Audio Engineering Society

Convention Paper

Presented at the 145th Convention
2018 October 17–20, New York, NY, USA

Audio Forensic Gunshot Analysis and Multilateration

Robert C. Maher and Ethan R. Hoerr

Electrical & Computer Engineering, Montana State University, Bozeman, MT USA 59717-3780
Correspondence should be addressed to R.C. Maher (rob.maher@montana.edu)

ABSTRACT

This paper considers the opportunities and challenges of acoustic multilateration in gunshot forensics cases. Audio forensic investigations involving gunshot sounds may consist of multiple simultaneous but unsynchronized recordings obtained in the vicinity of the shooting incident. The multiple recordings may provide information useful to the forensic investigation, such as the location and orientation of the firearm, and if multiple guns were present, addressing the common question “who shot first?” Sound source localization from multiple recordings typically employs time difference of arrival (TDOA) estimation and related principles known as multilateration. In theory, multilateration can provide a good estimate of the sound source location, but in practice acoustic echoes, refraction, diffraction, reverberation, noise, and spatial/temporal uncertainty can be confounding.

1 Introduction

Audio forensic evidence is of increasing importance in law enforcement investigations because of the growing use of personal audio/video recorders carried by officers on duty, the routine use of dashboard audio/video systems in police cruisers, and the increasing likelihood that images and sounds from criminal incidents will be captured by public or private surveillance systems or mobile handheld devices. In some cases, gunshots and other firearm sounds may be captured by these recording devices [1, 2, 3].

In the United States, criminal actions involving firearms are of ongoing concern to law enforcement and the public. For example, the Bureau of Justice Statistics reported over 467,000 individuals experienced a non-fatal criminal incident involving a firearm in 2011, and over 11,000 individuals died from firearm homicides that year [4].

In some cases, the sound of gunfire may be recorded simultaneously by two or more different microphones located in proximity to the scene of the shooting. Audio forensic examiners may be asked to reconstruct a shooting scene based on these multiple audio recordings using the relative time-of-arrival of the gunshot sound at the several microphone locations—assuming the recordings are somehow synchronized in time, and the precise spatial location of the microphones is known.

Although commercial gunshot detection and localization systems are available using deliberately deployed and time-synchronized microphones at known locations [5], in the case of forensic reconstruction it is more likely that an audio forensic examiner will need to use an *ad hoc* collection of recordings from mobile audio recording devices that happened to be in acoustic range of the gunshot. Along with the need to estimate proper time synchronization, the examiner will also have to estimate the microphone positions, potential reflections, and the local speed of sound determined by an estimate of the air temperature.

This paper is organized as follows. First, we review the mathematical formulation of acoustic *multilateration*, which is the proper general term describing techniques to estimate the location of a sound source given a set of spatially distributed observing microphones. Next, we present an example scenario involving gunshot observations and representative forensic questions typical of audio forensic cases. We simulate acoustical observations with various degrees of uncertainty about the time synchronization and microphone positions. Finally, we summarize the practical considerations with which audio forensic examiners need to be aware.

2 Acoustic multilateration

When a source produces a sound pulse, the pulse will generally arrive at slightly different times at two spatially separated microphones. The time of arrival difference is due to the difference in path length from the source to each of the two microphones, assuming a constant speed of sound. The use of the *time difference of arrival* (TDOA) at two known microphone locations identifies a locus of possible positions of the source with respect to the two microphones. The possible source locations comprise all positions that have the same *difference* in distance from the two microphones that result in the measured TDOA. The source position estimation based upon TDOA is known as *multilateration* [6].[1]

For example, consider a simplified two-dimensional (planar) case with two microphones located at known Cartesian coordinates (x_1, y_1) and (x_2, y_2), respectively, and the sound source at an unknown location (x, y), as shown in Figure 1. The pulse produced by the source at (x, y) arrives at the Mic 1 position Δt seconds before it arrives at the Mic 2 position.

