# Acoustic Source Localization in a Network of Doppler Shift Sensors

David Lindgren, Mehmet B. Guldogan, Fredrik Gustafsson, Hans Habberstad and Gustaf Hendeby



Original Publication:

David Lindgren, Mehmet B. Guldogan, Fredrik Gustafsson, Hans Habberstad and Gustaf Hendeby, Acoustic Source Localization in a Network of Doppler Shift Sensors, 2013, 16th International Conference on Information Fusion.

# Acoustic Source Localization in a Network of Doppler Shift Sensors

David Lindgren* Mehmet B. Guldogan† Fredrik Gustafsson‡ Hans Habberstad* Gustaf Hendeby*
*Swedish Defence Research Agency, FOI, Email: name.familyname@foi.se
†Turgut Ozal University, Turkey, Email: bguldogan@turgutozal.edu.tr
‡Linköping University, Sweden, Email: fredrik@isy.liu.se

***Abstract*—It is well-known that the motion of an acoustic source can be estimated from Doppler shift observations. It is however not obvious how to design a sensor network to efficiently deliver the localization service. In this work a rather simplistic motion model is proposed that is aimed at sensor networks with realistic numbers of sensor nodes. It is also described how to efficiently solve the associated least squares optimization problem by Gauss-Newton variable projection techniques, and how to initiate the numerical search from simple features extracted from the observed frequency series. The methods are demonstrated on real data by determining the distance to a passing propeller-driven aircraft and by localizing an all-terrain vehicle. It is concluded that the processing components included are fairly mature for practical implementations in sensor networks.**

## I. Introduction

The motion of a passing acoustic source can be estimated from the observed Doppler shift, even by a single sensor.

We consider an acoustic sensor network with one or more synchronized nodes with known positions. It is assumed that the Doppler frequency of a passing acoustic source can be detected and measured by each sensor. It is further assumed that the source motion obeys a parameterized model. We address the problem of identifying the model parameters, from which the source trajectory is deduced. Specifically, the case with few sensors is in focus—say one, or perhaps three sensors.

The problem of localizing moving acoustic sources based on Doppler shift observations has received fairly much attention in the literature, especially in a period around 1990. Still, there are aspects within the area that are not entirely covered, and today it would be particularly interesting to revive the interest for the acoustic Doppler phenomenon with the background of both the last decade of advancement of wireless sensor network technologies and new opportunities to solve the associated numerical problems with state-of-the-art algorithms.

A number of authors have considered the case with a single sensor by which the source passage distance and assumed constant speed are estimated [1], [2], [3], [4]. It appears, however, that the estimation problem is inherently nonlinear in the parameters, and a rather wide range of different numerical approaches are proposed to pursue the, in various senses, optimal estimate. Most commonly, the parameters are given by minimizing some variation of a least squares criterion.

A complication in the acoustic case is the *retardation*; the effect that, when the sound reaches the sensors, the source has already moved to a new position. In the radar case, this effect can usually be disregarded since the propagation speed of the electromagnetic waves is many orders of magnitudes larger than the speed of most sources of practical interest (the "stop-and-go approximation"), see [5, pp. 97]. Disregarding the retardation in the acoustic case may, depending on geometry and source speed, however result in significant estimation bias.

The approximate retardation-free model can apparently be parameterized with only two nonlinear parameters (instead of three or four), for instance, an $x$ and a $y$ coordinate on the source trajectory. This low parameter dimension facilitates an exhaustive grid search for the (least squares) minimum, which avoids otherwise problematic local minima. This is elegantly demonstrated by [2]. By also using the frequency time derivative as a measurement, the grid search can be conducted in only one dimension, see [6]. Actually, by disregarding the retardation, an exact solution based on a deterministic model and three frequency measurements can be found by calculating the roots of a third order polynomial, see [1].

Another estimation technique is to match an observed frequency series with a library of precomputed Doppler shift profiles. This way, a complex retardation model may be used as a basis, and the estimation step can be implemented on a simple platform (wireless sensor) with computational limitations, as showed by [4].

