determined.

The shock wave (SW) is formed by supersonic bullets. The SW has (approximately) the shape of an expanding cone, with the bullet trajectory as axis, and reaches only microphones that happens to be located inside the cone. The SW propagates at the speed of sound in direction away from the bullet trajectory, but since it is generated by a supersonic bullet, it always reaches the microphone before the MB, if it reaches the microphone at all. A number of SW detections may primarily reveal the direction to the shooter. Extra observations or assumptions on the ammunition are generally needed to deduce the distance to the shooter. The SW detection is also more difficult to utilize than the MB detection, since it depends on the bullet's speed and ballistic behavior.

Figure 1 shows an acoustic recording of gunfire. The first pulse is the SW, which for distant shooters significantly dominates the MB, not the least if the bullet passes close to the microphone. The figure shows real data, but a rather ideal case. Usually, and particularly
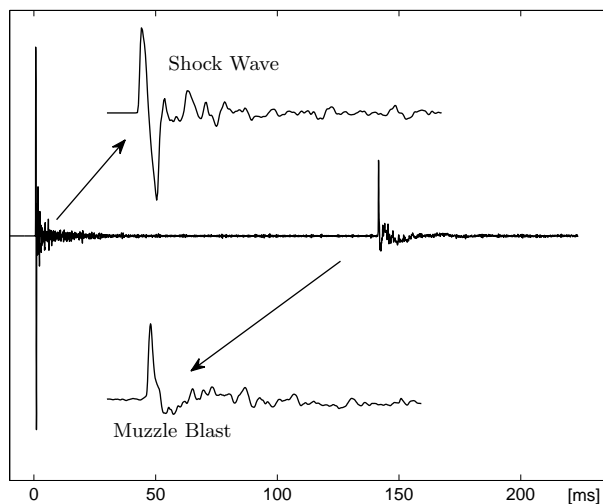

Figure 1: Signal from a microphone placed 180 m from a firing gun. Initial bullet speed is 767 m/s. The bullet passes the microphone at a distance of 30 m. The shockwave from the supersonic bullet reaches the microphone before the muzzle blast.

in urban environments, there are reflections and other acoustic effects that make it difficult to accurately determine the MB and SW times. This issue will however not be treated in this work. We will instead assume that the detection error is stochastic with a certain distribution. A more thorough analysis of the SW propagation is given in [16].

Of course, the MB and SW (when present) can be used in conjunction with each other. One of the ideas exploited later is to utilize the *time difference* between the MB and SW detections. This way, the localization is independent of the clock synchronization errors that are always present in wireless sensor networks.

3 Estimation Framework

It is assumed throughout this work that

1. the coordinates of the microphones are known with negligible error,

2. the arrival times of the MB and SW at each microphone are measured with significant synchronization error, and that

3. the shooter position and aim direction are the sought parameters.

Thus, assume there are M microphones with known positions $\{p_k\}_{k=1}^M$ in the network detecting the muzzle blast. Without loss of generality, the first $S \leq M$ ones also detect the shock wave. The detected times are denoted $\{y_k^{MB}\}_1^M$, and $\{y_k^{SW}\}_1^S$, respectively. Each detected time is subject to a detection error $\{e_k^{MB}\}_1^M$ and $\{e_k^{SW}\}_1^S$, different for all times, and a clock synchronization error $\{b_k\}_1^M$ specific for each microphone.

The firing time t_0, shooter position $x \in \mathbb{R}^3$ and shooting direction $\alpha \in \mathbb{R}^2$ are unknown parameters. Also the bullet speed v and speed of sound c are unknown. Basic signal models for the detected times as a function of the parameters will be derived in the next section. The notation is summarized in Table 1.

The derived signal models will be of the form

$$y = h(x, \theta; p) + e, \tag{1}$$

where y is a vector with the measured detection times, h is a nonlinear function with values in $\mathbb{R}^{M+S}$, and where θ represents the unknown parameters apart from x. The error e is assumed to be stochastic, see Section 4.5. Given the sensor locations in $p \in \mathbb{R}^{M \times 3}$, nonlinear optimization can be performed to estimate x, using the nonlinear least squares (NLS) criterion

$$\hat{x} = \operatorname{argmin}_x \min_\theta V(x, \theta; p), \tag{2a}$$

$$V(x, \theta; p) = ||y - h(x, \theta; p)||_R^2. \tag{2b}$$

Here, argmin denotes the minimizing argument, min the minimum of the function, and $||v||_Q^2$ denotes the Q-norm, that is, $||v||_Q^2 \triangleq v^T Q^{-1} v$. Whenever Q is omitted, $Q = I$ is assumed. The loss function norm R is chosen by consideration of the expected error characteristics. Numerical optimization, for instance the Gauss-Newton method, can here be applied to get the NLS estimate.

In the next section it will become clear that the assumed unknown firing time and the inverse speed of sound enter the model equations linearly. To exploit this fact we identify a sub-linear structure in the signal model and apply the weighted least squares method to the parameters appearing linearly, *the separable least squares method*, see for instance [17]. By doing so, the NLS search space is reduced which in turn significantly reduces the computational burden. For that reason, the signal model (1) is rewritten as

$$y = h_N(x, \theta_N; p) + h_L(x, \theta_N; p)\theta_L + e. \tag{3}$$

Note that θ_L enters linearly here. The NLS problem can then be formulated as

$$\hat{x} = \operatorname{argmin}_x \min_{\theta_L, \theta_N} V(x, \theta_N, \theta_L; p),$$

$$V(x, \theta_N, \theta_L; p) = ||y - h_N(x, \theta_N; p) - h_L(x, \theta_N; p)\theta_L||_R^2.$$