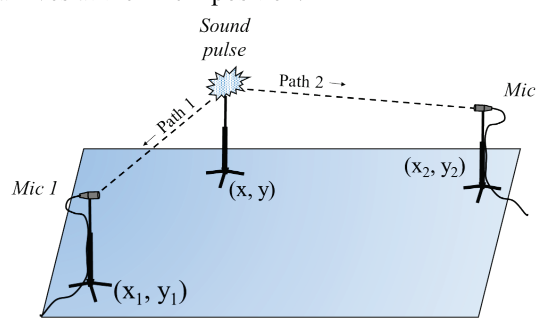

Figure 1: Example geometry for unknown source position and known microphone pair.

In general, the absolute time at which the sound pulse occurs at (x, y) is not known, only the relative time of arrival at the two microphones, Δt. Using the speed of

[1] The TDOA multilateration approach is sometimes erroneously referred to as *triangulation*, a different procedure using *angle* measurements.

sound, c, in meters per second, the path length difference ($path_2 - path_1$) is $c\Delta t$ meters. Thus, any point (x, y) for which the distance ($path_2 - path_1$) is $c\Delta t$ could be the location of the pulse source.

From the example geometry,

$$|path_1| = \sqrt{(\mathrm{x} - \mathrm{x}_1)^2 + (\mathrm{y} - \mathrm{y}_1)^2} \quad (1)$$

$$|path_2| = \sqrt{(x - x_2)^2 + (y - y_2)^2} \quad (2)$$

and

$$|path_2| - |path_1| = c\Delta t \quad (3)$$

Accordingly, the unknown (x, y) coordinates of the pulse source need to satisfy Equation 3, given the TDOA Δt. Since there are two unknowns and one equation, there are many possible solutions. The solutions can be found numerically or analytically.

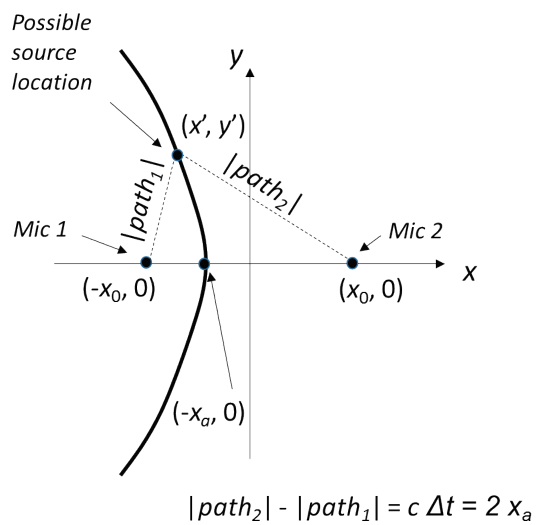

Figure 2: Multilateration geometry with translated and rotated coordinate system.

To find the analytical solution, it is convenient mathematically to rotate and translate the coordinate system so that the origin is placed at the middle of a line connecting the two microphones, and representing the x-axis (abscissa) of the Cartesian system. This geometry, plan view, is shown in Figure 2. Any point (x', y') along the curve is a possible source location. For example, in this coordinate system the point ($-x_a$, 0) represents a solution such that $c\Delta t = (x_0+x_a) - (x_0-x_a)$, or $x_a = c(\Delta t)/2$.

In the new coordinate system, the equations equivalent to Equations 1-3 become:

$$|path_1| = \sqrt{(x' + x_0)^2 + (y')^2} \quad (3)$$

$$|path_2| = \sqrt{(x' - x_0)^2 + (y')^2} \quad (4)$$

and

$$|path_2| - |path_1| = c\Delta t = 2x_a \ . \quad (5)$$

Inserting (3) and (4) into (5), squaring and expanding to isolate the radical, then squaring again to simplify into a form with x' and y', the result is the familiar *hyperbola* formula:

$$\frac{(x')^2}{x_a^2} - \frac{(y')^2}{(x_0^2 - x_a^2)} = 1 \ , \quad (6)$$

where

$$x_a^2 = \frac{(c\Delta t)^2}{4} \ . \quad (7)$$

A compact calculation based upon aligning the origin of the coordinate system at the closest sensor is also possible [7].