The validity of a retardation model has been demonstrated by having a single sensor estimating the distance to, and speed of, a passing propeller-driven aircraft, see [3].

There are also examples of sensor networks with multiple sensors that use Doppler shift for source localization. A way to do this is to model the measured frequency difference (in effect of the Doppler shift) among every sensor pair in the network. This rich set of measurements may then be informative enough to identify more complex motion models, for instance ballistic trajectories, see [7]. Another (suboptimal) way to configure the network is to have each sensor make a local estimate of the passage distance and speed, and then transmit the results to a fusion node that in turn aggregates the individual estimates to a more informative output in terms of global trajectory coordinates. This was demonstrated by [2] and later by [8].

There are many reasons for keeping the number of sensors in a network moderate, such as purchase, deployment, and maintenance costs. Balancing costs, in terms of sensor count, with performance is an essential part of the sensor system design that requires careful analysis. For instance, with as few as five Doppler sensors (with a non-degenerate geographical distribution) it is possible to uniquely and instantaneously determine the source position, see [9]. With multiple consecutive

observations, fewer sensors, perhaps three, suffices.

The scientific contribution of this work includes two motion models for acoustic Doppler measurements. The first model (the single sensor model) is however the same as in [4], although derived under different assumptions. The second model (the sensor network model) we have not seen elsewhere. The models explicitly account for the retardation effect. It is described how the model parameters are efficiently identified by separable nonlinear least squares optimization, and proposed how to initialize the numerical search. The models are validated on aircraft passages and also demonstrated on an all-terrain vehicle.

The identification is formulated as a batch processing problem, which means that the source has more or less passed the sensors before the parameters are estimated. This is suitable for implementation on lightweight platforms (wireless sensor nodes) in applications where a delay of a few seconds is acceptable. An application could be a ground sensor network that from a limited set of alternatives decides which way passing vehicles take. Another application is a sensor that can determine the distance to passing aircraft [3], [8] or rockets [4]. Also passive sonar has been mentioned in the literature. Moreover, we presented a tracking application for multiple ground vehicles with a probability hypothesis density (PHD) filter using Doppler measurements in a acoustic sensor network, see [10]. For more references on Doppler-based tracking we refer to that work.

The paper is organized as follows. Section II; Model descriptions and general assumptions. Section III; Estimation framework with initial estimate, separable least squares, and briefly about frequency estimation and sensor network aspects. Section IV; Model evaluation on data from an aircraft and an all-terrain vehicle. Section V; Conclusions and future work.

## II. Signal Model

The Doppler model is based on the source-observer *propagation delay*, $t_c(t)$, which leads to a comprehensive description of the relation between source motion and observed Doppler shift, both in the time and frequency domain. By our definition, the propagation delay is with respect to the observed signal $y(t)$, and it is assumed that it satisfies

$$t_c(t) = \mathrm{sol}_{\tau>0}\{\tau = \tfrac{1}{c} z(t-\tau)\}. \tag{1}$$

Here, $c$ is the assumed known speed of sound and $z(t)$ is the source-observer distance. The solution (solubility) of this equation depends on source motion (how the motion is modeled).

### A. Doppler Model

A moving sinusoidal source with constant frequency $f$ is assumed to give the signal observation

$$y(t) = h(t) + e(t), \tag{2a}$$
$$h(t) \sim \cos[2\pi f \cdot (t - t_c(t)) + \phi], \tag{2b}$$

where $e(t)$ is a background noise process and $\phi$ is the constant phase. The relation of proportionality ($\sim$) emphasizes that the amplitude of the sinusoid is disregarded in this work. The observed frequency follows as the time derivative of the cosine argument,

$$y_f(t) = h_f(t) + \epsilon(t), \tag{3a}$$
$$h_f(t) = f \cdot \left(1 - \dot{t}_c(t)\right), \tag{3b}$$

where dot notation, "$\dot{t}_c(t)$", is used to denote time derivative. Consider for a moment the special case with strict radial motion, $z(t) = vt$. Then (1) is solved by $t_c(t) = vt/(c+v)$, and $\dot{t}_c(t) = v/(c+v)$. In this special case, (3b) gives the classical expression for Doppler shift. The error $\epsilon(t)$ depends on the signal-to-noise ratio (SNR) at the observer. Error analysis and the exploitation of error characteristics are parts of ongoing research, an we aim to report interesting results in publications to come.

