IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 May 2026, accepted 10 June 2026, date of publication 12 June 2026, date of current version 2 July 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3703320

RESEARCH ARTICLE

Acoustic Shooter Localization With Consistency-Function Fusion in Sparse Sensor Networks

ÁKOS LÉDECZI[1], (Senior Member, IEEE), JÁNOS SALLAI[2], PÉTER VÖLGYESI[1], MIKLÓS MARÓTI[3], WILL HEDGECOCK[1], (Member, IEEE), AND GYULA SIMON[4], (Senior Member, IEEE)

[1]Institute for Software Integrated Systems, Vanderbilt University, Nashville, TN 37212, USA
[2]Xtremis AI, Washington, DC 20007, USA
[3]University of Szeged, 6720 Szeged, Hungary
[4]Obuda University, 1034 Budapest, Hungary

Corresponding author: Ákos Lédeczi (akos.ledeczi@vanderbilt.edu)

This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) Network Embedded Systems Technology (NEST), Advanced Soldier Sensor Information Systems and Technology (ASSIST) and Transformative Apps Programs; and in part by the National Science Foundation under Grant 2312391 and Grant 2527317.

ABSTRACT Acoustic localization of supersonic rifle fire in urban environments in military and security operations faces challenges from multipath propagation and non-line-of-sight detections. Physics-based methods predominate because the geometric ground-truth data required for machine learning is hard to collect at scale. Building on consistency-function fusion–an approach that identifies shooter hypotheses by counting supporting sensor observations rather than minimizing measurement residuals–this paper compares two sparse-network architectures: a wearable multi-channel-array system and a single-microphone deployment exploiting geometric shockwave constraints. We introduce two algorithmic refinements: a caliber-aware deceleration correction that removes a systematic position bias growing quadratically with shooter range, and a coplanar-geometry recovery for the multi-channel pipeline that handles trajectories wide of the sensor field. Under noiseless evaluation on the Aberdeen 2006 geometry the refinements drive position error from ∼17 m to ∼1 m for the multi-channel pipeline and from ∼19 m to ∼2 m for the single-channel pipeline. A high-fidelity simulator with a noise budget representing the Aberdeen hardware reproduces the published position MAE values to within a few percent for both architectures. Under mobile-deployment noise the multi-channel pipeline retains both better position accuracy (23 vs 26 m all-range MAE) and a 17 percentage-point localization-rate advantage that widens to nearly 29 points with moderate acoustic multipath. Architecture choice thus emerges as an environment-dependent trade-off–single-channel for cost-constrained line-of-sight deployments, multi-channel for cluttered environments–rather than a universal hierarchy.

INDEX TERMS Shooter localization, measurement fusion, acoustic sensing, wireless sensor networks.

NOMENCLATURE

AoA	Angle of Arrival.
BC	Ballistic Coefficient (G1 reference projectile).
MAE	Mean Absolute Error.
MB	Muzzle Blast.
MC	Multi-Channel (four-microphone array per node).
RIPS	Routing Integrated Time Synchronization.
SC	Single-Channel (one microphone per node).
SW	Shockwave.
TDoA	Time Difference of Arrival.
ToA	Time of Arrival.

The associate editor coordinating the review of this manuscript and approving it for publication was Dominik Strzalka.

I. INTRODUCTION

Accurate localization of supersonic rifle fire has been a critical objective for military force protection, counter-sniper operations, and security applications for several decades.

© 2026 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

Unlike subsonic handgun fire typical of law enforcement scenarios, supersonic projectiles generate ballistic shockwaves in addition to muzzle blasts; these phenomena provide rich geometric constraints for trajectory estimation. The ability to rapidly identify both the projectile path and the hostile shooter position is essential for force protection. While various technological approaches have been explored – including optical, infrared, and radar methods – acoustic systems remain the most widely deployed solution due to their omnidirectional coverage, passive operation, and exploitation of well-understood sound propagation principles.

Despite the recent surge in machine learning (ML) applications for acoustics, data-driven methods have not yet been applied to the specific problem of shooter *localization*. Deep-learning models have achieved remarkable success in gunshot *detection* and weapon *classification* [1], [2], demonstrating the potential of learned representations for acoustic event analysis. However, to our knowledge, no published work has demonstrated ML-based supersonic gunshot localization. This gap likely reflects practical constraints: unlike classification tasks, which can use crowd-sourced audio of muzzle blasts (e.g., YouTube datasets), localization requires precise geometric ground truth across diverse acoustic environments – data that are expensive and logistically difficult to collect at scale. Physics-based fusion algorithms, which exploit invariant shockwave geometry rather than learned features, thus remain the predominant approach for robust gunshot localization.

Acoustic localization in urban environments faces significant measurement challenges regardless of the architecture chosen. Multipath propagation, non-line-of-sight detections, and rapid succession of shots introduce large, non-Gaussian errors that cause traditional least-squares fusion techniques to fail. Conventional systems typically require expensive sensor arrays with precise geometric calibration and ongoing maintenance, limiting their practical deployment. Over the past two decades, our research group has developed a fusion methodology – consistency-function fusion – that reframes localization from solving an overdetermined equation system to identifying shooter position hypotheses supported by the maximum number of sensor observations [3]. The approach inherently rejects outliers without an explicit outlier-detection step and tolerates uncalibrated commodity microphones, GPS-class sensor positions, and the occasional bad detection.

The consistency-function paradigm originated with a dense network of single-channel sensors deployed at the McKenna MOUT urban training site [3] and was subsequently adapted to two architectures suitable for sparse, soldier-wearable deployments: a multi-channel system with four-microphone arrays per node that extract angle-of-arrival information [4], and a single-microphone system that elevates shockwave duration from a post-localization caliber estimate to a primary geometric constraint [5]. The dense system is described here as the canonical setting where the paradigm originated; we do not re-evaluate it because it shares neither the deployment density nor the sensor mobility of the other two and so cannot be compared on equal footing.

This paper focuses on the two sparse architectures and makes the following four contributions:

- **Two algorithmic refinements**: a caliber-aware deceleration correction (Section VII-A) that removes a previously unrecognized systematic position bias growing quadratically with shooter range; and a coplanar-geometry recovery for the multi-channel pipeline (Section VII-B) that handles the small fraction of shots whose trajectory passes wide of the sensor field, a failure mode acknowledged in the original Aberdeen evaluation [4].
- **Significant position-error reduction** evaluated under various geometries and noise budgets. On the Aberdeen 2006 geometry under noiseless conditions the two refinements together drive all-range position error from 16.6 m to 1.0 m for the multi-channel pipeline and from 19.5 m to 2.4 m for the single-channel pipeline – essentially convergent recovery of ground truth at every range bin (Section VIII-B).
- **A high-fidelity simulator** that produces synthetic per-(shot, sensor) detections in the same format as the field data, allowing both pipelines to run on simulator and field input through identical code paths. With a calibrated noise budget reflecting the documented Aberdeen 2006 hardware, the simulator-driven pipelines reproduce the published Aberdeen position MAE (19.17 m vs published 23.00 m for the multi-channel system; 3.01 m vs 2.96 m for the single-channel system on the published $\leq$130 m subset) within their 95% confidence intervals (Section VIII-C).
- **A comparative study of the two architectures** under three noise scenarios: noiseless, Aberdeen-calibrated, and a more demanding mobile-deployment profile that adds GPS-class sensor positioning error, compass-class orientation error, and acoustic multipath. The multi-channel pipeline dominates on both metrics under mobile-deployment noise: better position accuracy (23 vs 26 m all-range MAE) and a 17 percentage point localization-rate advantage that widens to nearly 29 percentage points when acoustic multipath is added (Sections VIII-D–VIII-E).

The headline finding is that architecture choice is **not a universal hierarchy** but an environment-dependent trade-off. The single-channel architecture's hardware simplicity and elimination of orientation tracking deliver comparable accuracy to the multi-channel system on benign data at materially lower cost, while the multi-channel architecture's larger evidence pool delivers meaningfully higher localization rate in cluttered, multipath-rich environments. The original Aberdeen result that single-channel sensors achieved better accuracy than multi-channel arrays on the same field data [5] holds under the conditions of that test (an open range with surveyed sensor positions and calibrated orientations), but does not generalize to deployments where soldier-worn

sensors operate without those calibration aids and where multipath is part of the environment.

The remainder of the paper is organized as follows. Section II reviews the acoustic phenomena and core measurements that underpin both architectures. Section III introduces the consistency-function fusion paradigm in its canonical dense single-channel setting. Sections IV and V describe the sparse multi-channel and sparse single-channel pipelines respectively. Section VI summarizes the published Aberdeen 2006 baselines for both sparse architectures. Section VII introduces the two algorithmic refinements. Section VIII presents the simulator-based comparative evaluation across different scenarios. Section IX surveys related work, and Section X discusses the architecture trade-off and avenues for future work.

II. BACKGROUND

A firearm discharge generates two acoustic phenomena exploitable for localization: the muzzle blast and, when the projectile is supersonic, the ballistic shockwave (Figure 1). The muzzle blast is a spherical pressure wave that propagates omnidirectionally from the shooter at the speed of sound ($c \approx 343$ m/s). The ballistic shockwave, also known as the sonic boom or N-wave, forms a conical Mach wave trailing the projectile when its velocity exceeds c. At a sensor it presents as a characteristic N-shaped pressure pulse: a sharp positive overpressure (sub-microsecond rise time), a rapid transition through ambient pressure, a negative underpressure phase, and a return to ambient. The total duration is typically 100–600 μ s, depending on projectile characteristics and the perpendicular distance from the trajectory to the sensor (the *miss distance*). See Figure 2.

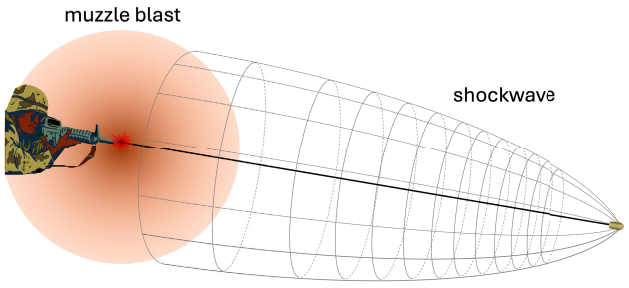


FIGURE 1. Acoustic phenomena associated with a gunshot with a supersonic projectile.

Stoughton [6], building on Whitham's work [7], derived the shockwave duration T_{shock} as a function of projectile and propagation geometry:

$$T_{shock} = \frac{1.82\, Mb^{1/4}}{c(M^2-1)^{3/8}} \frac{d}{l^{1/4}} \tag{1}$$

where M is the projectile Mach number, l is the projectile length, d is the projectile diameter (caliber), and b is the miss distance. The dominant scaling is with caliber: bullet diameter is directly proportional to shockwave duration.

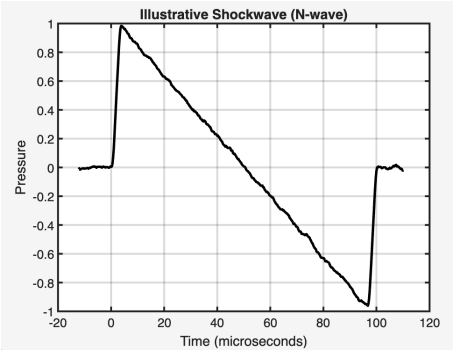


FIGURE 2. Ballistic Shockwave (N-wave) Illustration.

Both the muzzle blast and the shockwave can be detected with a microphone. The shockwave's sub-microsecond rising edge yields microsecond-precision time-of-arrival (ToA). The muzzle blast's leading edge is sharp in the near-field but progressively broadens with distance as high-frequency content attenuates; at typical engagement ranges of 50–300 m it still provides tens-of-microseconds onset precision. A small array of microphones with known geometry sharing a clock can additionally estimate angle-of-arrival (AoA) from either signal via standard time-difference-of-arrival (TDoA) methods. Six measurable quantities are therefore available, each carrying distinct geometric information about the shooter, the trajectory, or the bullet:

- **Muzzle blast ToA** – combined across multiple sensors, multilaterates the shooter position.
- **Muzzle blast AoA** – bearing from the sensor to the shooter.
- **Shockwave ToA** – constrains the bullet's emission point along the trajectory.
- **Shockwave AoA** – bearing from the sensor to the wavefront emission point on the trajectory.
- **Shockwave duration** – N-wave length set by Whitham's equation. Constrains the miss distance between the trajectory and the sensor.
- **Shockwave–muzzle-blast TDoA at one sensor** – closed-form shooter-range constraint, given bullet velocity and miss distance.

The remainder of this section reviews the principal sensor configurations used by acoustic shooter-localization systems; for clarity some illustrations restrict the geometry to a plane, but all results generalize to 3D.

A. SINGLE MICROPHONE ARRAY

A single 4-microphone array (3 mics for 2D) measures all six quantities. The MB and SW ToAs and AoAs alone are sufficient to recover the shot time and shooter location: the angle α between the muzzle blast (MB) and shockwave (SW) arrival directions, combined with the TDoA between the two events, yields the shot time via $t = (t_s - t_m \cos\alpha)/(1-\cos\alpha)$. The MB AoA then provides bearing to the shooter; the recovered shot time provides range. Figure 3 illustrates the

geometry, and [4] gives the 3D derivation. This formulation assumes constant projectile velocity, an approximation we revisit in Section VII-A.