From (6) and (7), it becomes apparent that the multilateration expression will have a singularity if $c(\Delta t)$ becomes equal to the microphone spacing, $2x_0$. In other words, referring to Figure 2, the time difference of arrival (Δt) will have its maximum value if the sound source is in-line with the x-axis at a position $|x'| > x_0$ and $y'=0$, for which the hyperbola devolves into a line segment. In general, if the microphones are closely spaced (small x_0), the acceptable range of Δt is correspondingly small ($\Delta t < 2\,x_0/c$), so the precision of the time difference measurement is particularly important, as will be considered later.

With $N > 3$ microphones in a two-dimensional plane, there are N-1 TDOA values, and these mic-to-mic differences can be used to create N-1 hyperbolas. In principle, the analytical procedure resulting in Equations 6 and 7 can be repeated for each time difference, and the resulting intersection point of the hyperbolas would, at least theoretically, represent the location of the sound source.

In practice, the audio recordings used to determine the TDOA will contain noise, and therefore the time estimates will have some level of uncertainty. A timing discrepancy for the N-1 Δt values means that the solution to the intersection of multiple hyperbolas is also uncertain. The solution of multiple non-linear equations in the presence of noise is typically handled with numerical solvers rather than a closed-form analytical solution, and such methods are applicable in the case of TDOA multilateration.

If the sound source is not an impulse, the time delay estimate can be obtained using cross-correlation of the received signals. The reliability of the time estimates will depend upon the characteristics of the signal and any interfering noise for the correlation operation.

3 The forensic situation

As described in the introduction, a relatively common contemporary scenario in audio forensic gunshot investigations involves several audio recordings of a shooting incident captured simultaneously by multiple unsynchronized recording systems positioned in an indiscriminate manner. If the incident takes place out of the view of the camera(s), or occurs at night or otherwise without good recorded images, the audio information may be very important to the investigation. Such a scenario (see Figure 3) could involve the sound of a gunfire recorded simultaneously by several dashboard camera systems in law enforcement vehicles, body recorders worn by officers, nearby commercial or residential surveillance systems, and even mobile recording devices such as cell phone video [8, 9, 10, 11].

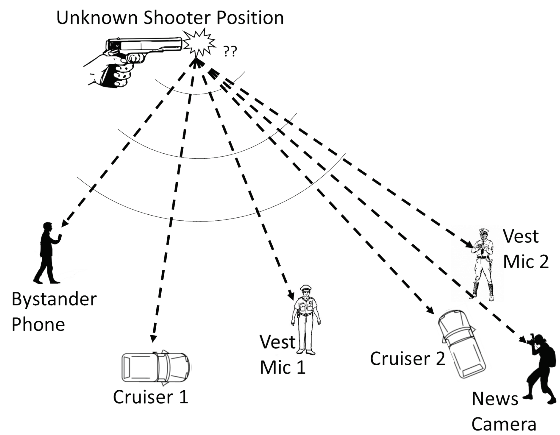

Figure 3: Scenario with multiple simultaneous but unsynchronized audio recordings from spatially-distributed microphones.

Questions that may arise for the audio forensic examiner could include [12]:

- *What is the estimated location of the first gunshot with respect to a reference position, based upon the audio evidence?*
- *Were the second and subsequent shots from the same position as the first shot?*

3.1 Position estimation

In order to attempt multilateration using the available audio recordings, the first requirement is obtaining a reliable estimate of the position of each recording device at the incident scene. Depending upon the circumstances and the type of recordings available, the position determination may be relatively straightforward, or it may be quite ambiguous. The audio forensic examiner would need to determine an appropriate spatial reference point, such as an individual with a microphone who is visible in one or more of the dashcam videos, and who is standing at a spatially identifiable spot, such as near a street sign or fence post. The other recording devices would then need to be located with respect to the known reference point.