### B. Single Sensor Linear Motion Model

As described, the estimation of the source motion is based on parameterized motion models. How flexible these models are allowed to be, depends on the number of observers or sensors. For instance, with only a single sensor, the motion is here restricted to constant speed along a straight line, see Fig. 1. Needless to say, violations to the model assumption may result in various levels of estimation error (bias).

The unknown model parameters are the emitted frequency $f$, the normalized speed $v$, the passage distance $d$, and the passage instant $t_0$. The model parameterization uses the normalized speed $\bar{v} = v/c$ and normalized distance $\bar{d} = d/c$. The set of unknown parameters is collected in the vector

$$\theta = \begin{bmatrix} f & \bar{v} & \bar{d} & t_0 \end{bmatrix}^T. \tag{4}$$

From Fig. 1 it is evident that the source-target distance is

$$z(t) = \sqrt{d^2 + (t-t_0)^2 v^2} = c\sqrt{\bar{d}^2 + (t-t_0)^2 \bar{v}^2} \tag{5a}$$

and the solution to (1) gives the propagation delay with accompanied time derivative as

$$t_c(t) = \frac{\sqrt{\bar{d}^2 + \bar{v}^2(t-t_0)^2 - \bar{d}^2\bar{v}^2} - \bar{v}^2(t-t_0)}{1-\bar{v}^2}, \tag{5b}$$
$$\dot{t}_c(t) = \frac{1}{1-\bar{v}^2}\left[\frac{\bar{v}^2(t-t_0)}{\sqrt{\bar{d}^2 + \bar{v}^2(t-t_0)^2 - \bar{d}^2\bar{v}^2}} - \bar{v}^2\right], \tag{5c}$$

cf. [4].

### C. Sensor Network Circular Motion Model

With a sensor network with three or more sensors, more measurements are of course available, and a more flexible (general) motion model where the source follows a circular path with arbitrary radius may be proposed. The inverse radius $g = 1/r$ is one of the estimation parameters, and $g = 0$ is thus equivalent to the special case of motion along a straight line. By assuming that the source and sensors are all restricted to the same plane, there are six unknown parameters,

$$\theta = \begin{bmatrix} f & v & \alpha_0 & p_x & p_y & g \end{bmatrix}^T, \tag{6}$$

where furthermore $\alpha_0$ is the angle on the circular path and $p_0 = [p_x\ p_y]^T$ is the position, both defined at the time $t = 0$, see Fig. 2. The plane restriction may be suitable for

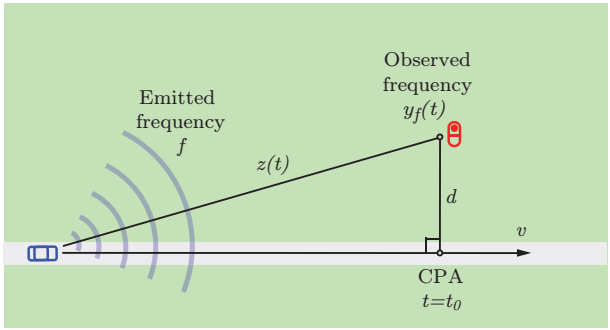


Fig. 1: Single sensor linear motion model. It is assumed that the source, here illustrated by a car on a road, moves along a straight line at constant speed $v$ emitting a constant frequency $f$. At $t = t_0$ the source attains its closest point of approach (CPA) with respect to the observer, for instance the microphone. $z(t)$ is the source-observer distance. The object is to determine $d$ and $v$ by observing the Doppler variations in the received frequency $y_f(t)$.