Since θ_L enters linearly it can be solved for by linear least squares (the arguments of $h_L(x, \theta_N; p)$ and $h_N(x, \theta_N; p)$ are suppressed for clarity):

$$\hat{\theta}_L = \operatorname{argmin}_{\theta_L} V(x, \theta_N, \theta_L; p)$$
$$= (h_L^T R^{-1} h_L)^{-1} h_L^T R^{-1} (y - h_N), \tag{4a}$$

$$P_L = (h_L^T R^{-1} h_L)^{-1}. \tag{4b}$$

Here, $\hat{\theta}_L$ is the weighted least squares estimate and P_L is the covariance matrix of the estimation error. This simplifies the nonlinear minimization to

$$\hat{x} = \operatorname{argmin}_x \min_{\theta_N} V(x, \theta_N, \hat{\theta}_L; p) \tag{5a}$$

$$= \operatorname{argmin}_x \min_{\theta_N}$$

$$||y - h_N + h_L(h_L^T R^{-1} h_L)^{-1} h_L^T R^{-1}(y - h_N)||_{R'}^2,$$
$$R' = R + h_L P_L h_L^T. \tag{5b}$$

This general separable least squares (SLS) approach will now be applied to four different combinations of signal models for the MB and SW detection times.

4 Signal Models
4.1 Muzzle Blast Model (MB)

According to the clock at microphone k, the *muzzle blast (MB) sound* is assumed to reach p_k at the time

$$y_k = t_0 + b_k + \frac{1}{c}||p_k - x|| + e_k. \tag{6}$$

The shooter position x and microphone location p_k are in $\mathbb{R}^n$, where generally $n = 3$. However, both computational and numerical issues occasionally motivate a simplified plane model with $n = 2$. For all M microphones, the model is represented in vector form as

$$y = b + h_L(x; p)\theta_L + e, \tag{7}$$

where

$$\theta_L = \begin{bmatrix} t_0 & 1/c \end{bmatrix}^T, \tag{8a}$$

$$h_{L,k}(x; p) = \begin{bmatrix} 1 & ||p_k - x|| \end{bmatrix}^T, \tag{8b}$$

and where y, b, and e are vectors with elements y_k, b_k, and e_k, respectively. 1_M is the vector with M ones, where M might be omitted if there is no ambiguity regarding the dimension. Furthermore, p is M-by-n, where each row is a microphone position. Note that the *inverse* of the speed of sound enters linearly. The $\cdot_L$ notation indicates that $\cdot$ is part of a linear relation, as described in the previous section. With $h_N = 0$ and $h_L = h_L(x; p)$, (5) gives

$$\hat{x} = \operatorname{argmin}_x ||y - h_L(h_L^T R^{-1} h_L)^{-1} h_L^T R^{-1} y||_{R'}^2, \tag{9a}$$

$$R' = R + h_L(h_L^T R^{-1} h_L)^{-1} h_L^T. \tag{9b}$$

Here, h_L depends on x as given in (8b).

This criterion has computationally efficient implementations, that in many applications make the time it takes to do an exhaustive minimization over a, say, 10-meter grid acceptable. The grid-based minimization of course reduces the risk to settle on suboptimal local minimizers, which otherwise could be a risk using greedy search methods. The objective function does, however, behave rather well. Figure 2 visualizes (9a) in logarithmic scale for data from a field trial (the norm is $R' = I$). Apparently, there are only two local minima.

Table 1: Notation. MB, SW, and MB–SW are different models, and L/N indicates if model parameters or signals enter the model linearly (L) or nonlinearly (N).

| Variable	MB	SW	MB–SW	Description
M				Number of microphones
S				Number of microphones receiving shock wave, $S \leq M$
x	N	N	N	Position of shooter, $\mathbb{R}^n$ ($n = 2, 3$)
p_k	N	N	N	Position of microphone k, $\mathbb{R}^n$ ($n = 2, 3$)
y_k	L	L	L	Measured detection time for microphone at position p_k
t_0	L	L		Rifle or gun firing time
c	L	N	N	Speed of sound
v		N	N	Speed of bullet
α		N	N	Shooting direction, $\mathbb{R}^{n-1}$ ($n = 2, 3$)
b_k	L	L		Synchronization error for microphone k
e_k	L	L	L	Detection error at microphone k
r		N	N	Bullet speed decay rate
d_k				Point of origin for shock wave received by microphone k
β				Mach angle, $\sin \beta = c/v$
γ				Angle between line of sight to shooter and shooting angle
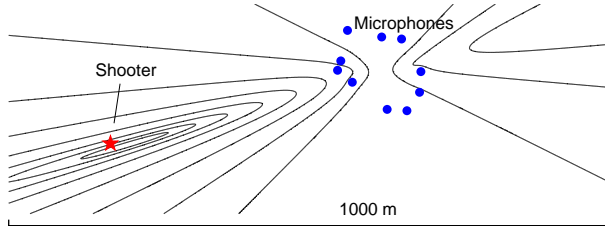
Figure 2: Level curves of the muzzle blast localization criterion based on data from a field trial.

4.2 Shock Wave Model (SW)

In general, the bullet follows a ballistic three-dimensional trajectory. In practice, a simpler model with a two-dimensional trajectory with constant deceleration might suffice. Thus, it will be assumed that the bullet follows a straight line with initial speed v_0, see Figure 3. Due to air friction, the bullet decelerates, so when the bullet has traveled the distance $||d_k - x||$, for some point d_k on the trajectory, the speed is reduced to

$$v = v_0 - r||d_k - x||, \qquad (10)$$

where r is an assumed known ballistic parameter. This is a rather coarse bullet trajectory model, compared with, for instance, the curvilinear trajectories proposed by [18], but we use it here for simplicity. This model is also a special case of the ballistic model used in [19].

The shock wave from the bullet trajectory propagates at the speed of sound c with angle β_k to the bullet heading. β_k is the *Mach angle* defined as

$$\sin \beta_k = \frac{c}{v} = \frac{c}{v_0 - r||d_k - x||}. \qquad (11)$$

d_k is now the point where the shock wave that reaches

microphone k is generated. The time it takes the bullet to reach d_k is

$$\int_0^{||x - d_k||} \frac{d\xi}{v_0 - r \cdot \xi} = \frac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||}. \qquad (12)$$

This time and the wave propagation time from d_k to p_k sum up to the total time from firing to detection,

$$y_k = t_0 + b_k + \frac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||} + \frac{1}{c}||d_k - p_k|| + e_k,$$

according to the clock at microphone k. Note that the variable names y and e for notational simplicity have been reused from the MB model. Below, also h, θ_N, and θ_L will be reused. When there is ambiguity, a superscript will indicate exactly which entity that is referred to, for instance, y^{MB}, h^{SW}.