In some circumstances there may be survey information, diagrams, and maps prepared by crime scene analysts. In other circumstances there may be still photographs of the scene, witness recollections, or other sources of spatial information. A common challenge is that recordings from vest cameras or mobile handheld devices often take place while the individual is moving, not stationary in a fixed and identifiable location. The audio forensic examiner will need to determine all of the information and uncertainty associated with the estimated recording positions.

3.2 Synchronization

In order to attempt multilateration using the recordings, the second requirement is *synchronization*. The multiple recordings generally do not have a common clock, although dashboard camera systems often have at least two audio channels recorded synchronously. One channel is typically from a microphone inside the cabin of the vehicle, and the other channel is often fed by a wireless microphone worn by the law enforcement officer. The various recordings may also have different sampling rates and formats.

The first step is to collect and catalog all of the available recordings. Files that are in compressed format (e.g., MP3 or a proprietary codec) need to be decoded into standard PCM .wav files. It is also convenient to perform high-quality sampling rate conversion to get all of the available audio at a common high sampling rate, such as 48 kHz.

The second step is to start with recording channels that are known to be synchronized, such as the dashboard recordings mentioned above. If multiple dashcams were at the scene, a useful strategy is to identify an audible signal that is common to all of the dashcams. One possible solution is to identify a signal transmitted from the law enforcement dispatch center and picked up simultaneously by the radios in the various cruisers: the cabin microphone in each dashboard recorder would capture the dispatch signal. The audible signal common to the cabin recordings is used as a time reference to align the associated recordings.

If no dispatcher radio signal is available, and for other bystander recordings, the examiner would need to attempt to find another common signal from a known location at the scene, and arrange to use that signal for time alignment.

3.3 TDOA estimation

In the case of gunshots, the onset of the sound may be clearly observed in each of the synchronized audio recordings. The audio forensic examiner would need to use a waveform display program to identify the corresponding time instants, or use a cross-correlation procedure to find the time delay corresponding to the best alignment.

However, there are often significant problems in trying to interpret distorted recordings if the microphone system was close enough to the sound that the signal has overloaded the recording system. For example, the output of some audio recording systems essentially drops out when presented with extremely loud sounds (Figure 4).

If the recording device uses an audio speech encoding system (e.g., VSELP), the recording may not be able to represent an impulsive signal such as a gunshot in a manner suitable for time-of-arrival determination. In other cases the microphone may not be in a direct line-of-sight position with respect to the gunshot, so the received signal is actually reflected off nearby surfaces, diffracted around obstacles, or otherwise traveled over a path not radial to the sound source. The audio forensic examiner needs to consider all of these issues [12].