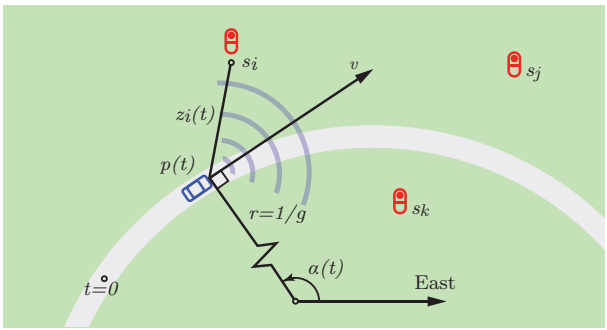


Fig. 2: Sensor network circular motion model. It is assumed that the source follows a circular trajectory $p(t)$ with radius $r = 1/g$, where $g$ is a model parameter that may take any value on the real axis. Model parameters are also the initial source location $p_0 = p(0) = [p_x\ p_y]^T$ and origin angle $\alpha_0 = \alpha(0)$.

the application of ground sensors for ground vehicle monitoring. Airborne sources would call for a three-dimensional model with eight parameters and minor model generalizations obvious to the practitioner. However, in this work we focus on the plane model. Define for notational convenience the rotated projection and orthogonal operator as

$$R(\alpha) \triangleq \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}, \quad R_\perp(\alpha) \triangleq \begin{bmatrix} -\sin\alpha \\ \cos\alpha \end{bmatrix}. \tag{7}$$

By fundamental trigonometric arguments, the source location is then

$$p(t) = p_0 - \tfrac{1}{g}R(\alpha_0) + \tfrac{1}{g}R(\alpha_0 + tgv). \tag{8}$$

Note that an equivalent form, although less intuitive, that avoids potential numerical problems in the limit $g \to 0$ (movement along straight line) is

$$p(t) = p_0 - tv \begin{bmatrix} \cos\alpha_0 \sin\frac{tgv}{2} \operatorname{sinc}\frac{tgv}{2\pi} + \sin\alpha_0 \operatorname{sinc}\frac{tgv}{\pi} \\ \sin\alpha_0 \sin\frac{tgv}{2} \operatorname{sinc}\frac{tgv}{2\pi} - \cos\alpha_0 \operatorname{sinc}\frac{tgv}{\pi} \end{bmatrix}.$$

The distance and relative speed between the source and the $i$th sensor have the forms

$$z_i(t) = \|p(t) - s_i\|, \tag{9a}$$

$$\dot{z}_i(t) = v \cdot \frac{(p(t) - s_i)^T R_\perp(\alpha_0 + tgv)}{\|p(t) - s_i\|}, \tag{9b}$$

where $s_i$ is the sensor coordinate. As with the straight line case, the propagation delay $t_{c,i}(t)$ is implicitly given by (1). In contrast, however, it here appears difficult to express $t_{c,i}(t)$ on a closed form, so it is proposed that a numerical solution is pursued using the Newton-Raphson method. Note first that $t_{c,i}(t) \approx \frac{1}{c}z_i(t)$ is an approximate solution that serves well to initialize the Newton-Raphson iterations. Define

$$F(t_c) \triangleq t_c^2 c^2 - z^2(t - t_c), \tag{10}$$

where the sensor index $i$ has been omitted to moderate the notational complexity. The positive solution to $F(t_c) = 0$ is no doubt equivalent to (1). The Newton-Raphson implementation would then be given by

$$t_c^0 = \tfrac{1}{c}z(t), \tag{11a}$$

$$t_c^+ = t_c - \frac{F(t_c)}{F'(t_c)} = t_c - \frac{t_c^2 c^2 - z^2(t - t_c)}{2t_c c^2 - 2v(p(t - t_c) - s_i)^T R_\perp(\alpha_0 + (t - t_c)gv)}. \tag{11b}$$