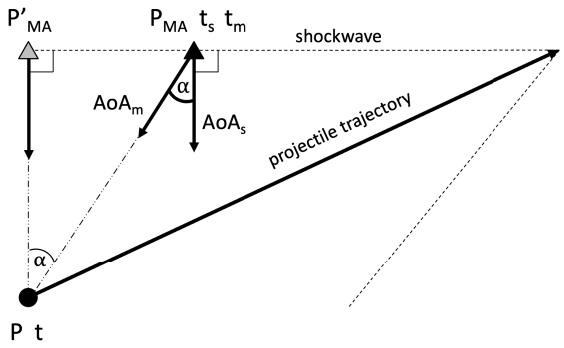


FIGURE 3. Geometry of a shot from point *P* measured by a microphone array at point P_{MA}.

B. DISTRIBUTED SENSING

Distributed sensors with synchronized clocks at known positions enable various complementary approaches.

Multilateration on muzzle blast ToAs. With at least three non-collinear sensors (four in 3D), MB TDoA between sensor pairs constrains the shooter through hyperbolic intersections. Additional sensors improve accuracy and enable outlier rejection. The approach faces two operational challenges: distant shooters produce weak MB signals (reducing detection reliability), and when sensors are clustered relative to a distant shooter the hyperbolic intersections occur at shallow angles, producing severe geometric dilution of precision.

Trajectory recovery from shockwave ToAs. The projectile trajectory itself – the source of the shockwave – typically passes near or through the protected area where sensors are deployed, producing higher-SNR detections than the distant MB. Pairs of single-channel sensors yield shockwave-direction estimates from their TDoAs (Figure 4); three or more sensors resolve the inherent two-fold ambiguity. Two practical limitations remain: the basic three-sensor approach assumes all sensors lie on the same side of the trajectory, and the Mach angle depends on the unknown projectile velocity. Sensors detecting the shockwave on *both* sides of the trajectory enable trajectory reconstruction through geometric consistency, simultaneously resolving the side ambiguity and recovering projectile velocity. Numerical optimization extends the same principle to arbitrary sensor placement [8]. The trajectory does not directly yield bearing or range; combining it with at least two MB ToAs produces the shooter location through hyperbola-trajectory intersection.

Multi-array deployments enable a closed-form trajectory recovery from any pair of arrays via the Danicki formula [9], which fails when the two SW AoAs become near-parallel (i.e., when the trajectory passes outside the line connecting the two arrays). Multiple arrays also enable caliber and weapon classification by combining the shockwave-duration constraint of Equation 1 with the recovered trajectory and projectile speed; we revisit this in Section IV-E.

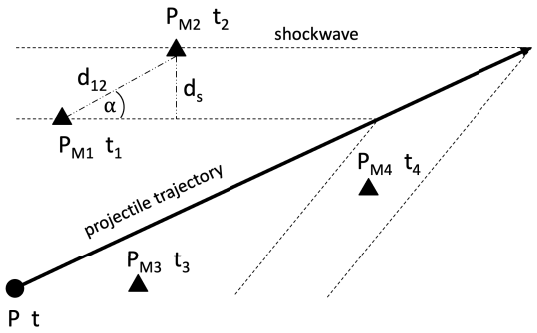


FIGURE 4. Shockwave ToAs measured by distributed microphones.

An alternative line of work [10], [11] uses the shockwave-muzzle-blast TDoA *at each sensor*, which is measured locally and removes the need for inter-node time synchronization, at the cost of requiring both events to be detected at every contributing sensor.

C. A SINGLE MICROPHONE

A single microphone can measure MB and SW ToAs and the shockwave duration T_{shock}. Used alone, the duration enables caliber estimation once other unknowns in Whitham's formula are constrained [12]; conversely, with a known or assumed weapon, the duration estimates the miss distance b [13]. Combining MB-SW TDoA at a sensor with a miss-distance estimate and an assumed projectile velocity yields a closed-form shooter range from the Mach-cone geometry; [5] reported a 4.7% mean range error for 168 AK-47 shots with ranges 50–130 m using this single-sensor approach. The known-weapon assumption and absence of bearing estimation limit single-sensor practicality, but the result motivates the sparse single-channel architecture (Section V), which fuses these single-sensor estimates across a network without assuming the weapon a priori.

D. MEASUREMENT CHALLENGES

The clean derivations above assume well-behaved measurements. In practice – particularly in urban environments – multipath propagation from building reflections, non-line-of-sight detections due to obstructions, and overlapping shots from multiple near-simultaneous events introduce large, non-Gaussian errors into both ToA and AoA measurements. Traditional least-squares multilateration is highly sensitive to these outliers: a single erroneous detection can severely degrade localization accuracy. The consistency-function fusion methodology discussed in this paper addresses this by reframing localization from minimum-residual estimation to maximum-consensus search, inherently rejecting outliers without requiring explicit outlier detection or distributional assumptions.

III. CONSISTENCY-FUNCTION FUSION

The two sparse architectures evaluated in this paper share a common fusion methodology: the consistency function. Originally developed for a dense network of single-channel sensors deployed at the McKenna MOUT urban training site [3], [14], the methodology has since been adapted to the sparse multi-channel and sparse single-channel pipelines (Sections IV–V) by extending the function to incorporate richer per-sensor evidence (per-array shooter estimates, AoAs, shockwave-derived range estimates, and so on). This section describes the methodology in its original ToA-only form because the algebra is cleanest there; the sparse-pipeline extensions reduce to the same maximum-consensus search over a different evidence set. The dense single-channel system itself was extensively evaluated in [14]; we do not re-evaluate it here because it does not share the deployment density or sensor mobility of the two sparse architectures, so it cannot be compared on equal footing in the simulation studies of Section VIII.

The original ToA-only consistency function [14] operates as follows. Sensors detect N muzzle blast events with detection times t_i, $i = 1, 2, \ldots, N$, at sensor positions $p_i = (x_i, y_i, z_i)$. Detections typically correspond to a single shot but may include multiple simultaneous shots, and individual measurements may be correct or contain outliers from reverberations, line-of-sight blockages, or wrong-shot mispairings. Traditional closed-form and iterative techniques exhibit bias unless outliers are removed prior to estimation [15], [16]; the consistency-based approach achieves accurate estimation even when outliers outnumber correct measurements.

If the shooter is at position (x, y, z) and the shot time is t then theoretically the ith measurement is

$$t_i(x, y, z, t) = t + \frac{d_i(x, y, z)}{c} \tag{2}$$

where

$$d_i(x, y, z) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \tag{3}$$

is the distance between the ith sensor and shot position (x, y, z). For perfect measurements $t_i(x, y, z, t) - t_i$ would be 0. For real measurements, let us define the error tolerance τ, which is a cumulative uncertainty, stemming from the inherent limitations in sensor positioning, time synchronization, event detection, and time measurement. Let us also define the tolerance limit function $\Lambda_\tau(x)$:

$$\Lambda_\tau(a) = \begin{cases} 1, & \text{if } |a| <= \tau \\ 0, & \text{otherwise} \end{cases}$$

The value of $\Lambda_\tau\left(t_i(x, y, z, t) - t_i\right)$ is 1 for correct, but inherently somewhat imperfect, real measurements, while for outliers, it is 0. The consistency function is defined as follows:

$$C_\tau(x, y, z, t) = \sum_{i=1}^{N} \Lambda_\tau\left(t_i\left(x, y, z, t\right) - t_i\right) \tag{4}$$

The function $C_\tau(x, y, z, t)$ provides the number of consistent measurements that support the hypothesis that the shot was taken in position (x, y, z) at time t, with uncertainty τ. When evaluated in the true position and time of the shot, $C_\tau(x, y, z, t)$ provides the number of correct sensor measurements (maximum N). If evaluated at an arbitrary point (x, y, z, t), the number of supporting sensors will be small (if any), thereby resulting in a small value for $C_\tau(x, y, z, t)$. Consequently, the estimate of the shot location and shot time will be at the global maximum of the consistency function.

Notes:

- The consistency function maximum typically spans a region (max area) rather than a single point, with size dependent on τ. The location estimate is computed as the centroid of this region or through local optimization starting from it.
- Multiple shots appear as distinct peaks with different shot times, enabling multi-shot localization. Echoes produce peaks sharing the shot time of direct detections but with smaller consistency values, allowing straightforward filtering.
- Multipath effects may create local maxima, but the global maximum consistently corresponds to the true shooter position. While artificially constructed outlier sets can theoretically defeat the algorithm, field tests demonstrate remarkable resilience even when outliers outnumber valid measurements [3].

The determination of the global maximum of the integer-valued, non-continuous consistency function, which may possess several local maxima, poses a significant challenge. The brute-force evaluation of function (4) on a grid is a theoretical possibility, however, it is not feasible in real-life scenarios. Alternatively, a fast search method may be applied, which finds the global maximum of (4) on a grid.

The first step toward a fast search method is an important observation: the maximum of the 4-dimensional consistency function (4) is the same as the maximum of the 3-D consistency function, defined as follows:

$$C_\tau(x, y, z) = \max_t \sum_{i=1}^{N} \Lambda_\tau\left(t_i\left(x, y, z, t\right) - t_i\right) \tag{5}$$

Using (2), let us define variable $t_i'(x, y, z) = t_i - d_i(x, y, z)/c$, which is the estimate of the shot time from the measurement of the ith sensor, provided the shot was taken at position (x, y, z). Now the 3-D consistency function (5) can be written as $C_\tau(x, y, z) = \max_t \sum_{i=1}^{N} \Lambda_\tau\left(t - t_i'(x, y, z)\right)$. From this form it is apparent that finding $C_\tau(x, y, z)$ is equivalent to finding a value t_0, around which the most $t_i'(x, y, z)$ values are concentrated in a radius of τ. This t_0 is an estimate of the true shot time, while the number of $t_i'(x, y, z)$ values in its vicinity is the value of the consistency function. Since the total number of values t_i' is N, the calculation of the 3-D

consistency function is cheap at any point (x, y, z) of space (e.g., can be performed in N steps on a sorted list of t_i).

Figure 5 illustrates the construction for six sensors with one outlier (t_6). At the true position the five non-outlier shot-time estimates cluster within 2τ, giving consistency 5; at an arbitrary point the values split into smaller groups, giving lower consistency.

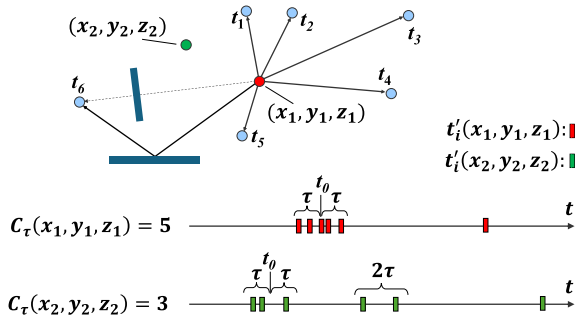


FIGURE 5. Calculation of the consistency function in two points. Measurement t_6 is an outlier. At the true position (x_1, y_1, z_1) there is a strong consistent group of five t_i' values, while in another point (x_2, y_2, z_2) the maximum group size is three.

Reducing the search to a 3-D space removes the time grid but still requires evaluation over the spatial volume of interest (typically $\sim$ 100 $\times$ 100 $\times$ 10 m at $\sim$0.1 m resolution), which is impractical to brute-force. The consistency function admits an efficient best-first branch-and-bound search via an upper-bound inequality on cubic cells [17]: for a cube of side L, $\max_{(x,y,z)\in\text{cube}} C_\tau(x, y, z) \le C_{\tau'}(x_c, y_c, z_c)$ with $\tau' = \tau + L\sqrt{3}/c$. The search starts from a coarse partition of the search space, evaluates the upper bound on each cell, subdivides the highest-bound cell into octants, and continues until the highest-bound cell is at the target resolution Δ. The same structure is reused in the multi-channel pipeline of Section IV for richer evidence sets. The fast method's computational cost is roughly 8 orders of magnitude below a linear grid search [14].

When multiple shots are present, it may be necessary to determine multiple local maxima of the consistency function. In such cases, first the global maximum is found and then the measurements contributing to it are removed from the measurement set. Subsequently, a new consistency function is defined using the remaining measurements and the algorithm is repeated while the consistency function shows significant peaks. Notice that subsequent searches are faster because the number of measurements is reduced. Echoes and true shots can be distinguished as follows: The estimated time of shot t_0 is stored for each local maximum. If the time values for two local maxima are closer than a few times τ, then the smaller peak probably corresponds to an echo.

IV. SPARSE MULTI-CHANNEL SENSORS

The dense single-channel paradigm requires a pre-deployed network, which is impractical in combat scenarios where sensor recovery or redeployment is infeasible. A soldier-wearable variant moves with the warfighter, but the available sensor density is set by personnel distribution rather than deliberate placement — typically an order of magnitude lower than a static network. To preserve localization accuracy with fewer nodes, each must extract richer per-event information; this motivates equipping each sensor with a four-microphone array that yields angle-of-arrival (AoA) in addition to time-of-arrival. The wearable paradigm also introduces a hardware requirement absent from fixed deployments: each node integrates an IMU and digital compass to transform local AoA into the global coordinate frame required for multi-sensor fusion.

A. MEASUREMENT FUSION OVERVIEW

Each of the sensors detects all or a subset of the events related to the gunshot: shockwave ToA and AoA, muzzle blast ToA and AoA, the TDoA between muzzle blast and shockwave and the shockwave duration.