It is a little bit tedious to calculate d_k. The law of sines gives

$$\frac{\sin(90° - \beta_k - \gamma_k)}{||d_k - x||} = \frac{\sin(90° + \beta_k)}{||p_k - x||}, \qquad (13)$$

which together with (11) implicitly defines d_k. We have not found any simple closed form for d_k, so we solve for d_k numerically, and in case of multiple solutions keep the admissible one (which turns out to be unique). γ_k is trivially induced by the shooting direction α (and x, p_k). Both these angles thus depend on x implicitly.

The vector form of the model is

$$y = b + h_N(x, \theta_N; p) + h_L(x, \theta_N; p)\theta_L + e, \qquad (14)$$

where

$$h_L(x, \theta_N; p) = 1,$$
$$\theta_L = t_0,$$
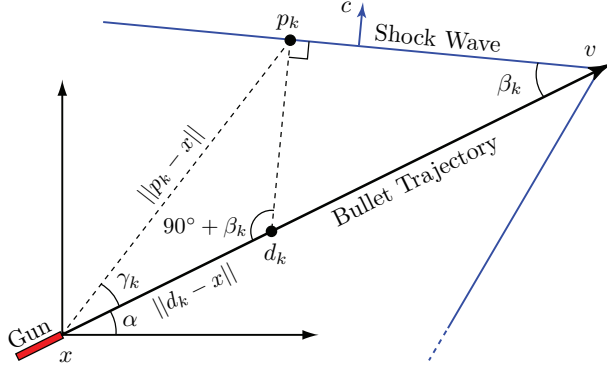$$\theta_N = \begin{bmatrix} 1/c & \alpha^T & v_0 \end{bmatrix}^T,$$

Figure 3: Geometry of supersonic bullet trajectory and shock wave. Given the shooter location x, the shooting direction (aim) α, the bullet speed v, and the speed of sound c, the time it takes from firing the gun to detecting the shock wave can be calculated.

and where row k of $h_N(x, \theta_N; p) \in \mathbb{R}^{S \times 1}$ is

$$h_{N,k}(x, \theta_N; p_k) = \tfrac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||} + \tfrac{1}{c}||d_k - p_k||,$$

and d_k is the admissible solution to (11) and (13).

4.3   Combined Model (MB;SW)

In the MB and SW models, the synchronization error has to be regarded as a noise component. In a combined model, each pair of MB and SW detections depend on the same synchronization error, and consequently the synchronization error can be regarded as a parameter (at least for all sensor nodes inside the SW cone). The total signal model could be fused from the MB and SW models as the total observation vector

$$y^{MB;SW} = h_N^{MB;SW}(x, \theta_N; p) + h_L^{MB;SW}(x, \theta_N; p)\theta_L \tag{15}$$

where

$$y^{MB;SW} = \begin{bmatrix} y^{MB} \\ y^{SW} \end{bmatrix} \tag{16}$$

$$\theta_L = \begin{bmatrix} t_0, b_M^T \end{bmatrix}^T \tag{17}$$

$$h_L^{MB;SW}(x, \theta_N; p) = \begin{bmatrix} 1_{M,1} & I_M \\ 1_{S,1} & \begin{bmatrix} I_S & 0_{S,M-S} \end{bmatrix} \end{bmatrix}, \tag{18}$$

$$\theta_N = \begin{bmatrix} \alpha^T & v_0 & c \end{bmatrix}^T \tag{19}$$

$$h_{N,k}^{MB;SW} = \begin{bmatrix} \tfrac{1}{c}||p_k - x||^T & h_{N,k}^{SW}(x, \theta_N; p) \end{bmatrix}. \tag{20}$$

Here, b_M denotes the set of clock synchronization errors for the M MB observations, and b_S denotes the S clock synchronization errors for the SW observations. Usually, b_S is a subset of b_M.

4.4   Difference Model (MB–SW)

Motivated by accurate localization despite synchronization errors, we study the *MB–SW model*:

$$
\begin{aligned}
y_k^{MB-SW} &= y_k^{\text{MB}} - y_k^{\text{SW}} \\
&= h_L^{\text{MB}}(x; p)\theta_L^{\text{MB}} - h_N^{\text{SW}}(x, \theta_N^{\text{SW}}; p) \\
&\quad - h_L^{\text{SW}}(x, \theta_N; p)\theta_N^{\text{SW}} + e_k^{\text{MB}} - e_k^{\text{SW}}, \quad (21)
\end{aligned}
$$

for $k = 1, 2 \ldots S$. This rather special model has also been analyzed in [12, 15]. The key idea is that y is by cancellation independent of both the firing time t_0 and the synchronization error b. The drawback, of course, is that there are only S equations (instead of a total of $M + S$) and the detection error increases, $e_k^{\text{MB}} - e_k^{\text{SW}}$. However, when the synchronization errors are expected to be significantly larger than the detection errors, and when also S is sufficiently large (at least as large as the number of parameters), this model is believed to give better localization accuracy. This will be investigated later.