Convergence to machine precision is typically attained in less than 10 iterations. Once given $t_c(t)$, the chain rule gives in turn

$$\dot{t}_c(t) = \frac{\dot{z}(t - t_c(t))}{c + \dot{z}(t - t_c(t))}. \tag{12}$$

Assuming motion along a straight line ($g \equiv 0$) features a reduced parameter space and also other possible simplifications. Each sensor $i$ can then be recast to the single sensor framework where (5) gives the propagation delay on a closed form, which in turn gives computational advantages. The details of deriving $d_i$ and $t_{0,i}$ from (8) for this special case are however not given here.

## III. Least Squares Estimation

The (unweighted) least squares estimate of (4) or (6) is generally expressed

$$\hat{\theta} = \arg\min_\theta \sum_{i=1}^{n} \left[y_f(t_i) - h_f(t_i;\theta)\right]^2. \tag{13}$$

Here, a time discrete representation is used, where a set of $i = 1, 2, \ldots, n$ frequency measurements are given at distinct time instances, $t_i$. To our knowledge, (13) can for the current application only be solved by iterative numerical algorithms. It is, however, observed that (13) is not generally free from local minima, a fact that complicates the numerical approach significantly. In this section we propose both a robust initial estimate and review numerical algorithms tailored for the nonlinear least squares problem at hand.

### A. Initial Estimate

Robust initial parameter values are here proposed for the single sensor model (5), which assumes motion along a straight line. Initial values for the sensor network model are not explicitly developed here, but it is proposed that the sensor network parameters (6) are fitted to a set of single sensor estimates as already described in [8], [2]. The linear motion assumption may, if necessary, be enforced by setting $g \equiv 0$. http://www.vinnova.se/sv/Ansoka-och-rapportera/Utlysningar/

Thus, assume the single sensor model (5) and define the asymptotic frequencies as

$$f_a = \lim_{t\to-\infty} f(1-\dot{t}_c(t)) = \frac{f}{1-\frac{v}{c}}, \tag{14a}$$

$$f_b = \lim_{t\to\infty} f(1-\dot{t}_c(t)) = \frac{f}{1+\frac{v}{c}}, \tag{14b}$$

see Fig. 3. Then

$$c\cdot\frac{f_a-f_b}{f_a+f_b} = c\cdot\frac{\frac{f_a}{f_b}-1}{\frac{f_a}{f_b}+1} = c\cdot\frac{\frac{c+v}{c-v}-1}{\frac{c+v}{c-v}+1} = v, \text{ and} \tag{15}$$

$$\frac{f_a+f_b}{2} = \frac{0.5f}{1-\frac{v}{c}} + \frac{0.5f}{1+\frac{v}{c}} = \frac{f}{1-(\frac{v}{c})^2}. \tag{16}$$

Also define the observed frequency derivative extremum

$$f_o^* = \min_t \dot{h}_f(t) = \min_t -f\ddot{t}_c(t)$$

$$\min_t -\frac{fd^2v^2}{c(t^2v^2+d^2(1-v^2/c^2))^{3/2}} \approx \frac{-fv^2}{cd}, \tag{17}$$

with reference to (3b) and (5c). If $f_a$, $f_b$, and $f_o^*$ can be directly extracted from $y(t)$, for instance, as the mean of the first and last few samples, and as the slope of a line fit around $t_0$, then the following entities give a suboptimal parameter estimate to be used for initializing a numeric search;

$$v^0 = c\cdot\frac{f_a-f_b}{f_a+f_b}, \tag{18a}$$

$$f^0 = \frac{2f_af_b}{f_a+f_b}, \tag{18b}$$

$$d^0 = \frac{-f^0(v^0)^2}{cf_o^*}. \tag{18c}$$

A start solution for $t_0$ could be obtained by extracting the zero crossing of $y(t)-f^0$, for instance

$$t_0^0 = -\frac{d^0}{c} + \max_\tau \sum_{t=1}^{\tau}[y_f(t)-f^0] - \sum_{\tau+1}^{n}[y_f(t)-f^0]. \tag{18d}$$

The first term $d^0/c$ accounts for the propagation delay at the closest point of approach (CPA).