The measurement fusion algorithm processes these detections from the distributed sensor network to estimate the shooter's position, projectile trajectory, bullet caliber, and weapon type through a sequential multi-stage computation:

- When a single sensor detects both shockwave and muzzle blast AoAs, estimate the shooter range using the combined measurements and the TDoA between the two acoustic events.
- Fuse shockwave AoAs from multiple distributed sensors across the network to compute the projectile trajectory.
- Estimate the shooter position using an extended version of the consistency function-based approach described in Section III. If the trajectory could not be computed in the previous step, use the estimated shooter position with available shockwave data to compute the trajectory.
- If the trajectory has been successfully estimated, determine the bullet caliber by analyzing shockwave duration from sensors at known miss distances from the computed trajectory.
- If caliber has been determined, classify the weapon type by comparing the estimated projectile velocity profile against known ballistic signatures of different weapons.

B. RANGE COMPUTATION

The range to the shooter can be estimated using the muzzle blast and shockwave ToAs and AoAs as was shown in Section II-A. [4] showed that this also works for a pair of sensors: if one measured the AoA of the shockwave only and the other measured the muzzle blast AoA, the range can still be computed using a simple formula. This step of the measurement fusion proceeds by computing all possible range estimates using this method.

C. TRAJECTORY ESTIMATION

As was mentioned in Section II-B, there is a closed form formula to compute the trajectory from two AoA

estimates [9]. However, if the angles are too similar, the error can be large. For example, a common situation arises when all the sensors and the trajectory are roughly in the same plane, and the sensors are all on the same side of the trajectory. In this case, the trajectory cannot be computed accurately. Further complicating the situation is the fact that many times only three microphones detect the shockwave resulting in two AoA candidates symmetrical to the plane of the microphones.

The trajectory estimation proceeds as follows. We gather all pairs of AoA detections where the two angles differ by more than an empirically derived threshold. Applying Danicki's method to these pairs generates numerous trajectory candidates, many of which are spurious solutions arising from the geometric ambiguity inherent in three-microphone detections. To identify the true trajectory from this candidate set, we employ a clustering-based outlier rejection approach followed by averaging of the inlier trajectories.

The filtering algorithm operates as follows. For any pair of trajectory candidates, we quantify their similarity using the scalar product of their direction vectors:

$$D(i, j) = \mathbf{v}_i \cdot \mathbf{v}_j$$

where $\mathbf{v}_i$ and $\mathbf{v}_j$ are unit vectors parallel to trajectories i and j, respectively. We then construct a neighborhood for each candidate trajectory:

$$N(i) := \{j \mid D(i, j) < R\}$$

where R is a predefined similarity threshold. The neighborhood with maximum cardinality defines the core set C, which we assume contains the correct trajectory along with other candidates consistent with it; all trajectories outside this core set are classified as outliers. The final trajectory estimate is obtained by averaging the direction vectors of all candidates in the core set C.

D. SHOOTER POSITION ESTIMATION

Several measurement types constrain shooter location:

- Range estimates from individual sensors detecting both shockwave and muzzle AoAs, or from two arrays detecting shockwave and muzzle blast AoAs, respectively,
- Muzzle blast AoAs providing bearing estimates, and
- Muzzle blast ToAs (four sufficient for multilateration).

While these measurements have different error characteristics, all are susceptible to echoes. Therefore, standard nonlinear optimization techniques (e.g., Newton-Raphson) could fail given the potentially large number of outliers.

We extend the consistency function from Section III to incorporate range estimates and muzzle blast AoAs in addition to ToAs, and search for the shooter location by a best-first branch-and-bound procedure over a spatial octree.

For each candidate 3-D box B, the consistency $C(B)$ is defined as the weighted count of measurements admitting *some* shooter position inside B. Each evidence class contributes with its own (empirically tuned) weight: an MB ToA contributes if a shot time can be chosen such that some $S \in B$ produces that ToA within the detection tolerance; an MB AoA contributes if the bearing ray from its sensor intersects B; a per-array shooter estimate contributes if the estimate lies inside B. These are exact, lightweight per-box predicates; no actual shooter position is computed during scoring.

The search begins with a single box covering the area of interest and maintains a priority queue of unexplored boxes ordered by $C(B)$. At each step the highest-scoring box is dequeued, subdivided into eight equal sub-boxes (octree split), and the children are scored and inserted back into the queue. The procedure terminates when the box at the front of the queue has reached the target spatial resolution; that box is returned as the shooter estimate. Because C is not strictly monotone in box size — a parent box that includes many feasible shooter positions may have children that include only a few — a high-scoring parent can have lower-scoring children, in which case other previously-enqueued boxes resurface to the front of the queue. This is the implicit backtracking that lets the algorithm escape local maxima of C without ever requiring an explicit search over the full grid.

To evaluate consistency for candidate box B: Each muzzle blast ToA maps to an interval of feasible shot times (since B is a region, not a point). We count the maximum overlap among these intervals, following Section III. The consistency score increases when range estimates fall within B, and muzzle blast AoAs contribute proportionally to the number of rays intersecting B. Empirically determined weights balance the contributions of these heterogeneous measurement types.

If a trajectory estimate exists, shooter location is constrained to a line, enabling the algorithm to search along the trajectory rather than the entire area—substantially accelerating computation.

Conversely, if shooter position is known but trajectory is not, we define a consistency function over projectile velocity vector $\mathbf{v}$ as the count of shockwave AoAs consistent with the velocity (which determines the Mach angle). Maximizing this function yields the trajectory estimate.

E. CALIBER ESTIMATION

Once the trajectory is known, the only unknowns remaining in Whitham's formula (Equation 1) are the projectile diameter and length, since we can calculate both (1) the miss distance from the sensor location and (2) the Mach number from the shockwave angle of arrival. In practice, only a few discrete projectile types are encountered in operational scenarios. Table 1 lists the six weapons and four distinct projectile types used in our experiments. Note that while the AK-47 and M240 share the same 7.62 mm caliber, their projectile lengths differ significantly due to the different cartridges (7.62×39 mm versus 7.62×51 mm NATO).

Figure 6 plots the measured shockwave duration as a function of estimated miss distance, with different markers corresponding to various weapons. The miss distances shown are estimates provided by the measurement fusion algorithm rather than ground truth measurements. For these experiments, we deployed ten stationary sensors, each equipped

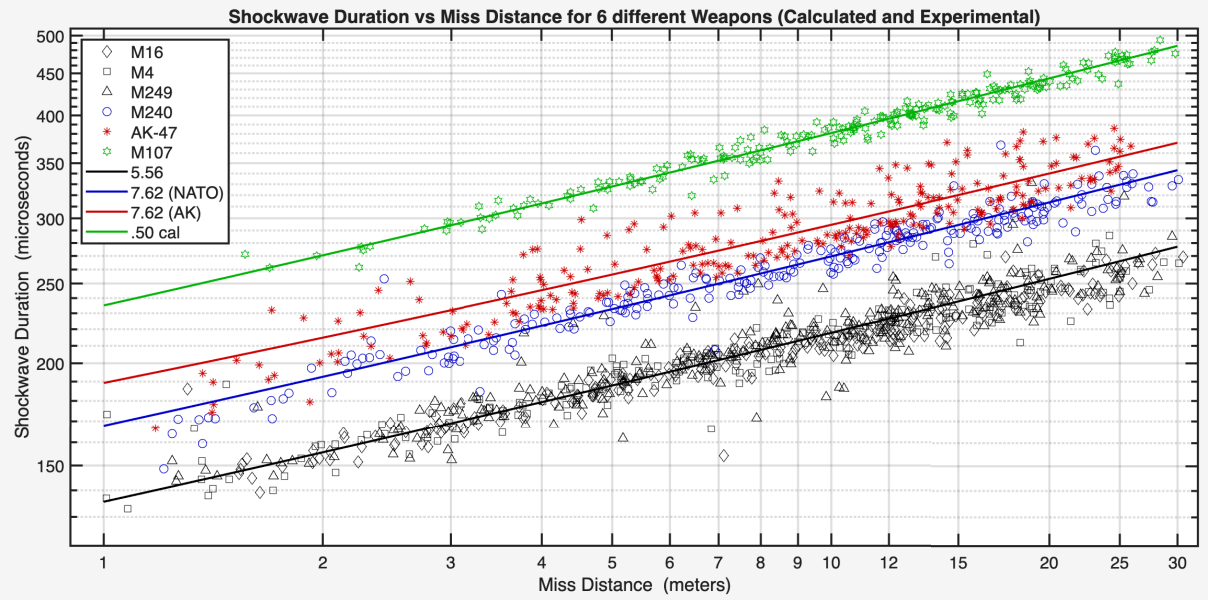


FIGURE 6. Shockwave duration as a function of miss distance. Each measured shot is represented by multiple markers corresponding to sensors detecting it. Lines show Whitham's theoretical formula (Equation 1) evaluated at each projectile's nominal Mach number with an empirically fitted offset per projectile type. The vertical scatter of points around each curve therefore reflects both measurement noise and the genuine variation in T from Mach number changing across shooter ranges (T depends on both miss distance and projectile speed).

TABLE 1. Weapons and ammunition types used.

Weapon	Caliber	Cartridge	Projectile Length	Muzzle Velocity
M16	5.56 mm	5.56×45 mm	23 mm	Mach 2.8
M4	5.56 mm	5.56×45 mm	23 mm	Mach 2.5
M249	5.56 mm	5.56×45 mm	23 mm	Mach 2.6
AK-47	7.62 mm	7.62×39 mm	32 mm	Mach 2.1
M240	7.62 mm	7.62×51 mm	27 mm	Mach 2.5
M107	.50 cal	12.7×99 mm	58 mm	Mach 2.6

with four microphones. We filtered the data to include only measurements where at least three of the four microphones at each node agreed on the shockwave duration to within 5 microseconds, ensuring measurement quality.

The solid lines in Figure 6 represent Equation 1 with empirically determined per-projectile offsets (+30 μs for 5.56 mm, +37 μs for 7.62 × 51 mm, +53 μs for 7.62 × 39 mm, +48 μs for.50 cal) fitted by least squares. The most likely physical origin of the caliber-dependent residual is the bullet's base-wake structure, which Whitham's slender-body derivation does not capture and which varies systematically with projectile base geometry (flat-based vs. boattailed) [18], [19]. Also, controlled measurements of Stoughton [6] found that Whitham's formula overestimates duration for 5.56 mm and underestimates it for 7.62 mm.

For caliber estimation in practice, we use the estimated shockwave duration and miss distance pairs from multiple sensors. Each shot was detected by an average of 5 to 6 sensors, providing sufficient redundancy to reject at least one outlier measurement. We compute the root mean square error between each sensor's measurement of a given shot and each theoretical curve (corresponding to different projectiles) and select the one that provides the best match.

F. WEAPON CLASSIFICATION

Within a caliber class, weapons can be further discriminated by muzzle velocity, which is approximately constant for a given weapon-ammunition combination and recoverable from the per-sensor projectile-velocity estimates extrapolated back to the muzzle via the deceleration model. For each candidate weapon, the predicted per-sensor velocities (using its known muzzle velocity and deceleration function at the estimated shooter ranges) are compared against the measured values, and the minimum-RMSE weapon is selected. Discrimination is reliable when the within-caliber muzzle velocities are well-separated (AK-47 Mach 2.1 vs. M240 Mach 2.5; M16 vs. M4/M249) but limited when they are not (M4 vs. M249 within the 5.56 mm NATO class).

V. SPARSE SINGLE-CHANNEL SENSORS

The multi-channel mobile system achieved soldier-wearable protection but introduced significant complexity through orientation tracking requirements. Digital compasses and inertial measurement units proved vulnerable to interference from metallic objects and electromagnetic sources—precisely the environments where soldiers operate. This section demonstrates that comparable localization accuracy can be achieved with simpler single-channel sensors.

The key innovation elevates shockwave duration from secondary characteristic to primary geometric constraint [5]. The Whitham equation (Equation 1) relating shockwave duration to miss distance can be rearranged to express miss

distance as a function of measured shockwave duration and an unknown weapon-specific scaling factor. By recognizing that this scaling factor remains constant across all sensors detecting the same projectile, the consistency function search can jointly optimize trajectory parameters and the scaling factor without requiring prior weapon knowledge.

By simplifying Equation 1

$$T_{shock} = \frac{1.82\, Mb^{1/4}}{c(M^2-1)^{3/8}} \frac{d}{l^{1/4}} \approx \frac{1.82\, d}{c} \left(\frac{Mb}{l}\right)^{1/4} \quad (6)$$

then rearranging it and grouping the weapon-specific and speed-dependent terms into a scaling factor k, we get

$$b \approx kT_{shock}^4 \quad (7)$$

where k encapsulates the projectile characteristics and ballistic properties and b is the miss distance. This rearrangement is crucial because it transforms the shockwave duration measurement—which can be reliably obtained from a single microphone—into a direct geometric constraint: each sensor's measured shockwave duration T_{shock} defines a circle of radius kT_{shock}^4 centered at that sensor's location, and the projectile trajectory must be tangent to this circle. The beauty of this formulation is that while the scaling factor k remains unknown initially, it has the same value for all sensors since they are measuring the shockwave duration of the same projectile. Hence, it can be jointly optimized along with the trajectory parameters during the search process. This effectively converts what was previously a classification problem into a geometric localization constraint that can be fused with time-of-arrival measurements, enabling trajectory estimation without requiring prior knowledge of the weapon type nor orientation tracking.

A. MEASUREMENT FUSION OVERVIEW

The sensor fusion algorithm operates on three types of acoustic measurements extracted from distributed single-channel microphone nodes:

- shockwave time-of-arrival (ToA),
- muzzle blast time-of-arrival (ToA), and
- shockwave duration.

Note that not every sensor necessarily detects all three measurement types for a given gunshot. Sensors may observe only the shockwave, only the muzzle blast or none of them depending on their position relative to the shooter and trajectory.