There are no parameters in (21) that appear linearly everywhere. Thus, the vector form for the MB–SW model can be written

$$y^{MB-SW} = h_N^{MB-SW}(x, \theta_N; p) + e, \tag{22}$$

where

$$
\begin{aligned}
h_{N,k}^{MB-SW}(x, \theta_N; p_k) &= \\
&= \tfrac{1}{c}||p_k - x|| - \tfrac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||} - \tfrac{1}{c}||d_k - p_k||,
\end{aligned}
$$

and $y = y^{\text{MB}} - y^{\text{SW}}$, $e = e^{\text{MB}} - e^{\text{SW}}$. As before, d_k is the admissible solution to (11), (13). The MB–SW least squares criterion is

$$\hat{x} = \arg\min_{x, \theta_N} ||y^{MB-SW} - h_N^{MB-SW}(x, \theta_N; p)||_R^2, \tag{23}$$

which requires numerical optimization. Numerical experiments indicate that this optimization problem is more prone to local minima, compared to (9a) for the MB model, therefore good starting points for the numerical search are essential. One such starting point could for instance be the MB estimate $\hat{x}^{\text{MB}}$. Initial shooting direction could be given by assuming, in a sense, the worst possible case; that the shooter aims at some point close to the center of the microphone network.

4.5   Error Model

At an arbitrary moment, the detection errors and synchronization errors are assumed to be independent stochastic variables with normal distribution,

$$e^{\text{MB}} \sim \mathcal{N}(0, R^{\text{MB}}), \tag{24a}$$

$$e^{\text{SW}} \sim \mathcal{N}(0, R^{\text{SW}}), \tag{24b}$$

$$b \sim \mathcal{N}(0, R^b). \tag{24c}$$

For the MB–SW model the error is consequently

$$e^{\text{MB–SW}} \sim \mathcal{N}(0, R^{\text{MB}} + R^{\text{SW}}). \qquad (24\text{d})$$

Assuming that $S = M$ in the MB;SW model, the detection error covariance can be expressed in a simple manner as

$$R^{MB;SW} = \begin{bmatrix} R^{\text{MB}} + R^b & R^b \\ R^b & R^{\text{SW}} + R^b \end{bmatrix}. \qquad (24\text{e})$$

Note that the correlation structure of the clock synchronization error b enables estimation of these. Note also that the (assumed known) total error covariance, generally denoted R, dictates the norm used in the weighted least squares criterion. R also impacts the estimation bounds. This will be discussed in the next section.

4.6 Summary of Models

Four models with different purposes have been described in this section:

- **MB**. Given that the acoustic environment enables reliable detection of the muzzle blast, the MB model promises the most robust estimation algorithms. It also allows global minimization with low-dimensional exhaustive search algorithms. This model is thus suitable for initialization of algorithms based on the subsequent models.

- **SW**. The SW model extends the MB model with shooting angle, bullet speed, and deceleration parameters, which provide useful information for sniper detection applications. The SW is easier to detect in disturbed environments, particularly when the shooter is far away and the bullet passes closely. However, a sufficient number of microphones are required to be located within the SW cone, and the SW measurements alone can not be used to determine the distance to the shooter.

- **MB;SW**. The total MB;SW model keeps all information from the observations, and should thus provide the most accurate and general estimation performance. However, the complexity of the estimation problem is large.

- **MB–SW**. All algorithms based on the models above require that the synchronization error in each microphone is either negligible or can be described with a statistical distribution. The MB–SW model relaxes such assumptions by eliminating the synchronization error by taking differences of the two pulses at each microphone. This also eliminates the shooting time. The final model contains all interesting parameters for the problem, but only one nuisance parameter (actual speed of sound, which further may be eliminated if known sufficiently well).

The different parameter vectors in the relation $y = h_L(\theta_N)\theta_L + h_N(\theta_N) + e$ are summarized in Table 2.

5 Cramér-Rao Lower Bound

The accuracy of any unbiased estimator $\hat{\eta}$ in the rather general model

$$y = h(\eta) + e \qquad (25)$$

is, under not too restrictive assumptions [20], bounded by the Cramér-Rao bound,

$$\text{Cov}(\hat{\eta}) \geq \mathcal{I}^{-1}(\eta^o), \qquad (26)$$

where $\mathcal{I}(\eta^o)$ is the Fisher's information matrix evaluated at the correct parameter values η^o. Here, the location x is for notational purposes part of the parameter vector η. Also the sensor positions p_k can be part of η, if these are known only with a certain uncertainty. The Cramér-Rao lower bound provides a fundamental estimation limit for unbiased estimators, see [20]. This bound has been analyzed thoroughly in the literature, primarily for AOA, TOA and TDOA, [21–23].

The Fisher information matrix for $e \sim \mathcal{N}(0, R)$ takes the form

$$\mathcal{I}(\eta) = \nabla_\eta [h(\eta)] R^{-1} \nabla_\eta^T [h(\eta)]. \qquad (27)$$

The bound is evaluated for a specific location, parameter setting, and microphone positioning, collectively $\eta = \eta^o$.

The bound for the localization error is

$$\text{Cov}(\hat{x}) \geq \begin{bmatrix} I_n & 0 \end{bmatrix} \mathcal{I}^{-1}(\eta^o) \begin{bmatrix} I_n \\ 0 \end{bmatrix}. \qquad (28)$$

This covariance can be converted to a more convenient scalar value giving a bound on the root mean square error (RMSE) using the trace operator:

$$\text{RMSE} \geq \sqrt{\frac{1}{n} \text{tr} \left(\begin{bmatrix} I_n & 0 \end{bmatrix} \mathcal{I}^{-1}(\eta^o) \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right)}. \qquad (29)$$

The RMSE bound can be used to compare the information in different models in a simple and unambiguous way, which does not depend on which optimization criterion is used or which numerical algorithm that is applied to minimize the criterion.

