Abstract – Shooter localization is considered in a wireless network of microphones. 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 are 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 we demonstrate that estimation based on these measurements can potentially also reveal the shooting direction beside the position, but again this requires good synchronization. We propose to base the estimation on the time difference of MB and SW at each sensor, which becomes insensitive to synchronization. Cramér-Rao lower bound analysis indicates how a lower bound of the root mean square error depends on the synchronization error for MB and MB-SW differences, respectively. Results from field trials with different type of ammunition show excellent accuracy of the proposed method for both the position and the aiming angle of the shooter.

Keywords: Localization, Sensor Networks, Acoustic, Synchronization

1 Introduction

Several acoustic shooter localization systems are today commercially available, see for instance [1–3]. Typically, one or more 4-element microphone arrays are used, each synchronously sampling acoustic phenomena associated with gunfire. An overview is found in [4]. Some of these systems are mobile, and in [5] it is even described how soldiers can carry the microphone arrays on their helmets.

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, of course, the sensor synchronization. For localization algorithms that rely on accurate timing like TOA and 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 [6–8], the synchronization procedures are associated with costs in battery life or communication resources that usually must be kept at a minimum.

In [9] the synchronization error impact on the localization ability of an urban network is studied by using Monte Carlo simulations. One of the results was that the inaccuracy increased significantly (> 2 m) for synchronization errors exceeding approximately 4 ms. 56 mote sensors were used.

In this paper we derive fundamental estimation bounds for shooter localization systems based on wireless sensor networks, and with the synchronization errors in focus. An accurate method independent of the synchronization errors will be proposed (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. Earlier in [10,11] we contributed with algorithms and bounds for sensor network based localization using energy measurements.

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 MB, SW and MB-SW, respectively. Section 5 derives expressions for the RMSE CRLB for MB and MB-SW, respectively, and provides numerical results from a realistic scenario. Section 6 presents the results from field trials.
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 the (approximative) shape of a cone, with the bullet trajectory as axis, and reaches only microphones that happens to be located in the cone. The SW propagates at the speed of sound, but since it is generated by a supersonic bullet, it always reaches the microphone before the MB, if it reaches the microphone at all. As with the MB, a sufficient number of SW detections may reveal the shooter’s position. The SW detection is however more difficult to utilize, since it depends on the bullet’s speed and ballistic behavior.

Figure 1 shows an acoustic taking 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, 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 text. We will instead assume that the detection error is stochastic with a certain distribution.

Of course, the MB and SW (when present) can be used in conjunction with each other. One of the ideas described and tested 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

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 $\{y_k^{MB}\}_{k=1}^M$ and $\{y_k^{SW}\}_{k=1}^M$, respectively. Each detected time is subject to a detection error $\{e_k^{MB}\}_{k=1}^M$ and $\{e_k^{SW}\}_{k=1}^M$, different for all times, and a clock synchronization errors $\{b_k\}_{k=1}^N$ specific for each microphone. The firing time $t_0$, shooter position $x$ and shooting angle $\alpha$ 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 next section. The notation is summarized in Table 1.

The derived signal models will be on the form

$$y = \varphi^T(x, \theta; p) + e, \quad \text{Cov}(e) = R,$$

where $\varphi$ is a nonlinear function and $\theta$ represents the unknown parameters apart from $x$. Given the sensor locations $\{p_k\}_{k=1}^M$, nonlinear optimization can be performed to estimate $x$, using the nonlinear least squares criterion

$$\hat{x} = \arg\min_{x, \theta} V(x, \theta; p), \quad V(x, \theta; p) = ||y - \varphi^T(x, \theta; p)||_R.$$  

Here, $||v||_R = \sqrt{v^T R^{-1} v}$. Numerical optimization, for instance using the Gauss-Newton method, can here be applied to get the NLS estimate.

However, to reduce the search space in NLS, and thus decreasing the computational burden, a separable least squares framework will be used. This means that we identify a sub-linear structure in the signal model and apply the weighted least squares method to the parameters appearing linearly. For that reason, the signal model (1) is rewritten as

$$y = \varphi_N^T(x, \theta_N; p) + \varphi_L^T(x, \theta_N; p)\theta_L + e.$$
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).