The sensor fusion algorithm executes four sequential steps [5]:

- Initial Shooter Localization (Optional). Multilateration is applied using muzzle blast ToA measurements to obtain an approximate shooter position. This step reduces the search space for subsequent trajectory estimation, but is skipped when insufficient muzzle blast detections are available or too many outliers are present.
- Projectile Trajectory Estimation. A search procedure minimizes a composite error function over the space of possible trajectories. The error function combines least-squares fitting of miss distances (computed from shockwave durations via the rearranged Whitham equation $b \approx kT^4$) with variance minimization of shockwave angles (derived from pairwise shockwave TDoAs). Both components incorporate outlier rejection mechanisms. (Section II-B)
- Final Shooter Localization. The trajectory estimation provides both projectile speed and miss distances for each sensor. Individual sensor-to-shooter ranges are computed using the TDoA-based approach without requiring weapon-specific assumptions. Constrained trilateration along the estimated trajectory yields the refined shooter position through overdetermined system optimization. (Section V-C)
- Caliber Estimation and Weapon Classification. With known projectile speed, the scaling factor k from the trajectory estimation constrains the remaining unknowns in the Whitham equation (caliber and projectile length) to a discrete set of possibilities. Combined with the estimated shooter range and projectile speed, these parameters enable weapon classification through muzzle velocity matching. (Section V-D)

If at least three muzzle blasts ToAs are available, the initial shooter localization can be done using traditional multilateration or the original consistency function-based method described in Section III. The accuracy of the result is not crucial; it is only used to constrain the trajectory search space.

B. TRAJECTORY SEARCH

We need to find the projectile trajectory that is maximally consistent with the observations. On the one hand, we have the shockwave duration measurements that constrain the miss distances. On the other hand, we have pairwise shockwave TDoAs that constrain the Mach angle. We need to jointly optimize the scaling factor k in Equation 7 and the Mach angle to minimize the error.

The first error component is the least-squares discrepancy between the actual miss distance from each sensor to the trajectory L and the miss distance predicted by Equation 7 for a candidate scaling factor k:

$$E_1(k, L) = \sum_{i=1}^{n} \left(b_i(L) - k\, T_i^4\right)^2, \quad (8)$$

where $b_i(L)$ is the miss distance from sensor M_i to trajectory L, T_i is the shockwave duration measured at M_i, and n is the number of sensors with valid shockwave detections.

The second component is based on the shockwave TDoAs. Consider Figure 7. Two microphones M_1 and M_2 positioned at miss distances b_1 and b_2 from the projectile trajectory detect the shockwave at times t_{sw1} and t_{sw2}, respectively. Without loss of generality, assume the time difference of

arrival $\Delta t = t_{sw2} - t_{sw1}$ is positive (the microphone labels can be exchanged if necessary to ensure this condition). At the instant t_{sw1} when the shockwave reaches microphone M_1, the projectile occupies position B_1 along its trajectory. Similarly, when the shockwave arrives at M_2 at time t_{sw2}, the projectile has advanced to position B_2. Given the projectile velocity v, the distance $|B_1B_2|$ traveled during the time interval Δt establishes a geometric constraint.

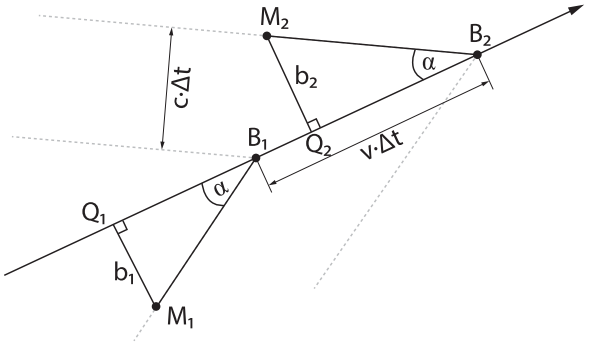


FIGURE 7. Shockwave angle calculation using TDoA.

Let Q_1 and Q_2 represent the orthogonal projections of the microphone positions onto the trajectory line. The geometric relationship can be expressed as:

$$|Q_1B_1| + v\Delta t = |Q_1Q_2| + |Q_2B_2| \tag{9}$$

Crucially, the same equation holds whether M_1 and M_2 are located on the same or opposite sides of the trajectory. Δt is measured while $|Q_1Q_2|$ is easily computed from the microphone positions and the given trajectory candidate. Finally, the distances $|Q_1B_1|$ and $|Q_2B_2|$ can be related to the shockwave cone angle α and the respective miss distances:

$$|Q_1B_1| = \frac{b_1}{\tan\alpha} \qquad |Q_2B_2| = \frac{b_2}{\tan\alpha} \tag{10}$$

Since the projectile velocity relates to the shockwave angle through the Mach cone geometry, substituting these relationships into Equation 9 yields a quadratic equation in $1/\tan\alpha$ with a unique physical solution:

$$\frac{b_1}{\tan\alpha} + c\sqrt{1 + \frac{1}{\tan^2\alpha}}\,\Delta t = |Q_1Q_2| + \frac{b_2}{\tan\alpha} \tag{11}$$

where c is the speed of sound. Applying Equation 11 to every microphone pair (M_i, M_j) yields a set of pairwise Mach-angle estimates $\{\alpha_{i,j}\}_{i<j}$ that should all agree in the noiseless case. We define the second error component as the variance of this set:

$$E_2(L) \;=\; \mathrm{Var}\big(\{\alpha_{i,j}\}_{i<j}\big). \tag{12}$$

The composite error function combines the two components multiplicatively:

$$E(k, L) \;=\; E_1(k, L) \cdot E_2(L). \tag{13}$$

The trajectory search must find a trajectory angle and one point on the trajectory. To define a set of points to choose from, we can exploit the geometric constraint that feasible trajectories must pass within the maximum shockwave detection range of the sensors that detected the shot. Hence, circles of radius b_{max} centered at each sensor position define the detection envelope. Any trajectory intersecting these circles must also intersect a larger circle centered at the centroid of the active sensor positions large enough to encompass all the smaller circles. This observation enables parameterization of all geometrically feasible trajectories using two variables: an angle $\phi \in [0, 2\pi]$ defining a pivot point on the encompassing circle's circumference, and the trajectory slope angle. In Figure 8, ϕ is the angular coordinate of each pivot P_j around the encompassing circle, and the slope angle is the bearing of the trajectory line drawn through that pivot; the figure illustrates one such candidate (the vertical trajectory through P_1, which happens to be tangent to every detection-range circle).

The search algorithm proceeds as follows. Without an initial shooter estimate, pivot points are distributed uniformly around one half of the encompassing circle's circumference (we do not consider the diametrically opposite of a pivot point). At each pivot, the trajectory slope is varied in discrete increments to define a set of candidate trajectories. For each candidate trajectory L, the composite error $E(k, L)$ from Equation 13 is evaluated over a discrete sweep of scaling-factor values k, and the minimum-error (k, L) pair at that pivot is recorded. The (k, L) pair yielding the overall minimum error across all pivots and slope angles is selected, with the projectile velocity determined as a byproduct of the optimization.

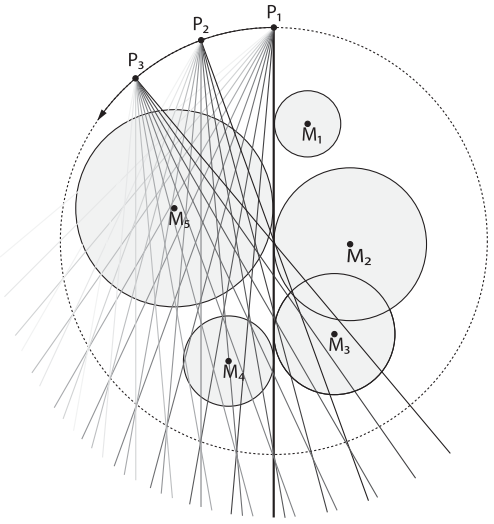


FIGURE 8. Trajectory search procedure. Notice how the vertical trajectory from P_1 is a tangent of every circle.

When an approximate shooter position estimate is available from the initial multilateration step (Section V-A), the search simplifies considerably. The trajectory is constrained to pass through the estimated shooter location, reducing the optimization to a single parameter—the slope angle. The

angular search range is further constrained to trajectories that pass within the maximum detection range b_{max} of the sensors. The accuracy of the initial estimate is not critical: the final shooter position is recomputed by constrained trilateration along the resulting trajectory in Section V-C, which can shift the localization along the trajectory direction away from the initial estimate without sacrificing accuracy.

C. SHOOTER POSITION

At this point, we have a good trajectory estimate and consequently, we know all miss distances. We also know the Mach angle estimate, and hence, the speed of the projectile. Recall Section II-C: the miss distance, the projectile speed and the TDoA between the muzzle blast and the shockwave are sufficient to calculate the range to the shooter with a closed-form formula. At this point, these are all known without assuming a known weapon type.

Therefore, we use the range estimates to carry out constrained trilateration along the trajectory, which does not suffer from the radial inaccuracy of the multilateration applied in the first step. A modified version of the consistency function can be used in this step, albeit with a much smaller, one-dimensional search space: slide a window along the trajectory and the value of the consistency function is the number of circles that intersect the trajectory within the window. The shooter position is taken at the window center with the maximum consistency value, provided this value is at least a minimum-inlier threshold (we use 5 in this work, matching the configuration of the original Aberdeen evaluation [5]); shots that never reach this threshold are reported as unlocalized.

Multilateration and trilateration handle clustered-sensor geometry differently when the shooter is distant: hyperbolic intersections from MB TDoAs produce uncertainty aligned *radially* with the shooter direction (Figure 9a), while circular range constraints produce uncertainty aligned *perpendicular* to it (Figure 9b). The hybrid approach exploits this complementarity: multilateration provides only a coarse initial estimate, and the final shooter is recovered by trilateration constrained to the estimated trajectory, converting the radial error pattern into the more favorable perpendicular one (Figure 9c).

D. CALIBER ESTIMATION AND WEAPON CLASSIFICATION

As a "side effect" of the trajectory estimation algorithm, the scaling factor k has been computed that tells us how the miss distances are related to the fourth power of the shockwave durations. The value of k carries information about the projectile, and, together with the bullet speed, allows for projectile classification (caliber and length).

Rearranging Equation 7, k can be expanded as

$$k = \frac{lc^5}{1.82^4 d^4 v} \tag{14}$$

where l is the length and d is the caliber of the bullet, and c and v are the sound and bullet speeds, respectively. Here, the projectile specific quantity is d^4/l, which we call the bullet coefficient C_{bullet}, computed as follows:

$$C_{bullet} = \frac{c^5}{1.82^4 vk} \tag{15}$$

At this point, all information is available to compute the bullet coefficient, which makes it feasible to differentiate between possible projectile types and provide a caliber estimate.

Weapon classification proceeds as described in Section IV-F. From the range to the shooter and the projectile speed over the sensors, the muzzle velocity can be estimated using the deceleration function of the given bullet type. The muzzle velocity and the bullet coefficient are sufficient to classify the rifle in most cases.

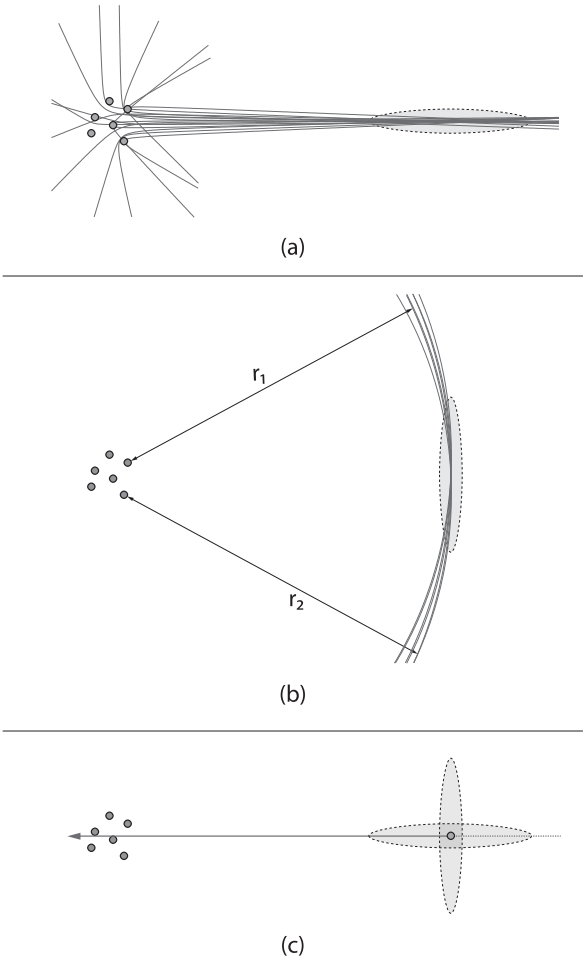


FIGURE 9. The nature of multilateration (a) and trilateration (b) errors and the result of their combined application (c).

E. EVIDENCE-CLASS COMPARISON WITH THE MULTI-CHANNEL ARCHITECTURE

Having described both sparse architectures, it is useful to compare them by the evidence classes each pipeline fuses. Table 2 lays out which measurement-derived quantities each architecture uses and at what stage of localization.

The asymmetry is structural: the multi-channel pipeline derives four independent evidence classes per array (per-array shooter, per-array range, MB AoA bearing, MB ToA), all of which feed simultaneously into a 3-D consistency search; the single-channel pipeline derives a smaller set of per-sensor scalars (one MB ToA, one SW ToA, one SW duration), with the final shooter location resting on a tight cluster of per-sensor ranges along the trajectory.