5.1 MB Case

For the MB case, the entities in (27) are identified by

$$\eta = \begin{bmatrix} x^T & \theta_L^T \end{bmatrix}^T, \qquad (30)$$
$$h(\eta) = h_L^{\text{MB}}(x; p)\theta_L, \qquad (31)$$
$$R = R^{\text{MB}} + R^b. \qquad (32)$$

Note that b is accounted for by the error model. The Jacobian $\nabla_\eta h$ is an M-by-$n+2$ matrix, n being the dimension of x. The LS solution in (4a) however gives a shortcut to an M-by-n Jacobian,

$$\nabla_x [h_L \hat{\theta}_L] = \nabla_x [h_L (h_L^T R^{-1} h_L)^{-1} h_L^T R^{-1} y^o] \qquad (33)$$

Model	Linear Parameters	Nonlinear Parameters	$\dim(\theta)$	$\dim(y)$
MB	$\theta_L^{MB} = \begin{bmatrix} t_0 & 1/c \end{bmatrix}^T$	$\theta_N^{MB} = [\ ]$	$2+0$	M
SW	$\theta_L^{SW} = t_0$	$\theta_N^{MB} = \begin{bmatrix} 1/c, \alpha^T, v_0 \end{bmatrix}^T$	$1+(n+1)$	S
MB;SW	$\theta_L^{MB;SW} = \begin{bmatrix} t_0 & b \end{bmatrix}^T$	$\theta_N^{MB;SW} = \begin{bmatrix} 1/c, \alpha^T, v_0 \end{bmatrix}^T$	$(M+1)+(n+1)$	$M+S$
MB–SW	$\theta_L^{MB-SW} = [\ ]$	$\theta_N^{MB-SW} = \begin{bmatrix} 1/c, \alpha^T, v_0 \end{bmatrix}^T$	$0+(n+1)$	S

Table 2: Summary of parameter vectors for the different models $y = h_L(\theta_N)\theta_L + h_N(\theta_N) + e$, where the noise models are summarized in (24). The values of the dimensions assume that the set of microphones giving SW observations is a subset of the MB observations.
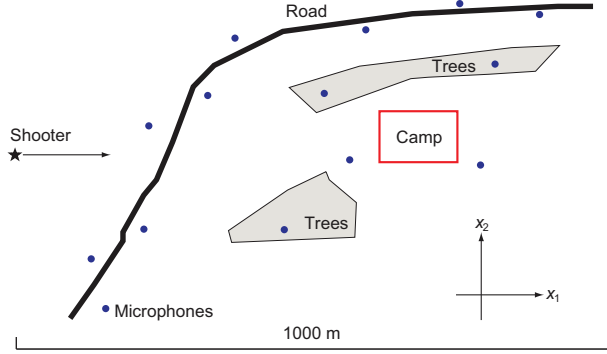

Figure 4: Example scenario. A network with 14 sensors deployed for camp protection. The sensors detect intruders, keep track on vehicle movements, and, of course, locate shooters.

for $y^o = h_L(x^o; p^o)\theta_L^o$, where x^o, p^o and θ^o denote the true (unperturbed) values. For the case $n = 2$ and known $p = p^o$, this Jacobian can, with some effort, be expressed explicitly. The equivalent bound is

$$\text{Cov}(\hat{x}) \geq \left[\nabla_x^T [h_L \hat{\theta}_L] R^{-1} \nabla_x [h_L \hat{\theta}_L] \right]^{-1}. \tag{34}$$

5.2 SW, MB;SW and MB–SW Cases

The estimation bounds for the SW, MB;SW and MB–SW cases are analogously to (28), but there are hardly any analytical expressions available. The Jacobian is probably best evaluated by finite difference methods.

5.3 Numerical Example

The really interesting question is how the information in the different models relate to each other. We will study a scenario where 14 microphones are deployed in a sensor network to support camp protection, see Figure 4. The microphones are positioned along a road to track vehicles and around the camp site to detect intruders. Of course, the microphones also detect muzzle blasts and shock waves from gunfire, so shooters can be localized and the shooter's target identified.

A plane model (flat camp site) is assumed, $x \in \mathbb{R}^2$,

$\alpha \in \mathbb{R}$. Furthermore, it is assumed that

$$R^b = \sigma_b^2 I, \quad \text{(synchronization error cov.)}$$
$$R^{\text{MB}} = R^{\text{SW}} = \sigma_e^2 I, \quad \text{(detection error cov.)}$$

and that $\alpha = 0$, $c = 330$ m/s, $v_0 = 700$ m/s, $r = 0.63$. The scenario setup implies that all microphones detect the shock wave, so $S = M = 14$. All bounds presented below are calculated by numerical finite difference methods.

MB model The localization accuracy using the MB model is bounded below according to

$$\text{Cov}(\hat{x}^{\text{MB}}) \geq (\sigma_e^2 + \sigma_b^2) \begin{bmatrix} 64 & -17 \\ -17 & 9 \end{bmatrix} \cdot 10^4. \tag{35}$$

The root mean square error (RMSE) is consequently bounded according to

$$\text{RMSE}\left(\hat{x}^{\text{MB}}\right) \geq \sqrt{\tfrac{1}{n} \text{tr}\, \text{Cov}\, \hat{x}^{\text{MB}}} \approx 606\sqrt{\sigma_e^2 + \sigma_b^2} \quad \text{[m]}. \tag{36}$$

Monte Carlo simulations (not described here) indicate that the NLS estimator attains this lower bound for $\sqrt{\sigma_e^2 + \sigma_b^2} < 0.1$ s. The dash-dotted curve in Figure 5 shows the bound versus σ_b for fix $\sigma_e = 500$ μs. An uncontrolled increase as soon as $\sigma_b > \sigma_e$ can be noted.

SW Model The SW model is disregarded here, since the SW detections alone contain no shooter distance information.