<table>
<thead>
<tr>
<th>Variable</th>
<th>MB</th>
<th>SW</th>
<th>MB–SW</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td>Number of microphones</td>
</tr>
<tr>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td>Number of microphones receiving shock wave</td>
</tr>
<tr>
<td>x</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Position of shooter</td>
</tr>
<tr>
<td>pk</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Position of microphone k</td>
</tr>
<tr>
<td>yk</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>Measured detection time for microphone at position pk</td>
</tr>
<tr>
<td>t0</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>Rifle or gun firing time</td>
</tr>
<tr>
<td>c</td>
<td>L</td>
<td>N</td>
<td>N</td>
<td>Speed of sound</td>
</tr>
<tr>
<td>v</td>
<td></td>
<td>N</td>
<td>N</td>
<td>Speed of bullet</td>
</tr>
<tr>
<td>α</td>
<td></td>
<td>N</td>
<td>N</td>
<td>Shooting angle</td>
</tr>
<tr>
<td>bk</td>
<td>L</td>
<td>L</td>
<td></td>
<td>Clock error for microphone k</td>
</tr>
<tr>
<td>ek</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>Detection error at microphone k</td>
</tr>
<tr>
<td>r</td>
<td></td>
<td>N</td>
<td>N</td>
<td>Bullet speed decay rate</td>
</tr>
<tr>
<td>dk</td>
<td></td>
<td></td>
<td></td>
<td>Point of origin for shock wave received by microphone k</td>
</tr>
<tr>
<td>β</td>
<td></td>
<td></td>
<td></td>
<td>Mach angle, ( \sin \beta = c/v )</td>
</tr>
</tbody>
</table>

The NLS problem can then be formulated as

\[ \hat{x} = \arg \min_{x, \theta_L} V(x, \theta_N; \theta_L; p), \]

\[ V(x, \theta_N, \theta_L) = \|y - \varphi_L(x, \theta_N) - \varphi_L^T(x, \theta_N)\theta_L\|_R. \quad (4) \]

Since \( \theta_L \) enters linearly it can be solved for by linear least squares (suppressing the arguments of \( \varphi_L(x, \theta_N, p) \) and \( \varphi_N(x, \theta_N, p) \))

\[ \hat{\theta}_L = \arg \min_{\theta_L} V(x, \theta_N, \theta_L; p) \]

\[ = \varphi_L R^{-1} \varphi_L^T \left( y - \varphi_N \right). \quad (5) \]

This simplifies the nonlinear minimization to

\[ \hat{x} = \arg \min_{x, \theta_N} V(x, \theta_N, \hat{\theta}_L; p) \]

\[ = \arg \min_{x, \theta_N} \|y - \varphi_N^T + \varphi_L (\varphi_L^T)^{-1} \varphi_L (y - \varphi_N)\|_R. \quad (6) \]

### 4 Signal Models

#### 4.1 Muzzle Blast model

According to the clock at microphone \( k \), the muzzle blast (MB) sound is assumed to reach \( pk \) at the time instant

\[ y_k = t_0 + b_k + \frac{1}{2} |p_k - x| + e_k. \quad (7) \]

Here, and throughout this text, \( ||\cdot|| \) will denote the vector 2-norm. The shooter position \( x \) and the microphone at location \( p_k \) are two-dimensional vectors in the examples to follow. The developed theory however works equally well in three dimensions. Collectively for all \( M \) microphones, the model is represented on vector form as

\[ y = b + \varphi_T^T (p - xT) \theta_L + e, \quad (8) \]

where

\[ \theta_L = [t_0 \ 1/c]^T, \quad (9a) \]

\[ \varphi_L (p - xT 1_M) \text{Col}_k = [1 \ ||p_k - x||]^T. \quad (9b) \]

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 is used as a parameter, since it enters linearly.

\( \varphi_L \) is 2-by-\( n \), where \( n = 2 \) for a plane model, and \( n = 3 \) for a model in the room. The \( \cdot_L \) notation indicates that \( \cdot \) is part of a linear relation, as described in the previous section. With \( \varphi_N = 0 \) and \( \varphi_L = \varphi_L(x; p) \), the derived least squares criterion is

\[ \hat{x} = \arg \min_{x} - y^T R^{-1} \varphi_L^T \left( \varphi_L R^{-1} \varphi_L^T \right)^{-1} \varphi_L R^{-1} y. \quad (10) \]

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 (10) in logarithmic scale for data from a field trial (\( R = \sigma^2 I \) assumed). It appears, that there are only two local minima.

At an arbitrary moment, the detection error and clock bias are assumed to be independent stochastic variables with normal distribution, \( e \in N(0, R^m) \), \( b \in N(0, R^b) \).

#### 4.2 Shock Wave Model

It is 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 trajec-
tory, the speed is reduced to

\[ v = v_0 - r||d_k - x||, \tag{11} \]

where \( r \) is an assumed known ballistic parameter. This is a very coarse bullet trajectory model, compared with, for instance, the curvilinear trajectories proposed by [12], but we propose it for simplicity. It gives quite an acceptable speed approximation for, say, \( v_0 > 700 \text{ m/s} \) and \( ||d_k - x|| < 400 \text{ m} \).