TABLE 2. Evidence classes used by the two sparse architectures. The multi-channel pipeline's 3-D consistency search fuses four classes simultaneously (per-array shooter, per-array range, MB AoA bearing, MB ToA cluster); the single-channel pipeline's localization rests on a smaller per-sensor scalar set, with the final shooter cluster requiring at least five mutually-consistent MB–SW-TDoA-derived ranges. This redundancy difference is the principal driver of the multi-channel pipeline's robustness advantage in cluttered acoustic environments (Section VIII-E).

Evidence class	MC sparse	SC sparse
Per-array shooter (single-array closed form)	yes	—
Per-array range (single-array closed form)	yes	—
MB AoA bearing	yes	—
MB ToA cluster (network multilateration)	yes	initial only
SW AoA pair → trajectory direction	yes	—
SW pairwise TDoA → Mach angle	—	yes
SW duration → miss distance ($b \approx kT^4$)	caliber only	primary trajectory
MB–SW TDoA → per-sensor range	via per-array	yes (final pos.)

The practical consequence becomes visible under acoustic multipath: a corrupted MB arrival on one sensor biases the SC pipeline's per-sensor range and may break the 5-inlier cluster; in the MC pipeline the same corruption affects one of four parallel evidence classes for that sensor, with the consistency search excluding the sensor from the consensus rather than failing it. Section VIII-E reports the empirical multipath robustness gap.

VI. ABERDEEN 2006 FIELD EVALUATION–BASELINE

This section reproduces the localization performance of the two sparse-architecture pipelines as originally published [4], [5], evaluated on the Aberdeen 2006 field test data. The numbers reported here serve as the baseline against which the algorithmic refinements introduced in Section VII are evaluated in Section VIII. Both pipelines are run with the parameters and conventions of the original publications; no refinements proposed in this paper are enabled.

A. FIELD TEST SETUP

A third-party evaluation of the sparse MC architecture was conducted by a team from NIST at the US Army Aberdeen Test Center [20]. Ten sensor nodes were deployed covering an approximately 30 × 30 m area, with positions surveyed and array orientations carefully calibrated. Five fixed targets were positioned behind the sensor network, with firing positions established at ranges of 50, 100, 200, and 300 m from the edge of the sensor field. Six weapon types were evaluated – AK-47 (7.62 × 39 mm M43), M240 (7.62 × 51 mm NATO M80), M16, M4, and M249 (5.56 × 45 mm M193/M855), and M107 (.50 cal M33) – producing 196 single-shot events. Figure 10 shows the deployment area.

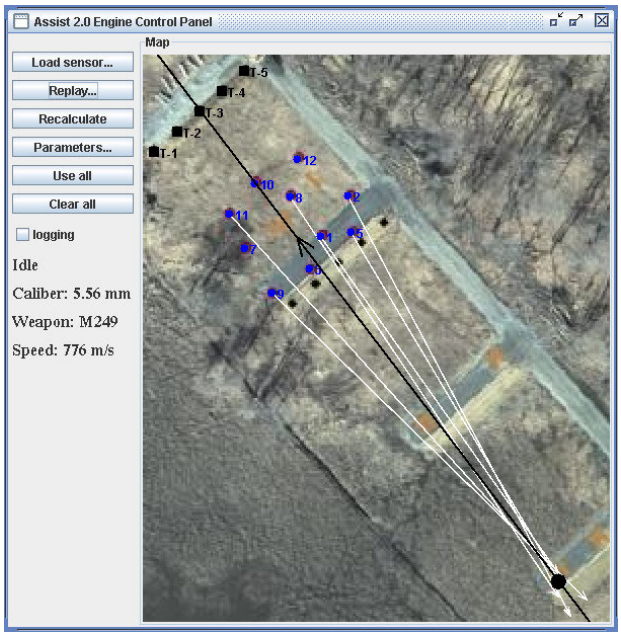


FIGURE 10. Overhead view of the Aberdeen Test Center deployment. Blue dots mark sensor locations covering a 30 × 30 m area. White arrows show per-array shooter estimates from individual sensors. The black line and circle mark the network-wide trajectory and shooter location estimates, both of which had essentially zero error in this case. The shooter was approximately 100 m from the nearest sensor.

The sensor hardware (described in Sections IV–V) used low-cost electret microphones sampled at 1 MSPS, with all signal processing performed on a Xilinx XC3S1000 FPGA. Time synchronization across the network used the Routing Integrated Time Synchronization (RIPS) protocol [21], achieving better than 50 μ s network-wide accuracy. All 196 shots were detected by the network; the shockwave detector produced no false detections, while the muzzle-blast detector occasionally triggered on artillery fire from a nearby range or on shockwave echoes. Because the sensor positions were surveyed and orientations calibrated, the principal noise sources affecting localization are measurement-side (per-sensor ToA estimation, AoA estimation, and SW duration estimation) rather than position or orientation uncertainty.

B. SPARSE MULTI-CHANNEL PIPELINE

The multi-channel pipeline (Section IV) successfully localized 189 of 196 shots (96% rate). The seven unlocalized shots shared a common geometric pattern: trajectories passing wide of the sensor field produced near-parallel shockwave AoAs across all sensors, leaving the trajectory-search step with no resolvable angular baseline. Table 3 summarizes performance by shooter range.

Two patterns visible in Table 3 motivate the refinements of Section VII. First, the trajectory azimuth and the perpendicular distance to the shooter remain tight (under 1.5° and under 7 m respectively) across all ranges, indicating

TABLE 3. Sparse multi-channel pipeline – baseline localization performance on Aberdeen 2006 (196 shots). Loc. Rate: shots producing a localization estimate; Cal. Acc.: caliber accuracy; Az. Err.: trajectory azimuth error; Perp. Err.: perpendicular distance from the estimated trajectory to the true shooter; Dist. Err.: total 3D distance error. Numbers reproduce [4].

Range (m)	Loc. Rate	Cal. Acc.	Az. Err. (deg)	Perp. Err. (m)	Dist. Err. (m)
50	93%	100%	0.86	0.91	2.2
100	100%	100%	0.66	1.34	8.7
200	96%	100%	0.74	2.71	32.8
300	97%	97%	1.49	6.29	70.6
All	96%	99.5%	0.88	2.47	23.0

that the trajectory *line* is well recovered. Second, total 3D distance error grows from 2.2 m at 50 m to 70.6 m at 300 m – approximately quadratic in range. The bullet's true trajectory is recovered, but the muzzle along that line is biased downrange. We address this in Section VII-A; the seven dropped shots are addressed in Section VII-B.

C. SPARSE SINGLE-CHANNEL PIPELINE

The single-channel pipeline (Section V) was evaluated on the same Aberdeen dataset. Because the muzzle-blast detection rate degrades sharply beyond 200 m, the original analysis [5] restricted attention to the 108 shots with shooter-to-sensor range $\leq$ 130 m, covering all six weapon types. To simulate single-channel operation on the multi-channel hardware, one microphone was selected from each of the 10 nodes; the trajectory error function was configured to require at least five consistent sensor observations.

The pipeline achieved a 99% trajectory success rate (107 of 108 shots) with 0.80° mean absolute azimuth error and a near-zero systematic bias (−0.1°). Of the 107 successful trajectories, 104 yielded shooter positions (96% overall localization rate) with mean position error 2.96 m. On the same 108-shot subset, the multi-channel pipeline produced 5.45 m mean position error. Section VIII revisits this head-to-head, separating the contributions of architecture, deceleration bias, and multipath.

D. CALIBER AND WEAPON CLASSIFICATION

Both pipelines achieved very high caliber accuracy: 99.5% for the multi-channel pipeline (188 of 189 localized shots) and 100% for the single-channel pipeline (104 of 104), confirming that shockwave-duration-based caliber estimation is robust under field conditions despite the empirical per-weapon offsets discussed in Section IV-E. Weapon discrimination within a caliber class was reliable for AK-47, M240, M16, and M107 but limited for the close-velocity 5.56-mm pair M4/M249. Detailed classification numbers are reported in [4] and [5].

E. SUMMARY

The published baselines establish that consistency-function fusion produces reliable trajectory estimates and high caliber accuracy across both sparse architectures, but reveal two specific shortcomings that have not previously been resolved: a quadratic-with-range bias in the estimated muzzle position, and complete failure on the small fraction of shots whose trajectory passes wide of the sensor field. Section VII introduces algorithmic refinements addressing each of these shortcomings; Section VIII then quantifies their impact on both architectures using a simulator calibrated against the empirical noise budget of the Aberdeen field measurements.

VII. ALGORITHMIC REFINEMENTS

We present two algorithmic refinements that address known limitations of the consistency-function pipelines as originally published. The first – a caliber-aware deceleration correction (Section VII-A) – applies to both sparse architectures and removes the systematic position bias that grows quadratically with shooter range, visible in Table 3. The second – a coplanar-geometry recovery (Section VII-B) – is specific to the sparse multi-channel pipeline and recovers the small fraction of shots whose trajectory passes wide of the sensor field, a failure mode acknowledged in the original Aberdeen evaluation [4]. Both refinements are gated behind explicit configuration flags; the baseline numbers reported in Section VI have neither enabled, while the simulator-based comparisons of Section VIII use both.

A. CALIBER-AWARE DECELERATION CORRECTION

For sensor a, the *emission point* P_a is the point along the bullet's trajectory from which the shockwave that ultimately reaches that sensor originates — i.e., the apex of the local Mach cone whose surface passes through sensor a. The *emission time* is the instant at which the bullet passes through P_a; equivalently, the instant at which the shockwave reaching sensor a begins propagating outward from the trajectory.

Both the closed-form range from MB-SW TDoA used in Section II-C and the per-array shooter formula used in Section IV treat the projectile as traveling at constant velocity from the muzzle to the emission point. Real bullets decelerate. Modeling the deceleration as linear, $v(s) = v_0 + \kappa s$ with $\kappa < 0$, the bullet reaches each sensor's emission point P_a at distance s_P from the muzzle with velocity $v_P = v_0 + \kappa s_P$, but its *mean* velocity over the segment $[0, s_P]$ is

$$\bar{v} = v_P + \tfrac{1}{2}|\kappa| s_P, \tag{16}$$

which is strictly larger than v_P. Consequently, the bullet arrives at P_a *earlier* than a constant-v_P bullet would: the difference between the actual emission time and the time a constant-v_P bullet would reach the same point is

$$\Delta t_{sw} = \frac{s_P(\bar{v} - v_P)}{v_P \bar{v}} = \frac{|\kappa| s_P^2}{2 v_P \bar{v}}. \tag{17}$$

Adding Δt_{sw} to the observed shockwave time-of-arrival on each sensor yields the time a constant-v_P bullet *would* have produced, so the original constant-velocity range and per-array shooter formulas then return the correct geometry without further modification.

The correction depends on per-sensor quantities (s_P, v_P) that are not directly measured. We obtain them from the prior pipeline output: the trajectory direction and shooter location from Sections IV-C–V-C fix s_P via the same Mach-cone geometry that produces it during forward simulation, and v_P comes from the bullet-velocity model evaluated at s_P. The deceleration coefficient κ is set to the caliber-class average from the weapon database; a per-weapon classification is not required because Equation 17 depends on κ only weakly through $\bar{v}$ in the denominator. This decouples the correction from the small fraction of shots where weapon classification within a caliber class is ambiguous (notably the close-velocity 5.56-mm pair M4/M249, see Section IV-F).

Algorithm

1) Run the baseline pipeline of Section IV or Section V to obtain an initial shooter location $\hat{S}$, trajectory direction $\hat{\mathbf{d}}$, and per-sensor emission distances $\{s_{P,a}\}$ and emission-point velocities $\{v_{P,a}\}$.
2) Run the caliber-classification stage (Section IV-E) to obtain the caliber class, then look up the caliber-mean deceleration coefficient $|\kappa|$.
3) For each sensor with valid shockwave and muzzle-blast detections, replace the observed $t_{sw,a}$ with $t_{sw,a}^{\text{corr}} = t_{sw,a} + \Delta t_{sw,a}$ using Equation 17.
4) Re-run the shooter-position step on the corrected times of arrival.

The correction is identical in form for the multi-channel and single-channel pipelines; only step 4 differs, since the multi-channel pipeline re-runs the constrained 1-D consistency search along the previously estimated trajectory, while the single-channel pipeline re-runs the constrained trilateration with the corrected per-sensor ranges. Steps 3–4 are iterated until the per-pass shooter shift falls below a small tolerance (5 cm in our implementation); convergence typically takes 2–3 iterations at short range and up to ∼5 at the long-range bin where the bullet operates near the Mach excess limit and the per-iteration s_P estimate refines more slowly.

B. COPLANAR-GEOMETRY RECOVERY FOR SPARSE MULTI-CHANNEL

The trajectory-estimation step of the sparse multi-channel pipeline (Section IV-C) recovers the bullet trajectory from the Danicki closed-form solution applied to pairs of shockwave angle-of-arrival measurements. When the trajectory passes through or near the sensor field, every pair contributes a different geometric constraint and the consensus among pair-derived candidates is sharp. When the trajectory passes *wide* of the sensor field, however, the per-sensor SW AoAs become nearly parallel: every Danicki pair sees the same near-degenerate angular geometry, so every pair returns a similar but uninformative candidate, and the pair-angle gate that filters near-degenerate pairs rejects the entire candidate set. The original Aberdeen evaluation [4] dropped 7 of 196 shots for this reason.