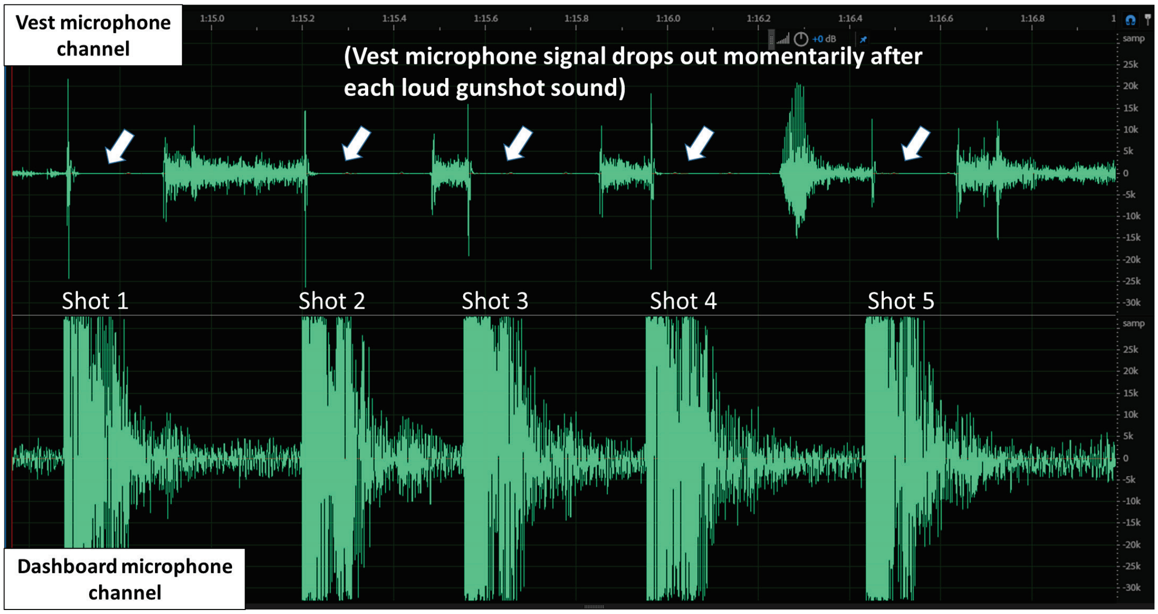

Figure 4: Example forensic audio recording with two simultaneous channels, from a confidential law enforcement source. Upper: a wireless vest microphone, showing signal drop outs due to overloading. Lower: a microphone in the cabin of the law enforcement vehicle, showing clipping but no drop out.

4 Example investigation

To contemplate several practical questions related to multilateration from heterogeneous audio recordings of a gunshot incident, consider the following two-dimensional example scenario.

- A gun is fired at arbitrary location: [-20 meters, +36 meters].
- The sound is observed at four arbitrary microphone locations, for what we will refer to as Configuration 1:
  [-7, -3], [-3, 1], [2, 0], and [5, -1] meters.
- The distances from the source to the microphones are:
  41.11 m, 38.91 m, 42.19 m, and 44.65 m (microphone 2 is closest).
- Using speed of sound c = 343.2 m/s (air temperature 20° C), the theoretical TDOAs with respect to microphone 2 are:
  $\Delta t_{1\rightarrow2}$= 6.408 ms, $\Delta t_{3\rightarrow2}$= 9.556 ms, and $\Delta t_{4\rightarrow2}$= 16.736 ms.

The plan view of this 2-D example configuration is shown in Figure 5.

4.1 Multilateration using Configuration 1

Now running the multilateration algorithm given the specific microphone coordinates and the expected TDOAs to simulate a forensic location estimation task, the calculated source position is [-18.92, 33.59] meters, which is an estimation error of [-1.08, 2.41] meters compared to the true location. The discrepancy is due to the numerical calculations used in the multilateration. Position discrepancies become more likely if the source is located in a positon that is in-line with the inter-sensor vector, because this orientation obscures the source distance.

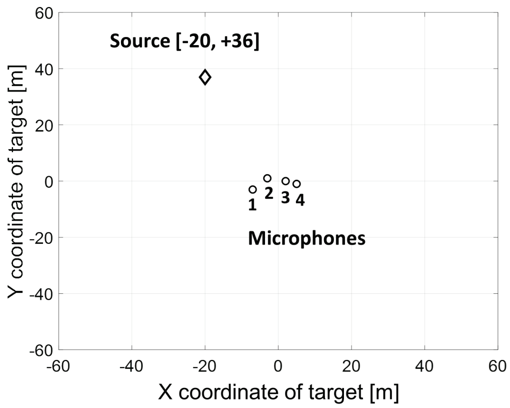

Figure 5: Example Configuration 1.