The shock wave that leaves the supersonic bullet trajectory propagates with the speed of sound \( c \) at the angle \( \beta_k \) relative the bullet heading. \( \beta_k \) is the Mach angle defined as

\[ \sin \beta_k = \frac{c}{v} = \frac{c}{v_0 - r||d_k - x||}. \tag{12} \]

\( d_k \) is now the coordinates of the (fictitious) point where the bullet generates that piece of the shock wave that eventually reaches microphone \( k \). The time it takes the bullet to reach \( d_k \) is

\[ \int_0^{||x-d_k||} \frac{d \xi}{v_0 - r \cdot \xi} = \frac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||}. \tag{13} \]

Adding this time to the time it takes the shock wave to propagate from \( d_k \) to \( p_k \) gives the total time from firing to detection,

\[ y_k = t_0 + b_k + \frac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||} + \frac{1}{r}||d_k - p_k|| + e_k, \tag{14} \]

according to the clock at microphone \( k \). Note that the variable names \( y \) and \( e \) for notational simplicity been reused from the MB model. Below, also \( \varphi, \theta_N, \) and \( \theta_L \) will be reused. When there is ambiguity, a superscript will indicate exactly which entity that is referred to, for instance, \( y^{\text{MN}}, \varphi^{\text{SW}}. \)

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

\[ \frac{\sin(90^\circ - \beta_k - \gamma_k)}{||d_k - x||} = \frac{\sin(90^\circ + \beta_k)}{||p_k - x||}, \tag{15} \]

which together with (12) implicitly defines \( d_k \). We have not found any simple closed form for \( d_k \) so we solve for \( d_k \) numerically. \( \gamma_k \) is trivially induced by the shooting direction \( \alpha \) (and \( x, p_k \)).

With notation similar to the one used with the muzzle blast model, the collective vector form is

\[ y = b + \varphi_N^T(x, \theta_N; p) + \varphi_L^T(x; p)\theta_L + e, \tag{16} \]

where

\[ \varphi_L(x; p) = 1, \tag{17} \]

\[ \theta_L = t_0, \tag{18} \]

\[ \theta_N = \left[ \frac{1}{c} \alpha^T \right] v_0^T, \tag{19} \]

and where the elements of \( \varphi_N(x, \theta_N; p) \in \mathbb{R}^{1 \times S} \) are

\[ \varphi_N(x, \theta_N; p_k) = \frac{1}{r} \log \frac{v_0}{v_0 - r||d_k - x||} + \frac{1}{r}||d_k - p_k||, \tag{20} \]

\( d_k \) being the reasonable solution to (12), (15).

It is assumed that the shock wave detection error is stochastic with normal distribution, \( e \in \mathcal{N}(0, \sigma_{SW}^2) \), independent of \( e^{\text{MN}} \) and \( b \).

### 4.3 MB–SW Model

With the motivation to maintain an accurate localization despite synchronization errors, we propose the MB–SW model:

\[ y_k = y_k^{\text{MN}} - y_k^{\text{SW}} \]

\[ = \varphi_L^{\text{MN}}(x; p_k)\theta_L^{\text{MN}} - \varphi_N^{\text{SW}}(x, \alpha^{\text{SW}}; p_k) - \varphi_L^{\text{SW}}(x; p_k)\theta_L^{\text{SW}} + e^{\text{MN}}_k - e^{\text{SW}}_k, \tag{21} \]

for \( k = 1, 2, \ldots S \). The key idea is that \( y \) is by cancelation independent of both the firing time to 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, \( e^{\text{MN}}_k - e^{\text{SW}}_k \).
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.

The Vector form for the MB–SW model is

$$y = \varphi_N^T(x, \theta_N; p) + e,$$

where

$$\varphi_N(x, \theta_N; p) = \frac{1}{c}||p_k - x|| - \frac{1}{r} \log \frac{v_0}{v_0 - r}||d_k - x|| - \frac{1}{c}||d_k - p_k||,$$

and $y = y_{\text{MB}} - y_{\text{SW}}$, $e = e_{\text{MB}} - e_{\text{SW}}$. As before, $d_k$ is the reasonable solution to (12), (15). The error $e$ has zero mean and covariance

$$R = R_{\text{MB}} + R_{\text{SW}}.$$

The MB–SW least squares criterion is

$$\hat{x} = \arg \min_{x, \theta_N} ||y - \varphi_N(x, \theta_N; p)||_R$$

which requires numerical optimization. A search-based method could use $\hat{x}_{\text{MB}}$ as a start value for the search. Initial shooting direction could be given by, for instance, assuming that the shooter aims at a point close to the microphone network.