The failure is geometric, not informational: although the pairwise constraints have collapsed, the network as a whole still receives valid muzzle-blast AoAs, valid muzzle-blast times of arrival, and valid per-array shooter estimates from sensors that detect both shockwave and muzzle blast. We exploit this by recovering the shooter position from the surviving evidence first, and then recovering the trajectory from a different geometric constraint that does not require non-degenerate AoA pairs.

Algorithm

1) Run a 3-D consistency search over the area of interest using only the per-array shooter position estimates, the muzzle-blast AoA bearings, and the muzzle-blast ToA cluster. This produces a coarse shooter location $\hat{S}_0$ that does not depend on the (degenerate) shockwave AoA pairs.
2) For each sensor with a valid shockwave detection, express the emission point P_a as a function of the unknown trajectory direction $\mathbf{d}$. Given $\hat{S}_0$ and a candidate $\mathbf{d}$, the trajectory line $\{\hat{S}_0 + s\,\mathbf{d} : s \geq 0\}$ is fixed; the per-sensor Mach angle is recovered implicitly from $v_P = c/|\hat{u}_{sw} \cdot \hat{\mathbf{d}}|$ at each candidate $\hat{\mathbf{d}}$; and the emission point $P_a(\mathbf{d})$ is the unique point on the trajectory line whose Mach cone (of that recovered half-angle) contains the sensor's shockwave AoA ray. The Mach angle therefore appears in the construction but is not an external assumption: no a priori bullet velocity or weapon-class knowledge is required.
3) Search over $\mathbf{d}$ to minimize the residual of the constraint that the per-sensor emission points $\{P_a(\mathbf{d})\}$ are collinear along a line through $\hat{S}_0$. This is a one-dimensional search over the trajectory's azimuth (the elevation is fixed by the residual at each azimuth).
4) With the recovered $(\hat{S}_0, \hat{\mathbf{d}})$, refine $\hat{S}$ by sliding a 1-D window along $\hat{\mathbf{d}}$ from $\hat{S}_0$ and selecting the position that maximizes the multi-channel consistency function from Section IV-C. The 1-D search is identical to the one used in the baseline pipeline, except the passed-in trajectory is the recovered $\hat{\mathbf{d}}$.

The collinearity criterion in step 3 is well-conditioned even when the AoAs themselves are nearly parallel: collinearity of $\{P_a\}$ requires only that the per-sensor emission points all lie on a common line, not that the rays meet at a sharp angle. As long as the sensors are not themselves collinear (which they are not on the Aberdeen 30×30 m deployment), enough geometric diversity remains in the per-sensor positions M_a to localize $\hat{\mathbf{d}}$ uniquely.

The recovery is intentionally narrow in scope: it fires only when Stage 2 of Section IV-C returns the specific failure "no consistent Danicki candidates," leaving all other pipeline behavior unchanged. On the Aberdeen 2006 dataset (196 shots) the baseline pipeline localizes 189 shots and drops 7 in the manner described above; enabling the recovery localizes 5 of those 7 additional shots. This matches the recovery rate of the post-hoc fix described in [4] but uses a

principled criterion (collinearity of emission points along a line through the prior $\hat{S}_0$) rather than the ad-hoc heuristics of that fix. The two shots that remain unlocalized fail the collinearity step itself: in those cases the muzzle-blast detections are insufficient to localize $\hat{S}_0$ reliably, so the recovery never reaches step 3.

VIII. SIMULATION-BASED EVALUATION

The Aberdeen 2006 field data (Section VI) is geometrically rich but covers a single environment, a fixed sensor deployment, and a single noise budget. To exercise the refinements of Section VII across a wider design space – in particular to vary noise levels, sensor geometry, and the presence of acoustic multipath – we built a high-fidelity simulator that produces synthetic per-(shot, sensor) detections in the same format as the field data so that both pipelines run on simulator and field input through identical code paths. This section describes the simulator (Section VIII-A), confirms with a noiseless test that both pipelines recover close to ground truth (Section VIII-B), then runs realistic noise scenarios: an Aberdeen-calibrated profile that reproduces the published Aberdeen results (Section VIII-C), a more demanding mobile-deployment profile (Section VIII-D), and the same profile with moderate acoustic multipath added (Section VIII-E).

A. SIMULATOR

The simulator is structured as two separable stages. The first is a deterministic ground-truth generator that produces idealized per-sensor detections from the configured shooter, target, and sensor geometry. The second is a probabilistic noise injector that perturbs those idealized detections to produce realistic measurements. Separating the stages means the physics is verifiable in isolation – the noiseless output can be checked against analytic predictions sensor by sensor (Section VIII-B) – and the same idealized ground truth can feed different noise budgets without re-running the physics.

The ground-truth stage implements the same physical models the localization pipelines invert, with no approximations. For each (shot, sensor) pair it computes the muzzle-blast time of arrival as r_{MB}/c; the shockwave emission point P_a as the iterative solution of the Mach-cone geometry $|s_Q - s_P| = c\,b/\sqrt{v_P^2 - c^2}$ along the trajectory, with bullet velocity following the linear deceleration model $v(s) = v_0 + \kappa s$ used in Section VII-A; the shockwave time of arrival by exact integration of the decelerating speed from muzzle to P_a plus the acoustic flight time from P_a to the sensor; shockwave and muzzle-blast angles of arrival as the unit vectors from P_a and from the shooter to the sensor, expressed in the sensor's local frame using its surveyed orientation; and the shockwave duration via Whitham's full N-wave formula (Equation 1) – not the simplified $b \approx kT^4$ form, which differs by the $(M^2-1)^{3/8}$ factor and biases miss distance by ∼9% at typical AK-47 Mach numbers. Gravitational drop is included as the parabolic offset $-\frac{1}{2}gt_{\mathrm{bullet}}^2$ along the local vertical (sub-meter at the ranges considered). The configured weapon parameters (v_0, κ, d, l) are the midpoints of the published manufacturer specifications, identical to the values used by both localization pipelines. Both v_0 and κ accept optional 1-σ Gaussian per-shot jitter to model lot-to-lot and round-to-round ammunition variability; the localization pipelines see only the nominal values, so non-zero jitter produces a per-shot trajectory-correlated residual that no per-sensor noise injector could reproduce.

Layered on top of the ground-truth stage, the noise injector models the principal degradation modes encountered in fielded deployments: Gaussian noise on sensor location, sensor orientation (heading/pitch/roll), per-node time synchronization, speed of sound (temperature bias plus configurable wind), shockwave and muzzle-blast times of arrival, shockwave duration, and angle of arrival (azimuth/elevation, multi-channel only); distance-dependent detection probability for both shockwave (driven by miss distance) and muzzle blast (driven by shooter-to-sensor range); false detections as a Poisson process per sensor per shot; and acoustic multipath echoes via Poisson reflectors with configurable spatial density and visibility constraints. Each error source is independently configurable and can be disabled. The configuration used for every simulation in this paper is included as a YAML snapshot in the corresponding output directory, so each result is reproducible to the parameter level.

B. NOISELESS GROUND-TRUTH RECOVERY

As a consistency check, we run both pipelines (with and without the refinements of Section VII) on the noiseless simulator output for the Aberdeen 2006 geometry. With no measurement noise, deviations from ground truth come entirely from the constant-velocity assumption baked into the per-array shooter formula and the closed-form range from MB-SW TDoA. Table 4 summarises position-error performance.

TABLE 4. Position error (m) on the noiseless Aberdeen 2006 geometry with refinements (Section VII) disabled vs enabled. Each cell shows the mean absolute error over the range bin with the 95% bootstrap confidence interval beneath in brackets. *N* is the number of shots in the bin. Trajectory azimuth MAE is 0.10° for the multi-channel pipeline and 0.29° for the single-channel pipeline (all-range, both modes); localization rate is 100% across all bins for both pipelines and both modes.

Range bin	Multi-channel		Single-channel		N
	Baseline	Refined	Baseline	Refined	
50 m	1.45 [1.36, 1.54]	0.30 [0.22, 0.39]	1.34 [1.17, 1.54]	0.54 [0.41, 0.72]	54
100 m	5.58 [4.96, 6.25]	0.47 [0.35, 0.61]	4.71 [4.21, 5.28]	0.85 [0.61, 1.13]	54
200 m	23.44 [21.25, 25.75]	1.53 [1.27, 1.84]	24.84 [21.39, 28.56]	3.13 [2.00, 4.43]	54
300 m	47.04 [41.94, 51.97]	2.04 [1.70, 2.39]	63.28 [49.62, 78.25]	6.59 [3.76, 9.99]	34
All	16.56 [14.09, 19.15]	0.99 [0.85, 1.14]	19.49 [15.66, 23.69]	2.39 [1.72, 3.16]	196

Two observations are immediate. First, with refinements off, position error grows quadratically with range, from

~1.4 m at 50 m to ~50–65 m at 300 m, exactly the deceleration-bias signature predicted in Section VI. Second, with refinements on, both pipelines converge to within a few meters of ground truth at every range bin – the all-range mean drops to 0.99 m for multi-channel and 2.39 m for single-channel, with 95% bootstrap confidence intervals uniformly under a few meters wide. The 300-m bin no longer dominates the residual: the iteration converges cleanly even for the AK-47 subset where the bullet has decelerated to Mach 1.29 at the emission point. Trajectory azimuth and localization rate are essentially perfect throughout, including in the baseline (refinements-off) configuration: the trajectory *line* itself is recovered correctly by the original pipeline, and the deceleration correction acts only on the shooter's position *along* that line — it removes the muzzle-position bias without affecting trajectory direction.

C. ABERDEEN-CALIBRATED NOISE REPRODUCES THE PUBLISHED RESULTS

The next step is to add a representative noise budget reflecting the Aberdeen 2006 deployment hardware, then verify that the simulator-driven pipelines reproduce the headline numbers reported in Section VI. Table 5 lists the noise injector settings; values were chosen to reflect the documented hardware (1 MSPS FPGA detector, RIPS time sync, surveyed sensor positions, calibrated orientations) and the per-quantity precision typical of low-cost electret microphones with FPGA-level edge detection.

TABLE 5. Aberdeen-calibrated noise settings used to corrupt the noiseless ground truth. Surveyed positions, calibrated orientations, and the absence of multipath/false detections reflect the documented Aberdeen 2006 hardware and venue conditions.

Noise source	1-σ value
Sensor position (surveyed)	0 m
Sensor orientation (calibrated)	0°
RIPS network time sync	20 µs
Speed of sound (temperature drift)	1.5 m/s
SW time of arrival (per microphone)	0.5 µs
MB time of arrival	0.4 ms
SW duration (relative)	5%
AoA azimuth (per array)	1.5°
AoA elevation (per array)	3.0°
Per-shot v_0 jitter (lot variability)	5 m/s
Per-shot κ jitter (lot variability)	0.05 m/s/m
Multipath echoes (open range)	disabled
False MB / SW detections	disabled

Table 6 confirms that with the calibrated noise budget the simulator-driven multi-channel pipeline reproduces the published Aberdeen position MAE (19.17 m vs published 23.00 m – the published value sits inside the 95% confidence interval [14.84, 23.94]) and trajectory azimuth MAE (0.85° vs 0.88°, well within the confidence interval), with localization rate matching at 97%. The single-channel pipeline reproduces the published position MAE essentially exactly (3.01 m vs 2.96 m) on the same ≤130 m subset; the slightly looser trajectory MAE (1.42° vs 0.80°) reflects that the simulator's 5% relative SW-duration noise propagates more conservatively than the empirical noise distribution of the original FPGA detector.

TABLE 6. Aberdeen-calibrated simulator output compared to the published Aberdeen field-test results. 95% confidence intervals beneath each simulator value (Wilson interval for localization rate, bootstrap for MAE). Multi-channel numbers are over the full 50–300 m range (196 shots). Single-channel numbers are restricted to the ≤130 m subset (108 shots), matching the analysis subset of the original published evaluation.

Metric	Multi-channel		Single-channel	
	Sim	Paper [4]	Sim (≤130 m)	Paper [5]
Localization rate	96.9% [93.5, 98.6]	96.0%	97.2% [92.1, 99.1]	96.3%
Trajectory az MAE	0.85° [0.70, 1.03]	0.88°	1.42° [1.14, 1.72]	0.80°
Position MAE (m)	19.17 [14.84, 23.94]	23.00	3.01 [2.54, 3.52]	2.96

The match required no per-quantity tuning to the field measurements – only the broad hardware-class assumptions in Table 5. Subsequent subsections use this noise model, with adjustments only as stated.

D. MOBILE-DEPLOYMENT NOISE: OPEN-TERRAIN SCENARIO

Aberdeen 2006 used surveyed sensor positions and calibrated array orientations – conditions favorable beyond what most fielded deployments achieve. To assess pipeline performance under representative mobile-deployment noise, we run a synthetic open-terrain scenario: 1000 shots, 12 sensors deployed in a ~25-m radius squad pattern, weapons drawn uniformly from the six tested at Aberdeen, shooter ranges from 50–200 m. Beyond the noise of Table 5, we add per-sensor GPS error (1.5 m horizontal, 3 m vertical), orientation error (5° heading, 3° pitch/roll), and a small false-detection rate (2% per sensor per shot for muzzle blast, 1% for shockwave). The orientation values represent a magnetometer-plus-tilt-corrected IMU on a stationary sensor in a benign outdoor environment; a raw uncalibrated magnetic compass can be substantially worse 10°–30° near ferromagnetic clutter, but with simple declination compensation and software-side tilt correction the chosen values are representative of what an attentive operator can achieve in practice. Sensitivity to larger orientation error is left as future work.