MB–SW Model The localization accuracy using the MB–SW model is bounded according to

$$\text{Cov}(\hat{x}^{\text{MB–SW}}) \geq \sigma_e^2 \begin{bmatrix} 28 & 5 \\ 5 & 12 \end{bmatrix} \cdot 10^5, \tag{37}$$

$$\text{RMSE}\left(\hat{x}^{\text{MB–SW}}\right) \geq 1430\sigma_e \quad \text{[m]}. \tag{38}$$

The dashed lines in Figure 5 correspond to the RMSE bound for four different values of σ_e. Here, the MB–SW model gives at least twice the error of the MB model, provided that there are no synchronization errors. However, in a wireless network we expect the synchronization error to be 10-100 times larger than the detection error, and then the MB–SW error will be substantially smaller than the MB error.
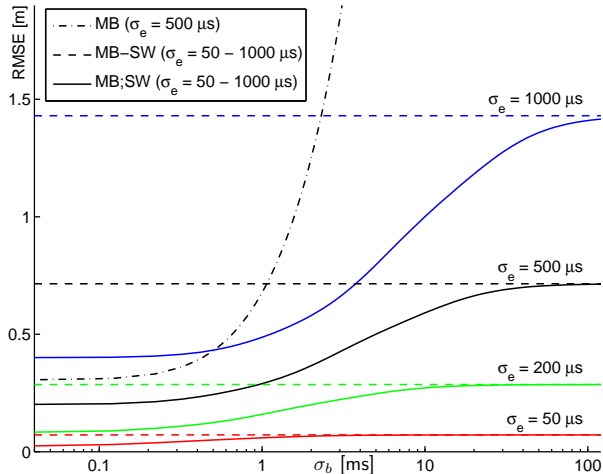
Figure 5: Cramér-Rao RMSE bound (29) for the MB (36), the MB–SW (38), and the MB;SW models, respectively, as a function of the synchronization error (STD) σ_b, and for different levels of detection error σ_e.

MB;SW Model The expression for the MB;SW bound is somewhat involved, so the dependence on σ_b is only presented graphically, see Figure 5. The solid curves correspond to the MB;SW RMSE bound for the same four values of σ_e as for the MB–SW bound.

Apparently, when the synchronization error σ_b is large compared to the detection error σ_e, the MB–SW and MB;SW models contain roughly the same amount of information, and the model having the simplest estimator, that is, the MB–SW model, should be preferred. However, when the synchronization error is smaller than 100 times the detection error, the complete MB;SW model becomes more informative.

These results are comparable with the analysis in [12] (Figure 4a), where an example scenario with 6 microphones is considered.

5.4 Summary of the CRLB Analysis

The synchronization error level in a wireless sensor network is usually a matter of design trade off between performance and battery costs required by synchronization mechanisms. Based on the scenario example, the CRLB analysis is summarized with the following recommendations:

- If $\sigma_b \gg \sigma_e$, then the MB–SW model should be used.

- If σ_b is moderate, then the MB;SW model should be used.

- Only if σ_b is very small ($\sigma_b \leq \sigma_e$), and the shooting direction is of minor interest, and performance may be traded for simplicity, then the MB model should be used.
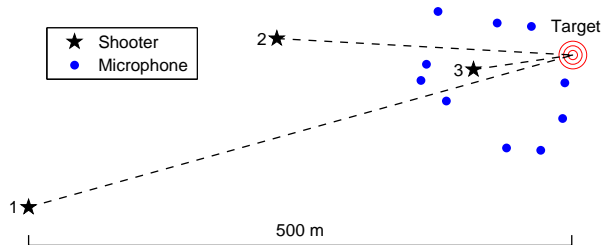

Figure 6: Scene of the shooter localization field trial. There are ten microphones, three shooter positions, and a common target.

6 Experimental Data

A field trial to collect acoustic data on non-military small arms fire is conducted. 10 microphones are placed around a fictitious camp, see Figure 6. The microphones are placed close to the ground and wired to a common recorder with 16-bit sampling at 48 kHz. A total of 42 rounds are fired from three positions and aimed at a common cardboard target. Three rifles and one pistol are used, see Table 3. Four rounds are fired of each armament at each shooter position, with two exceptions. The pistol is only used at position three. At position three, six instead of four rounds of 308W are fired. All ammunition types are supersonic. However, when firing from position three, not all microphones are subjected to the shock wave.

Light wind, no clouds, around 24 °C are the weather conditions. Little or no acoustic disturbances are present. The terrain is rough. Dense woods surround the test site. There is light bush vegetation within the site. Shooter position 1 is elevated some 20 m, otherwise spots are within ±5 m of a horizontal plane. Ground truth values of the positions are determined with less relative error than 1 m, except for shooter position 1, which is determined with 10 m accuracy.

6.1 Detection

The MB and SW are detected by visual inspection of the microphone signals in conjunction with filtering techniques. For shooter position 1 and 2, the shock wave detection accuracy is approximately $\sigma_e^{SW} \approx 80\,\mu s$, and the muzzle blast error σ_e^{MB} is slightly worse. For shooting position 3 the accuracies are generally much worse, since the muzzle blast and shock wave components become intermixed in time.

6.2 Numerical Setup

For simplicity, a plane model is assumed. All elevation measurements are ignored and $x \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$. Localization using the MB model (6) is done by minimizing (9a) over a 10 m grid well covering the area of

Table 3: Armament and ammunition used at the trial, and number of rounds fired at each shooter position. Also, the resulting localization RMSE for the MB–SW model for each shooter position. For the Luger Pistol the MB model RMSE is given, since only one microphone is located in the Luger Pistol SW cone.

Type	Caliber	Weight	Velocity	Sh. pos.	# Rounds	RMSE
308 Winchester	7.62 mm	9.55 g	847 m/s	1, 2, 3	4, 4, 6	19, 6, 6 m
Hunting Rifle	9.3 mm	15 g	767 m/s	1, 2, 3	4, 4, 4	6, 5, 6 m
Swedish Mauser	6.5 mm	8.42 g	852 m/s	1, 2, 3	4, 4, 4	40, 6, 6 m
Luger Pistol	9 mm	6.8 g	400 m/s	3	−, −, 4	−, −, 2 m

interest, followed by numerical minimization.

Localization using the MB–SW model (21) is done by numerically minimizing (23). The objective function is subject to local optima, therefore the more robust muzzle blast localization $\hat{x}$ is used as an initial guess. Furthermore, the direction from $\hat{x}$ toward the mean point of the microphones ("the camp") is used as initial shooting direction α. Initial bullet speed is $v = 800$ m/s and initial speed of sound is $c = 330$ m/s. $r = 0.63$ is used, which is a value derived from the 308 Winchester ammunition ballistics.