5 Cramér-Rao Lower Bound

The accuracy of localization based on the rather general model

$$y = \varphi(\eta) + e$$

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

$$\text{Cov} \hat{\eta} \geq \mathcal{I}^{-1}(\eta_0).$$

Here, the location $x$ is for notational purposes part of the parameter vector $\eta$. The Cramér-Rao lower bound provides a fundamental estimation limit for unbiased estimators, see [13]. This bound has been analyzed thoroughly in the literature, primarily for AOA, TOA, and TDOA, [14–16].

The square matrix $\mathcal{I}$ is the Fisher information matrix, which for $e \in \mathcal{N}(0, R)$ takes the form

$$\mathcal{I}(\eta) = \nabla_\eta \varphi^T(\eta) R^{-1} \nabla_\eta \varphi(\eta).$$

The bound is evaluated for a specific location, parameter setting and microphone positioning, collectively $\eta = \eta_0$.

5.1 MB Case

For the MB case, the entities in (28) are identified by

$$\eta = [x^T \theta_N^T]^T, \quad (29)$$

$$\varphi(\eta) = \varphi_N(x, \theta_N; p), \quad (30)$$

$$R = R_{\text{MB}}, \quad (31)$$

The Jacobian $\nabla_\eta \varphi$ is an $n+2$-by-$n+2$ matrix, $n$ being the dimension of $x$. The bound is

$$\text{Cov} \hat{x} \geq [I_n \ 0] \mathcal{I}^{-1} \begin{bmatrix} I_n \\ 0 \end{bmatrix}. \quad (32)$$

The LS solution in (5) however gives a shortcut to an n-by-n Jacobian,

$$\nabla_x \varphi_L \hat{\theta}_L = \nabla_x \varphi_L^T (\varphi_L R^{-1} \varphi_L^T)^{-1} \varphi_L R^{-1} y_0$$

for $y_0 = \varphi_L^T (p - 1 M x_0) \theta_L$, 0. At least for $n = 2$, this Jacobian can, with some effort, be expressed explicitly on a closed analytical form. The equivalent bound is

$$\text{Cov} \hat{x} \geq [\nabla_x \varphi_L^T \hat{\theta}_L]^T R^{-1} [\nabla_x \varphi_L^T \hat{\theta}_L]^{-1}. \quad (33)$$

5.2 MB–SW Case

For the MB–SW case,

$$\eta = [x^T \theta_N^T]^T, \quad (35)$$

$$\varphi(\eta) = \varphi_N(x, \theta_N; p), \quad (36)$$

$$R = R_{\text{MB–SW}}, \quad (37)$$

and the estimation bound is analogous to (32) above. The $2n + 1$-by-$2n + 1$ Jacobian $\nabla_\eta \varphi_N(x, \theta_N; p)$ is probably best evaluated by finite difference methods.

5.3 Numerical Example

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

$$R_{\text{MB}} + R_{\text{SW}} = (\sigma_e^2 + \sigma_d^2) I,$$

$$R_{\text{MB}} = R_{\text{SW}} = \sigma_e^2 I,$$

and that $\alpha = 0$, $\epsilon = 330$ m/s, $v_0 = 700$ m/s, $r = 0.63$. The scenario setup implies that all microphones detect the shock wave, $S = M$. 

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.

Figure 5: Scene of the shooter localization field trial. There are 10 microphones, 3 shooter positions, and a common target.

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

$$\text{Cov} \hat{x}_{MB} \geq \sigma_{det+b}^2 \begin{bmatrix} 64 & -17 \\ -17 & 9 \end{bmatrix} \cdot 10^4.$$ (38)

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

$$\text{RMSE} (\hat{x}_{MB}) \geq \sqrt{\text{tr} \text{Cov} \hat{x}_{MB}} \approx 858 \sigma_{det+b} \text{[m]}.$$ (39)

Monte Carlo simulations (not described here) indicate that the NLS estimator attains this lower bound for $\sigma_{det+b} < 0.1 \text{ s}$.

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

$$\text{Cov} \hat{x}_{MB-SW} \geq \sigma_{det}^2 \begin{bmatrix} 28 & 5 \\ 5 & 12 \end{bmatrix} \cdot 10^5.$$ (40)

RMSE ($\hat{x}_{MB-SW}$) $\geq 2000 \sigma_{det} \text{[m]}.$ (41)

Discussion In this example, 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.