Table 7 shows that under mobile-deployment noise both pipelines lose accuracy roughly uniformly across the engagement range, but the multi-channel pipeline retains a meaningful advantage on both metrics: position MAE 23.04 m vs 26.35 m all-range, and a 17.6 percentage point (pp) localization-rate gap whose 95% confidence interval [13.9, 21.2 pp] is entirely positive – even the worst-case lower bound has a 14 pp MC advantage. The single-channel pipeline's localization rate falls because the shooter-position step requires at least five consistent muzzle-blast inliers, and the higher MB-ToA noise (combined with false detections) degrades the per-sensor cluster faster than it degrades the multi-channel consensus search.

TABLE 7. Mobile-deployment open-terrain scenario, refinements enabled, 1000 shots. Each cell shows the point estimate with its 95% confidence interval beneath in brackets (Wilson interval for localization rate, bootstrap for position MAE). Adding GPS and orientation noise to the Aberdeen-calibrated profile reduces accuracy substantially for both pipelines, but the multi-channel architecture preserves a 17.6 pp [13.9, 21.2] advantage in localization rate.

Range bin	Multi-channel		Single-channel	
	Loc rate	Pos err (m)	Loc rate	Pos err (m)
50–100 m	75.8% [68.6, 81.7]	15.17 [12.08, 18.57]	71.4% [64.0, 77.8]	12.85 [11.21, 14.63]
100–150 m	80.5% [75.8, 84.5]	17.56 [15.19, 20.20]	63.2% [57.8, 68.3]	23.40 [20.69, 26.54]
150–200 m	76.4% [72.6, 79.8]	28.98 [26.08, 32.13]	54.5% [50.2, 58.7]	33.91 [30.49, 37.88]
All	77.6% [74.9, 80.1]	23.04 [21.15, 24.97]	60.0% [56.9, 63.0]	26.35 [24.33, 28.61]

The two refinements of Section VII contribute asymmetrically in this regime, and a baseline (refinements-off) run on the same dataset isolates their effects. The multi-channel pipeline's largest gain is in localization rate: 51.1% baseline → 77.6% refined (+26.5 pp), driven by the coplanar-geometry recovery (Section VII-B) rescuing wide-of-field shots that the baseline Stage 2 rejects outright; position MAE drops by an additional 27% (31.4 m → 23.0 m). The single-channel pipeline's localization rate is unchanged (60.0% both runs), because that rate is set by the 5-inlier MB-cluster requirement rather than by the deceleration correction; the correction reduces position MAE by 23% (34.3 m → 26.4 m).

E. MULTIPATH ROBUSTNESS

The most operationally relevant scenario for a fielded system is acoustic multipath: walls, vehicles, and dense vegetation that produce specular reflections of the muzzle blast. We model multipath as a per-shot population of acoustic reflectors generated as follows. The number of reflectors per shot is drawn from a Poisson distribution with mean $\lambda = 1.5$. Each reflector is placed at a uniformly random location between 10 m and 100 m from the sensor-field centroid, and a 90° visibility cone determines which sensors can hear its reflection (sensors outside that cone receive no echo from that reflector). For each (reflector, visible-sensor) pair, an MB echo is generated with probability $p = 0.25$: the simulator computes the geometric path from shooter to reflector to sensor, applies the corresponding propagation delay to the original muzzle-blast event, and inserts the resulting delayed arrival as a substitute MB detection on any sensor whose direct-path MB detection has dropped out (e.g., from the distance-dependent detection-probability model in Table 5). The effect is to replace a fraction of the missing-MB observations with echo-derived (incorrect) MB arrivals, modeling the realistic failure mode in which a sensor's direct line of sight to the shooter is occluded but a strong reflection still triggers the MB detector. The model captures the dominant per-shot echo-on-MB failure mode but is intentionally stylized — it does not aim to reproduce structured urban impulse responses or frequency-dependent attenuation.

TABLE 8. Effect of acoustic multipath on the open-terrain scenario, refinements enabled. 95% confidence intervals beneath each value (Wilson for rates, bootstrap paired-difference for the gap). Both pipelines lose localization rate, but the single-channel architecture loses nearly twice as much: the loc-rate gap widens from 17.6 pp to 28.5 pp; the two CIs do not overlap, confirming the widening is not a sampling artifact.

Quantity	Open terrain		+ multipath	
	Loc rate	Pos err (m)	Loc rate	Pos err (m)
Multi-channel	77.6% [74.9, 80.1]	23.04 [21.22, 24.92]	63.9% [60.9, 66.8]	25.30 [22.93, 27.90]
Single-channel	60.0% [56.9, 63.0]	26.35 [24.33, 28.62]	35.4% [32.5, 38.4]	28.80 [25.73, 32.45]
Loc-rate gap (MC − SC)	17.6 pp [13.9, 21.2]		28.5 pp [25.0, 32.0]	

Table 8 captures the central robustness result. Acoustic multipath drops the multi-channel localization rate by 14 pp (77.6→63.9%) and the single-channel rate by 25 pp (60.0→35.4%); the multi-channel pipeline retains nearly twice the localization rate of single-channel under multipath. The architecture gap widens from 17.6 pp [13.9, 21.2] to 28.5 pp [25.0, 32.0]; the two intervals do not overlap, so the additional widening induced by multipath is statistically distinct from sampling variation. Position error stays roughly constant for both pipelines but only because each loses different shots: multipath-corrupted shots simply fall out of the localization-rate denominator for single-channel, while the multi-channel consistency function is more tolerant of mixed-source MB arrivals because it weighs muzzle-blast evidence against shockwave evidence and does not require either to be unanimous. The refinements continue to contribute the same way they did without multipath: a baseline (refinements-off) run on the same urban dataset yields MC localization rate 41.9% and position MAE 34.0 m versus the refined 63.9% / 25.3 m (+22 pp loc rate, −26% pos MAE); SC localization rate is unchanged at 35.4% (the 5-inlier requirement is the limiting factor, not the deceleration correction) and SC position MAE improves 22% (36.8 m → 28.8 m). The refinements target the constant-velocity bias and the wide-trajectory failure mode rather than echo handling, so the multipath-driven localization rate drops above are residuals that survive after both refinements are applied.

The mechanism explains the asymmetry, and Table 2 (Section V-E) tabulates it: the single-channel pipeline derives the final shooter position from a tight cluster of per-sensor MB−SW−TDoA-derived ranges that requires at least five mutually-consistent inliers, and an MB echo behaves as a moderate-magnitude bias on that sensor's range estimate that may or may not be rejected by clustering. The multi-channel consistency function fuses four evidence classes (per-array shooter, per-array range, MB AoA bearing, MB ToA cluster) and accepts agreement among any subset, so a corrupted MB arrival on one sensor simply drops out of the consensus rather than failing a hard inlier count. This is the central

architectural advantage of the multi-channel approach in cluttered environments.

F. LOCALIZATION TIME

Localization latency was measured on the 1000-shot mobile-deployment dataset of Section VIII-D, with the deceleration correction and coplanar-geometry recovery enabled. The reference implementation is MATLAB R2025a running single-threaded on one core of an Apple M1; each shot was processed three times and the minimum wall-clock time taken, with I/O excluded and a warm-up pass run beforehand to amortize JIT compilation. Both pipelines localize a shot in the low tens of milliseconds per event at the median: ∼12 ms for the multi-channel pipeline and ∼23 ms for the single-channel pipeline, with p99 latencies under ∼45 ms in both cases. These figures are well below the ∼100 ms acoustic propagation time across a 30–40 m sensor field, so localization is not the rate-limiting step of an end-to-end system on commodity hardware.

IX. RELATED WORK

Acoustic gunshot localization methods can be broadly categorized by their sensing architecture (centralized versus distributed), measurement types (ToA, TDoA, AoA, signal features), and fusion methodology (analytical, optimization-based, or learning-based). Most traditional approaches derive cost functions from measurement physics and formulate localization as minimizing fit error between observations and model predictions. However, such measurement-centric fusion struggles with outlier measurements common in urban operational environments where multipath propagation and non-line-of-sight detections introduce large, non-Gaussian errors. This section reviews relevant work across these dimensions and positions our consistency-based approach within this landscape.

A. FUSION METHODOLOGY: MEASUREMENT-CENTRIC VERSUS CONSISTENCY-BASED APPROACHES

A critical distinction exists between measurement-centric and hypothesis-centric fusion paradigms. Nearly all prior work formulates localization by deriving cost functions from individual sensor measurements—whether through least-squares multilateration, maximum-likelihood estimation, or weighted nonlinear optimization—then solving for shooter parameters that minimize aggregate error. Recent research exemplifies this measurement-centric paradigm: Ospel et al. [22] derive cost functions from ray-casting-based urban propagation models using TDoA-defined hyperbolic surfaces or amplitude-difference constraints. Still et al. [23] develop a unified maximum-likelihood framework for heterogeneous observations including muzzle-blast AoA, shockwave AoA, and inter-event TDoA, with an associated Cramér-Rao bound characterizing estimator performance. Their fusion strategy extends the likelihood function to handle missing or partial measurements but remains fundamentally measurement-centric in its formulation.

Several papers propose extensions to this classical paradigm. Osborne et al. [24] address multi-shooter scenarios through assignment-based fusion, resolving measurement-to-shot ambiguities before joint estimation. George and Kaplan [25] formulate weighted nonlinear least-squares for soldier-worn distributed systems and explicitly identify that sensor-reported confidence weights can be unreliable—motivating what they term a "consistency-based weighting scheme." However, their consistency concept concerns *weight reliability* rather than global geometric consistency among hypothesized shooter states. Other recent work explores novel measurements such as N-shape shockwave duration to reduce local-minima sensitivity in asynchronous TDoA systems [26], or applies machine-learning methods to improve classification and wind-noise robustness [27], [28].

In contrast to these measurement-centric approaches, our consistency-based fusion evaluates candidate shooter hypotheses by computing a mutual-consistency score across all available measurements and sensor subsets. Rather than minimizing measurement residuals, the method searches for the hypothesis (shooter position, projectile trajectory, shot time) supported by the maximum number of observations within measurement uncertainty bounds. This reframing provides inherent outlier rejection without requiring explicit outlier detection algorithms or assumptions about error distributions, enabling principled fusion even when outlier measurements outnumber valid observations—conditions under which classical measurement-based fusion typically fails.

B. CLASSICAL AND DISTRIBUTED LOCALIZATION METHODS

Early research established the foundational approaches for acoustic gunshot localization using muzzle-blast and ballistic-shockwave measurements. Classical systems by Duckworth et al. [8], Mazurek et al. [29] demonstrated that ToA, TDoA, and AoA measurements from distributed microphone arrays or wearable nodes enable accurate 2D and 3D shooter localization. Damarla's comprehensive overview [30] situates these methods within the broader context of battlefield acoustics, covering impulse detection, classification, and localization across diverse operational scenarios.

The evolution toward distributed sensing architectures addresses limitations of centralized arrays, particularly vulnerability to single-point failures and coverage constraints. Distributed systems leverage spatial diversity across multiple independent observation points, enabling measurement fusion algorithms to identify and reject multipath-corrupted observations. However, distributed architectures introduce challenges including sensor positioning uncertainty, time synchronization requirements, and the need for robust fusion in the presence of heterogeneous measurement quality across nodes. Our three architectures span different design points in the sensor density versus per-sensor complexity trade space,

demonstrating how spatial diversity with simple sensors can achieve accuracy comparable to or exceeding systems with sophisticated per-sensor processing.

C. MACHINE LEARNING FOR DETECTION AND CLASSIFICATION

Machine-learning methods have become increasingly prominent in acoustic gunshot detection and classification, but the overwhelming focus remains on law-enforcement and forensic applications rather than military rifle fire. Recent reviews emphasize that ML pipelines are designed primarily to discriminate impulsive gunshot signatures from urban noise and to identify common pistols, revolvers, and shotguns encountered in criminal investigations [1]. Deep learning approaches, especially CNNs applied to spectrogram representations, have further improved classification of firearm category and caliber, but again using datasets dominated by law-enforcement-relevant handguns [2], [31], [32]. Critically, the training data—typically YouTube videos or smartphone recordings captured at close range—differ fundamentally from operational sensor recordings at extended distances (50–300m or more), where signal characteristics, propagation effects, and noise conditions are substantially different. Overall, the ML literature leaves a notable gap in methods tailored for military supersonic rifle fire, where shock-wave geometry and battlefield conditions fundamentally differ from the handgun-dominated datasets used in current systems.

D. POSITION OF THE PRESENT WORK

The consistency-based measurement fusion framework presented in this paper differs fundamentally from prior work in its approach to handling outlier measurements and sensor failures. While recent advances improve localization through asynchronous model-based estimation with additional observables [26], propagation-aware solvers for complex urban scenes [22], and ML-based detection and classification [2], these methods remain measurement-centric in their fusion formulation. Our approach applies the same geometric and ballistic constraints but replaces global error minimization with a robustness-focused consistency framework capable of accurate localization even when outliers outnumber valid measurements.

X. DISCUSSION

This paper has presented two refinements to the consistency-function fusion paradigm – a caliber-aware deceleration correction (Section VII-A) and a coplanar-geometry recovery for the multi-channel pipeline (Section VII-B) – and a comparative evaluation of the two sparse architectures across four scenarios spanning noiseless ground truth, Aberdeen-calibrated noise, mobile-deployment noise, and acoustic multipath. This section first reflects on what the comparison says about architecture choice, then summarizes known limitations and avenues for future work.