6.3   Results

Figure 7 shows, at three enlarged parts of the scene, the resulting position estimates based on the MB model (blue crosses) and based on the MB–SW (squares). Apparently, the use of the shock wave significantly improves localization at position 1 and 2, while rather the opposite holds at position 3. Figure 8 visualizes the shooting direction estimates, $\hat{\alpha}$. Estimate root mean square errors (RMSE) for the three shooter positions, together with the theoretical bounds (29) are given in Table 4. The practical results indicate that the use of the shock wave from distant shooters cut the error by at least 75%.

Synchronization and Detection Errors   Since all microphones are recorded by a common recorder, there are actually *no* timing errors due to inaccurate clocks. This is of course the best way to conduct a controlled experiment, where any uncertainty renders the dataset less useful. From experimental point of view, it is then simple to add synchronization errors of any desired magnitude off-line. On the dataset at hand, this is however work under progress. At the moment, there are apparently other sources of error, worth identifying. It should however be clarified, that in the final wireless sensor product, there will always be an unpredictable clock error. As mentioned, detection errors are present, and the expected level of these (80 μs) are used for bound calculations in Table 4. It is noted that the bounds are in level with, or below, the positioning errors.

There are at least two explanations for the bad performance using the MB–SW model at shooter posi-

tion 3. One is, that the number of microphones reached by the shock wave is insufficient to make accurate estimates. There are four unknown model parameters, but for the relatively low speed of pistol ammunition, for instance, only one microphone has a valid shock wave detection. Another explanation is, that the increased detection uncertainty (due to SW/MB intermix) impacts the MB–SW model harder, since it relies on accurate detection of both the MB and SW.

Model Errors   No doubt, there are model inaccuracies both in the ballistic and in the acoustic domain. To that end, there are meteorological uncertainties out of our control. For instance, looking at the MB–SW localizations around shooter position 1 in Figure 7 (squares), three clusters are identified that correspond to three ammunition types with different ballistic properties, see the RMSE for each ammunition and position in Table 3. This clustering or bias more likely stems from model errors than from detection errors, and could at least partially explain the large gap between theoretical bound and RMSE in Table 4. Working with three-dimensional data in the plane is of course another model discrepancy, that could have greater impact than we first anticipated. This will be investigated in experiments to come.

Numerical Uncertainties   Finally, we face numerical uncertainties. There is no guarantee that the numerical minimization programs we have used here for the MB–SW model really deliver the global minimum. In a realistic implementation, every possible a priori knowledge and also qualitative analysis of the SW and MB signals (amplitude, duration, caliber classification and so on) together with basic consistency checks are used to reduce the search space. The reduced search space may then be exhaustively sampled over a grid prior to the final numerical minimization. Simple experiments on an ordinary desktop PC indicate that with an efficient implementation, it is feasible to within the time frame of one second minimize any of the described model objective functions over a discrete grid with 10^7 points. Thus, by allowing–say–one second extra of computation time, the risk for hitting a local optima could be significantly reduced.
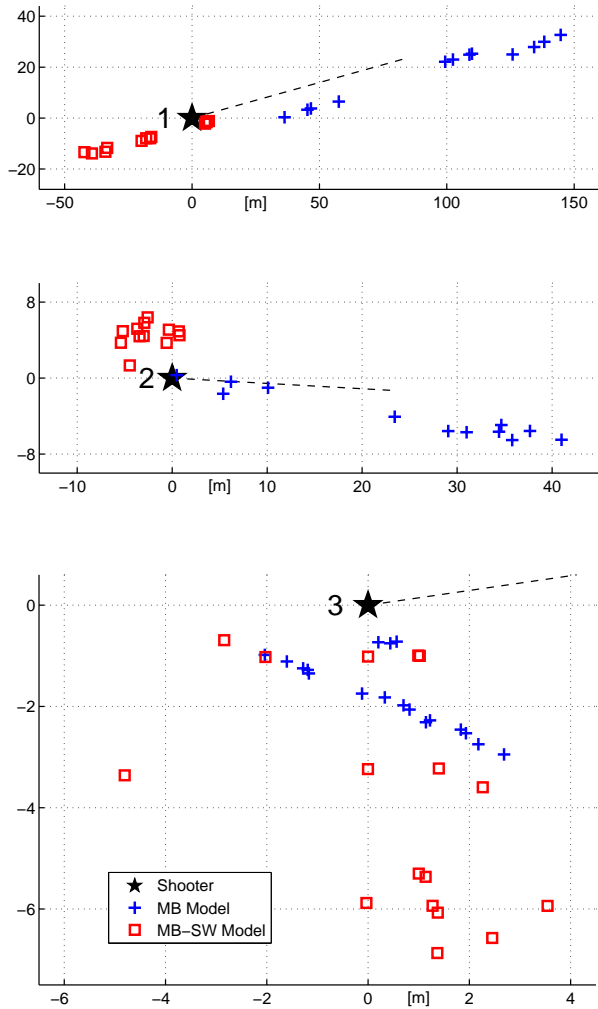
Figure 7: Estimated positions $\hat{x}$ based on the MB model and on the MB–SW model. The diagrams are enlargements of the interesting areas around the shooter positions. The dashed lines identify the shooting directions.

7 Conclusions

We have presented a framework for estimation of shooter location and aiming angle from wireless networks where each node has a single microphone. Both the acoustic muzzle blast (MB) and the ballistic shock wave (SW) contain useful information about the position, but only the SW contains information about the aiming angle. A separable nonlinear least squares (SNLS) framework was proposed to limit the parametric search space and to enable the use of global grid-based optimization algorithms (for the MB model), eliminating potential problems with local minima.