4.2 Multilateration using Configuration 2

Next, we keep the source location fixed, but consider a different scenario in which the microphones are spread slightly farther apart to new arbitrary locations, giving Configuration 2: [-10, -3], [-4, 1], [5, 0], and [10, -1] meters, as shown in Figure 6.

In this configuration the multilateration algorithm calculates the source position to be [-19.97, 35.88] meters, which gives an estimation error of [-0.03, 0.12] meters compared to the true location. Wider sensor spacing can provide greater inter-sensor time delay, and lowered sensitivity to timing and position errors.

4.3 Position uncertainty

However, if the microphone locations are uncertain, what does this do to the multilateration solution? Using Configuration 1 and Configuration 2, we perform a Monte Carlo simulation in which we randomly move each of the four simulated microphone positions within a 0.5 meter radius of the specified position, while keeping the source location unchanged. The TDOA information is then used to find the estimated source location, which is compared to the true location.

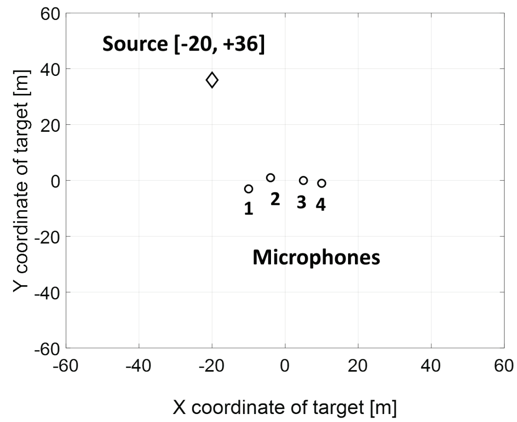

Figure 6: Example Configuration 2 (microphones spread slightly more in the horizontal direction).

While not a perfect simulation of the uncertainty found in a real forensic case, the results are intended to show how an imperfect set of microphone location estimates contribute to uncertainty in the multilateration result.

The estimated source location compared to the true location for Configuration 1 with the 0.5 meter microphone position uncertainty is shown in Figure 7.

The result is that the position discrepancies can give substantial estimation errors, although the magnitude of the error varies for different microphone locations. Additional review indicates that the estimated direction (azimuth) of the source is often nearly correct, which may be useful in certain circumstances.

Performing 1000 trials with Configuration 2 (microphones spaced slightly farther apart) and again moving the microphone positions randomly within a 0.5 meter radius, the estimated source location compared to the true location is shown in Figure 8.

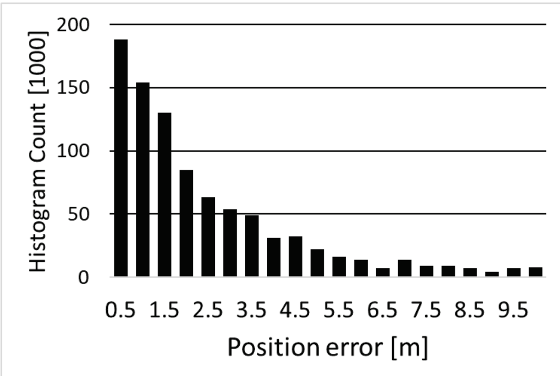

Figure 7: Simulated source location estimation error for Configuration 1 (in meters) for 1000 trials with 0.5 meter uncertainty in microphone positions.

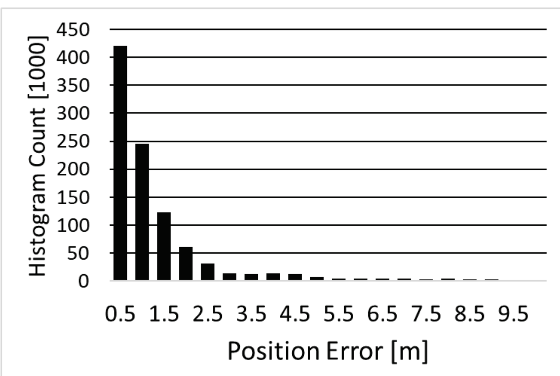

Figure 8: Simulated source location estimation error for Configuration 2 (in meters) for 1000 trials with 0.5 meter uncertainty in microphone positions.