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 5. 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 rifle and one pistol ammunition are used, see Table 2. Four rounds are fired of each ammunition type 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 of a horizontal plane. 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 muzzle blast and shock wave reception times are determined 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 $\sigma_{SW} \approx 80 \mu$s, and the muzzle blast error $\sigma_{MB}$ is slightly worse. For shooting position 3 the accuracies are generally 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 only the muzzle blast model (7) is conducted by minimizing (10) over a 10 m grid well covering the area of interest, followed by numerical minimization.

Localization using the difference model (21) is done by numerically minimizing (25). The objective function is subject to local optima, why 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 angle $\alpha$. Initial bullet speed is $v = 800 \text{ m/s}$ and
Table 2: Ammunition used at the trial. The table accounts for how many rounds that are fired from the different positions. Also, the resulting localization RMSE for the MB–SW model is noted. However, for the Luger Pistol the MB model RMSE is given instead, since only one microphone is located in the SW cone.

<table>
<thead>
<tr>
<th>Type</th>
<th>Caliber</th>
<th>Weight</th>
<th>Velocity</th>
<th>Positions</th>
<th>#Rounds</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>308 Winchester</td>
<td>7.62 mm</td>
<td>9.55 g</td>
<td>847 m/s</td>
<td>1, 2, 3</td>
<td>4, 4, 6</td>
<td>11 m</td>
</tr>
<tr>
<td>Heavy Bullet</td>
<td>9.3 mm</td>
<td>15 g</td>
<td>767 m/s</td>
<td>1, 2, 3</td>
<td>4, 4, 4</td>
<td>23 m</td>
</tr>
<tr>
<td>Swedish Mauser</td>
<td>6.5 mm</td>
<td>8.42 g</td>
<td>852 m/s</td>
<td>1, 2, 3</td>
<td>4, 4, 4</td>
<td>5.6 m</td>
</tr>
<tr>
<td>Luger Pistol</td>
<td>9 mm</td>
<td>6.8 g</td>
<td>400 m/s</td>
<td>3</td>
<td>-,-,4</td>
<td>1.4 m</td>
</tr>
</tbody>
</table>

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 6 shows, in three zoomed in parts of the scene, the resulting position estimates based on the MB model (blue crosses) and based on the MB–SW model (blue 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 7 visualizes the shooting direction estimates, \( \hat{\alpha} \). Estimate root mean square errors (RMSE) for the three shooter positions are given in Table 3. These statistics indicate that the use of the shock wave from distant shooters cut the error by at least 75%.

Since all microphones are recorded by a common recorder, there are actually no timing errors due to inaccurate clocks. But there are apparently other sources of error, worth identifying.

At least, there are two explanations for the bad performance using the MB–SW model at shooter position 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 difference model harder. This is illustrated by the numerical example in Section 5.3, where \( RMSE, \hat{x} \geq 858\sigma_{det+b} \) but \( RMSE, \hat{x} \geq 2000\sigma_{det} \).

Most likely, 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 6 (red squares), three clusters are identified that correspond to three ammunition types with different ballistic properties. This clustering indicates that the ballistic model is not perfect. Working with three-dimensional data in the plane is of course another model discrepancy.

Finally, we face numerical uncertainties. There is no guarantee that the numerical minimization programs we have used really deliver the global minimum.

Table 3: Localization and aim RMSE for the three different shooter positions using the MB and the MB-SW models. The aim error, given for the MB-SW estimates only, is with respect to the aim at \( \hat{x} \) against the target, not with respect to the true angle \( \alpha \). This way the ability to identify the target is assessed.

<table>
<thead>
<tr>
<th>Sh. pos.</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x^M )</td>
<td>105 m</td>
<td>28 m</td>
<td>2.4 m</td>
</tr>
<tr>
<td>( x^D )</td>
<td>26 m</td>
<td>5.7 m</td>
<td>5.2 m</td>
</tr>
<tr>
<td>Aim</td>
<td>0.041°</td>
<td>0.14°</td>
<td>17°</td>
</tr>
</tbody>
</table>

7 Conclusions

We have here presented a framework for estimation of shooter localization 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 SW for the aiming angle. A separable nonlinear least squares (SNLS) framework was proposed to limit the parametric search space and to enable the use for global grid-based (grid over the shooter position) optimization algorithms, 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, we propose to use the time difference of MB and SW at each microphone, where the clock offset is automatically eliminated. The effective number of measurements decreases in this approach, but as the CRLB analysis showed, the root mean square position error decreases significantly in practical networks.

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.
Figure 6: Estimated positions ˆ\(x\) 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.

Figure 7: Estimated shooting directions. The pistol ammunition is excluded since only one microphone detects the shock wave from the relatively slow pistol ammunition.

References