For a perfectly synchronized network, both MB and SW measurements should be stacked into one large signal model for which SNLS is applied. However, when the synchronization error in the network becomes comparable to the detection error for MB and SW, the performance quickly deteriorates. For that reason, the
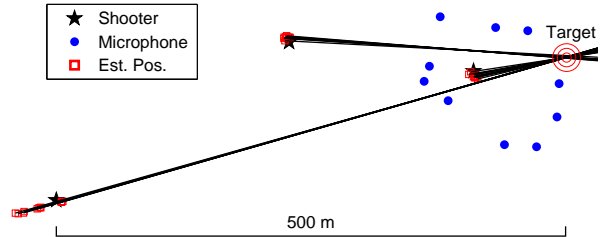

Figure 8: Estimated shooting directions. The relatively slow pistol ammunition is excluded.

Table 4: Localization RMSE and theoretical bound (29) for the three different shooter positions using the MB and the MB–SW models, respectively, beside the aim RMSE for the MB–SW model. The aim RMSE is with respect to the aim at $\hat{x}$ against the target, α', not with respect to the true direction α. This way the ability to identify the target is assessed.

Shooter position	1	2	3
RMSE($\hat{x}^{\text{MB}}$)	105 m	28 m	2.4 m
MB Bound	1 m	0.4 m	0.02 m
RMSE ($\hat{x}^{\text{MB–SW}}$)	26 m	5.7 m	5.2 m
MB–SW Bound	9 m	0.1 m	0.08 m
RMSE($\hat{\alpha}'$)	0.041°	0.14°	17°

time difference of MB and SW at each microphone is used, which automatically eliminates any clock offset. The effective number of measurements decreases in this approach, but as the CRLB analysis showed, the root mean square position error is comparable to that of the ideal stacked model, at the same time as the synchronization error distribution may be completely disregarded.

The bullet speed occurs as nuisance parameters in the proposed signal model. Further, the bullet retardation constant was optimized manually. Future work will investigate if the retardation constant should also be estimated, and if these two parameters can be used, together with the MB and SW signal forms, to identify the weapon and ammunition.

References

[1] J. Bédard and S. Paré, "Ferret, a small arms' fire detection system: Localization concepts," in *Proc. of the SPIE Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Defense and Law Enforcement II*, vol. 5071, 2003.

[2] J. A. Mazurek, J. E. Barger, M. Brinn, R. J. Mullen, D. Price, S. E. Ritter, and D. Schmitt,

"Boomerang mobile counter shooter detection system," in *In proc. of SPIE Sensors, and C3I Technologies for Homeland Security and Homeland Defense IV*, vol. 5778, Bellingham, WA, USA, 2005.

[3] D. Crane, "Ears-MM soldier-wearable gunshot/sniper detection and location system," *Defence Review*, 2008.

[4] "PILAR Sniper Countermeasures System," http://www.canberra.com, November 2008.

[5] J. Millet and B. Balingand, "Latest achievements in gunfire detection systems," in *Proc. of the RTO-MP-SET-107 Battlefield Acoustic Sensing for ISR Applications*, Neuilly-sur-Seine, France, 2006.

[6] P. Volgyesi, G. Balogh, A. Nadas, C. B. Nash, A. Ledeczi, K. Pence, T. Bapty, J. Scott, and The Nashville Police, "Shooter localization and weapon classification with soldier-wearable networked sensors," in *Proc. of MobiSys*, San Juan, Puerto Rico, 2007.

[7] A. Saxena and A. Ng, "Learning sound location from a single microphone," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, Kobe, 2009.

[8] W. S. Conner, J. Chhabra, M. Yarvis, and L. Krishnamurthy, "Experimental evaluation of synchronization and topology control for in-building sensor network applications," in *Proc. of WSNA*, San Diego, CA, USA, 2003.

[9] O. Younis and S. Fahmy, "A scalable framework for distributed time synchronization in multi-hop sensor networks," in *Proc. of SECON*, Santa Clara, CA, USA, 2005.

[10] J. Elson and D. Estrin, "Time synchronization for wireless sensor networks," in *Proc. of the Int. Parallel and Distributed Processing Symposium*, 2001.

[11] G. Simon, M. Maróti, Ákos Lédeczi, G. Balogh, B. Kusy, A. Nádas, G. Pap, J. Sallai, and K. Frampton, "Sensor network-based counter-sniper system," in *Proc. of SenSys*, Baltimore, MD, USA, 2004.

[12] G. T. Whipps, L. M. Kaplan, and R. Damarla, "Analysis of sniper localization for mobile asynchronous sensors," in *Proc. of the SPIE Signal Processing, Sensor Fusion, and Target Recognition XVIII*, vol. 7336, 2009.

[13] E. Danicki, "Acoustic sniper localization," *Archives of acoustics*, vol. 30, no. 2, pp. 233–245, 2005.

[14] L. M. Kaplan, T. Damarla, and T. Pham, "Qoi for passive acoustic gunfire localization," in *Proc. of the 5th IEEE Int. Conf. on Mobile Ad Hoc and Sensor Systems*, Atlanta, GA, 2008.

[15] D. Lindgren, O. Wilson, F. Gustafsson, and H. Habberstad, "Shooter localization in wireless sensor networks," in *Proc. of the 12th Int. Conf. on Information Fusion*, Seattle, WA, USA, 2009.

[16] R. Stoughton, "Measurements of small-caliber ballistic shock waves in air," *J. of the Acoustical Soc. of America*, vol. 102, no. 2, pp. 781–787, 1997.

[17] F. Gustafsson, *Statistical Sensor Fusion.* Lund: Studentlitteratur, 2010.

[18] E. Danicki, "The shock wave-based acoustic sniper localization," *Nonlinear analysis*, vol. 65, pp. 956–962, 2006.

[19] K. Lo and B. Ferguson, "A ballistic model-based method for ranging direct fire weapons using the acoustic muzzle blast and shock wave," in *Proc. of the Int. conf. on Intelligent Sensors, Sensor Networks and Information Processing*, 2008, pp. 453–458.

[20] S. Kay, *Fundamentals of signal processing – estimation theory.* Prentice Hall, 1993.

[21] N. Patwari, A. H. III, M. Perkins, N. Correal, and R. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. on Signal Processing*, vol. 51, no. 8, 2003.

[22] 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.

[23] 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.