As found with Configuration 1, Configuration 2 also shows substantial estimation errors due to the position discrepancy, but the sensitivity is reduced: more of the estimates are close to the true value with the wider spacing of Configuration 2.

From the simulation, it is clear that uncertainty in the microphone positions—even as seemingly small as ±0.5 meters—can lead to substantial errors in the multilateration estimate of the source position. An audio forensic examination involving *ad hoc* collections of uncertain microphone positions must proceed with caution regarding any conclusions about the estimated source location.

5 Conclusions

This paper has described several considerations regarding practical sound source localization from multiple recordings obtained assuming a two-dimensional geometry. The well-known multilateration procedure employs time difference of arrival (TDOA) information at the microphone positions. In theory, multilateration provides a good estimate of the sound source location. In practice, there can be difficulty in determining the TDOA values in the presence of noise, acoustic echoes, refraction, diffraction, and reverberation. As demonstrated using simulated geometries, uncertainty about the spatial location of the sensors can lead to errors in the multilateration results. An audio forensic examiner should take this uncertainty into account when performing calculations using data from multiple unsynchronized recording systems positioned in an indiscriminate manner [13].

References

[1] R.C. Maher, "Audio forensic examination: authenticity, enhancement, and interpretation," IEEE Sig. Proc. Mag., vol. 26, pp. 84-94 (2009).

[2] R.C. Maher, "Lending an ear in the courtroom: forensic acoustics," Acoustics Today, vol. 11, no. 3, pp. 22-29 (2015).

[3] R.C. Maher and S.R. Shaw, "Deciphering gunshot recordings," Proc. Audio Eng. Soc. 33rd Conf. Audio Forensics—Theory and Practice, Denver, CO (2008).

[4] Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice, https://www.bjs.gov/content/pub/press/fv9311pr.cfm

[5] G.L. Duckworth, D.C. Gilbert, and J.E. Barger, "Acoustic counter-sniper system," Proc. of SPIE, Command, Control, Communications, and Intelligence Systems for Law Enforcement, E.M. Carapezza and D. Spector, eds., pp. 262-275 (1997).

[6] W.R. Fried and M. Kayton, *Avionics Navigation Systems*, Wiley (1969).

[7] B.T. Fang, "Simple Solutions for Hyperbolic and Related Position Fixes," IEEE Trans. on Aerospace Elect. Sys., vol. 26, no. 5 (1990).

[8] B.M. Brustad and J.C. Freytag, "A survey of audio forensic gunshot investigations," Proc. Audio Eng. Soc. 26th Conf. Audio Forensics in the Digital Age, Denver, CO. (2005).

[9] R.C. Maher, "Modeling and signal processing of acoustic gunshot recordings," Proc. IEEE Signal Processing Society 12th DSP Workshop, Jackson Lake, WY, pp. 257-261 (2006).

[10] R.C. Maher, "Acoustical characterization of gunshots," in Proc. IEEE SAFE Workshop on Signal Processing Applications for Public Security and Forensics, Washington, D.C. (2007).

[11] S.D. Beck, H. Nakasone, and K.W. Marr, "Variations in recorded acoustic gunshot waveforms generated by small firearms," J. Acoust. Soc. Am., vol. 129, no. 4, pp. 1748-1759 (2011).

[12] R.C. Maher, "Gunshot recordings from a criminal incident: who shot first?" J. Acoust. Soc. Am., vol. 139, no. 4, p. 2024 (2016). Lay-language version: http://acoustics.org/2aaaa7-gunshot-recordings-from-a-criminal-incident-who-shot-first-robert-c-maher/

[13] National Academy of Sciences, *Strengthening Forensic Science in the United States: A Path Forward*, National Academy Press, Washington, DC. (2009).