A. ARCHITECTURE TRADE-OFF IS ENVIRONMENT-DEPENDENT

The single-channel architecture's principal advantage – one omnidirectional microphone per node, no orientation tracking, no per-array azimuth/elevation processing – translates into substantially lower hardware cost and simpler integration into existing soldier-worn equipment. The original Aberdeen evaluation [5] showed that this simplification did not cost accuracy: at $\leq$130 m on the Aberdeen test geometry the single-channel pipeline achieved 2.96 m position MAE, slightly better than the multi-channel system on the same data subset.

Our results corroborate this in clean line-of-sight conditions and reproduce the original Aberdeen numbers within a few percent (Section VIII-C) – both architectures land in the same $\sim$3–6 m position-error envelope on benign data. The architecture trade-off becomes apparent only when the noise environment is harder. Under mobile-deployment noise (Section VIII-D) the two architectures produce comparable position error at the engagement ranges that matter ($\sim$23–24 m all-range), but the multi-channel pipeline retains a 17 percentage-point advantage in localization rate, with much of the gap coming from shots where the multi-channel coplanar-geometry recovery can reconstruct a trajectory that the single-channel SW-duration constraint cannot. Adding acoustic multipath (Section VIII-E) widens that gap to nearly 29 percentage points – the multi-channel consistency function fuses four heterogeneous evidence classes and can absorb a small fraction of corrupted MB arrivals through outlier rejection on any one of them, while the single-channel pipeline's MB-cluster requirement of at least five mutually-consistent inliers is more brittle.

The implication is that architecture choice is not a universal hierarchy but an operational trade-off. For cost-constrained, line-of-sight engagements with surveyed or GPS-located sensors, the single-channel system delivers comparable accuracy at materially lower hardware cost and complexity. For deployments expected to encounter cluttered urban terrain, dense vegetation, or significant non-line-of-sight propagation, the multi-channel architecture's larger evidence pool delivers meaningfully higher localization rate.

B. LIMITATIONS AND FUTURE DIRECTIONS

The deceleration correction of Section VII-A compensates the constant-velocity assumption baked into the per-array shooter and per-array range formulas using the exact integrated time of a linearly-decelerating bullet. With damped fixed-point iteration it reduces noiseless position error by roughly an order of magnitude across all engagement ranges (16.6 $\rightarrow$ 1.0 m for multi-channel, 19.5 $\rightarrow$ 2.4 m for single-channel; Section VIII-B), including the long-range AK-47 subset where the bullet operates near Mach 1.3 at the emission point and the un-damped iteration would otherwise diverge.

The remaining limitation is the linear-deceleration model itself: a G1 ballistic-coefficient drag function would track velocity more accurately at low Mach excess, at the cost of the closed-form integrated correction that makes the present formulation attractive.

The evaluation is restricted to single-shot acoustic events. The consistency-function paradigm does not in principle require per-shot segmentation: the dense single-channel system at McKenna handled simultaneous shots from multiple shooters by identifying multiple distinct peaks of the consistency function [33]. Extending this to the sparse architectures — feed the full raw detection set to the consistency search, recover the strongest peak as the first shot, mark its contributing detections as consumed, and re-run — is a natural follow-up. The simulator's temporal-collision module, combined with the multi-shot generator already used here, is a ready test-bed for evaluating such an iterative peak extractor.

The most promising future-work direction is to fold the single-channel pipeline's distinguishing evidence class — the shockwave duration constraint $b \approx kT^4$ (Section V) — into the multi-channel pipeline's consistency function as a fifth evidence class. Because SW duration is determined by the N-wave shape and is largely independent of the MB ToA pathway, corrupted MB arrivals do not also corrupt the duration evidence, suggesting that the hybrid should widen the multi-channel loc-rate advantage under multipath beyond the 29 pp reported in Section VIII-E while tightening accuracy on benign data.

REFERENCES

[1] Y. Teng, K. Zhang, X. Lv, Q. Miao, T. Zang, A. Yu, A. Hui, and H. Wu, "Gunshots detection, identification, and classification: Applications to forensic science," *Sci. Justice*, vol. 64, no. 6, pp. 625–636, 2024.

[2] S. Raponi, I. Ali, and G. Oligeri, "Sound of guns: Digital forensics of gun audio samples meets artificial intelligence," 2020, *arXiv:2004.07948*.

[3] G. Simon, M. Maróti, Á. Lédeczi, G. Balogh, B. Kusy, A. Nádas, G. Pap, J. Sallai, and K. Frampton, "Sensor network-based countersniper system," in *Proc. 2nd Int. Conf. Embedded networked sensor Syst.*, Nov. 2004, pp. 1–12.

[4] P. Völgyesi, G. Balogh, A. Nádas, C. B. Nash, and Á. Lédeczi, "Shooter localization and weapon classification with soldier-wearable networked sensors," in *Proc. 5th Int. Conf. Mobile Syst., Appl. services*, Jun. 2007, pp. 113–126.

[5] J. Sallai, Á. Lédeczi, and P. Völgyesi, "Acoustic shooter localization with a minimal number of single-channel wireless sensor nodes," in *Proc. 9th ACM Conf. Embedded Networked Sensor Syst.*, Nov. 2011, pp. 96–107.

[6] R. Stoughton, "Measurements of small-caliber ballistic shock waves in air," *J. Acoust. Soc. Amer.*, vol. 102, no. 2, pp. 781–787, Aug. 1997.

[7] G. B. Whitham, "The flow pattern of a supersonic projectile," *Commun. Pure Appl. Math.*, vol. 5, no. 3, pp. 301–348, Aug. 1952.

[8] G. L. Duckworth, D. C. Gilbert, and J. E. Barger, "Acoustic counter-sniper system," *Proc. SPIE*, vol. 2938, pp. 262–275, Jul. 1997.

[9] E. Danicki, "Acoustic sniper localization," *Arch. Acoust.*, vol. 30, no. 2, pp. 233–245, 2005.

[10] T. Damarla, L. M. Kaplan, and G. T. Whipps, "Sniper localization using acoustic asynchronous sensors," *IEEE Sensors J.*, vol. 10, no. 9, pp. 1469–1478, Sep. 2010.

[11] D. Lindgren, O. Wilsson, F. Gustafsson, and H. Habberstad, "Shooter localization in wireless microphone networks," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, Dec. 2010, Art. no. 690732.

[12] B. M. Sadler, T. Pham, and L. C. Sadler, "Optimal and wavelet-based shock wave detection and estimation," *J. Acoust. Soc. Amer.*, vol. 104, no. 2, pp. 955–963, Aug. 1998.

[13] J. Sallai, P. Völgyesi, K. Pence, and Á. Lédeczi, "Fusing distributed muzzle blast and shockwave detections," in *Proc. 14th Int. Conf. Inf. Fusion*, Jul. 2011, pp. 1–8.

[14] Á. Lédeczi, A. Nádas, P. Völgyesi, G. Balogh, B. Kusy, J. Sallai, G. Pap, S. Dóra, K. Molnár, M. Maróti, and G. Simon, "Countersniper system for urban warfare," *ACM Trans. Sensor Netw.*, vol. 1, no. 2, pp. 153–177, Nov. 2005.

[15] M. Compagnoni, A. Pini, A. Canclini, P. Bestagini, F. Antonacci, S. Tubaro, and A. Sarti, "A geometrical–statistical approach to outlier removal for TDOA measurements," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 3960–3975, Aug. 2017.

[16] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.

[17] G. Simon and G. Zachár, "Fast and fault-tolerant passive hyperbolic localization using sensor consensus," *Sensors*, vol. 24, no. 9, p. 2891, Apr. 2024. [Online]. Available: https://www.mdpi.com/1424-8220/24/9/2891

[18] R. L. McCoy, *Modern Exterior Ballistics: The Launch and Flight Dynamics of Symmetric Projectiles*. Atglen, PA, USA: Schiffer Publishing, 1999.

[19] J. Sahu and K. R. Heavey, "Numerical Investigation of Supersonic Base Flow with Base Bleed," *J. Spacecraft Rockets*, vol. 34, no. 1, pp. 62–69, 1997, doi: 10.2514/2.3173.

[20] B. A. Weiss, C. Schlenoff, M. Shneier, and A. Virts, "Technology evaluations and performance metrics for soldier-worn sensors for ASSIST," in *Proc. Perform. Metrics Intell. Syst. Workshop*, 2006, pp. 157–164.

[21] J. Sallai, B. Kusy, Á. Lédeczi, and P. Dutta, "On the scalability of routing integrated time synchronization," in *Proc. Eur. Workshop Wireless Sensor Netw.* Berlin, Germany: Springer, 2006, pp. 115–131.

[22] M. Ospel, C. Blonde-Weinmann, P. Hamery, D.-O. Iacob, and S. Hengy, "Fast iterative shooter localization in urban terrain using ray casting," *J. Acoust. Soc. Amer.*, vol. 155, no. 5, pp. 3005–3013, May 2024.

[23] L. Still, M. Oispuu, and W. Koch, "Accuracy study on shooter localization using incomplete acoustic measurements," in *Proc. IEEE 23rd Int. Conf. Inf. Fusion (FUSION)*, Jul. 2020, pp. 1–8.

[24] R. W. Osborne, Y. Bar-Shalom, J. George, and L. Kaplan, "Data fusion from multiple passive sensors for multiple shooter localization via assignment," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–7.

[25] J. George and L. M. Kaplan, "Shooter localization using a wireless sensor network of soldier-worn gunfire detection systems," *J. Adv. Inf. Fusion*, vol. 8, no. 1, pp. 15–32, 2013.

[26] S. Koch, L. Still, M. Oispuu, and W. Koch, "Shooter localization based on TDOA and N-shape length measurements of distributed microphones," in *Proc. IEEE 24th Int. Conf. Inf. Fusion (FUSION)*, Nov. 2021, pp. 1–8.

[27] S. A. Qureshi, L. Hussain, H. M. Alshahrani, S. R. Abbas, M. K. Nour, N. Fatima, M. I. Khalid, H. Sohail, A. Mohamed, and A. M. Hilal, "Gunshots Localization and Classification Model Based on Wind Noise Sensitivity Analysis Using Extreme Learning Machine," *IEEE Access*, vol. 10, pp. 87302–87321, 2022, doi: 10.1109/ACCESS.2022.3198966.

[28] J. Svatos and J. Holub, "Impulse acoustic event detection, classification, and localization system," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–15, 2023.

[29] 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," *Proc. SPIE*, vol. 5778, pp. 264–282, May 2005.

[30] T. Damarla, *Battlefield Acoustics*. Cham, Switzerland: Springer, 2015.

[31] P. Doungpaisan and P. Khunarsa, "Deep spectrogram learning for gunshot classification: A comparative study of CNN architectures and time-frequency representations," *J. Imag.*, vol. 11, no. 8, p. 281, Aug. 2025.

[32] A. Shah, R. Singh, B. Raj, and A. Hauptmann, "Deciphering GunType hierarchy through acoustic analysis of gunshot recordings," 2025, *arXiv:2506.20609*.

[33] Á. Lédeczi, P. Völgyesi, M. Maróti, G. Simon, G. Balogh, A. Nádas, B. Kusy, S. Dóra, and G. Pap, "Multiple simultaneous acoustic source localization in urban terrain," in *Proc. 4th Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, 2005, pp. 491–496.



ÁKOS LÉDECZI (Senior Member, IEEE) received the bachelor's degree in electrical engineering from the Technical University of Budapest, in 1989, and the Ph.D. degree in electrical engineering from Vanderbilt University, in 1995. He is currently a Professor of computer science, computer engineering, and electrical engineering with Vanderbilt University. He also holds a courtesy appointment as a Professor of teaching and learning. His current research interests include cyber-physical systems and computer science education.



JÁNOS SALLAI received the bachelor's degree in computer science from the Technical University of Budapest, in 2002, and the Ph.D. degree in computer science from Vanderbilt University, in 2009. He is currently the Vice President of engineering of Xtremis AI, a startup company working on cutting edge RF spectrum intelligence. Previously, he was the Chief Scientist of Databuoy Corporation, creating acoustic gunshot location systems for law enforcement and critical infrastructure protection.



PÉTER VÖLGYESI received the M.S. degree in computer science from Budapest University of Technology and Economics, in 2000. Currently, he is a Principal Research Scientist with Vanderbilt's Institute for Software Integrated Systems. He is a Founder and the current CTO of Xtremis AI, a defense-focused spectrum intelligence company. His research interests include distributed RF sensing and software-defined radio systems.



MIKLÓS MARÓTI received the M.S. degree from the University of Szeged, in 1996, the Ph.D. degree in mathematics from Vanderbilt University, in 2002, and the Ph.D. degree in mathematics from the University of Szeged, in 2007. He is currently an Associate Professor and the Head of the Department of Algebra and Number Theory, University of Szeged. His research interests include universal algebra, lattice theory, logic, and applied areas, such as wireless sensor networks and RF spectrum sensing.



WILL HEDGECOCK (Member, IEEE) received the bachelor's degree in computer engineering and the Ph.D. degree in electrical engineering from Vanderbilt University, in 2008 and 2014, respectively. He is currently a Senior Research Scientist with Vanderbilt's Institute for Software Integrated Systems and an Adjunct Assistant Professor of computer science. His research interests include wireless sensor networks, cyber-physical systems, and resource-constrained localization techniques.



GYULA SIMON (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Budapest University of Technology, Hungary, in 1991 and 1998, respectively. Currently, he is a Professor with Obuda University, Budapest, Hungary. His research interests include adaptive signal processing, localization, and sensor networks.