### B. Separable Least Squares

Since the unknown frequency parameter $f$ enters the model equation linearly ($f$ is separable), it can be replaced by its implicit least squares solution, which reduces the parameter dimension and facilitates a more efficient nonlinear optimization step. Reformulate (3a) as

$$y_f(t) = \tilde{h}_f(t;\nu)f + \epsilon(t) \tag{19}$$

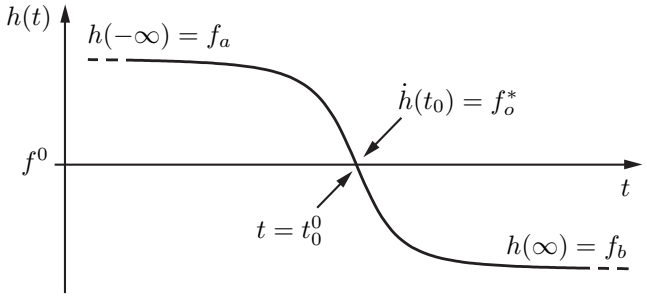


Fig. 3: Features extracted for the initial estimate for the single sensor linear motion model. The key features are the asymptotic frequencies, $f_a$, $f_b$, and the slope of the frequency curve at the closest point of approach, $f_o^*$.

or, on time discrete vector form as

$$Y_f = Gf + E, \tag{20}$$

where rows are time instances ($t_i$), $G = G(\nu) \in \mathcal{R}^n$, and $\nu \in \mathcal{R}^p$ is the part of $\theta$ that enters the model equation nonlinearly (all parameters but the frequency, $f$). As before, $\theta$ comprises the unknown parameters, for instance (4) or (6). Given any hypothesis on $\nu$, the (unweighted) least squares estimate of $f$ is

$$\hat{f} = (G^TG)^{-1}G^TY_f, \tag{21}$$

and inserted into (20) it gives

$$Y_f = G(G^TG)^{-1}G^TY_f + E. \tag{22}$$

The least squares estimate of $\nu$ may now be pursued by minimizing

$$||(I-G(G^TG)^{-1}G^T)Y_f||_2^2. \tag{23}$$

General optimization software can be used for this, for instance the Nelder-Mead minimization algorithm implemented in the fminsearch function in MATLAB.

However, if efficiency is of any importance, attention should instead be given the more specialized Gauss-Newton algorithm, which has well-documented merits as a least squares problem solver, see [11, pp. 259]. For the separable case at hand, the forms of both the objective function and its Jacobian are, notably, different compared to the case with ordinary nonlinear least squares. In [12] it is proposed that (23) is (equivalently) formed as

$$||P(\nu)Y_f||_2^2, \tag{24}$$

where $P = P(\nu)$ is orthogonal and spans exactly the nullspace of $G$. $P$ is computed by a QR factorization at each Gauss-Newton step. In [12] it is also proposed how to compute the Jacobian $J = \nabla(PY_f) \in \mathcal{R}^{n\times p}$ required by the Gauss-Newton algorithm. This variation of the algorithm is termed *variable projection* (Varpro). In the special case at hand (with only one separable parameter), the Jacobian is well approximated by

$$\nabla(PY_f) \approx -P\nabla GG^-Y_f, \tag{25}$$

where $G^-$ is a symmetric generalized inverse of $G$. This computationally efficient approximation was proposed in [13]. Computing $\nabla G$ in turn, is (in the application at hand) the

TABLE I: Comparison of computational efficiency between the Nelder-Mead (fminsearch in MATLAB) and Varpro (Gauss-Newton) based on the processing of 28 aircraft passages described in Section IV-A. The function tolerance is the algorithm stop criteria, namely the maximum iterative improvement in objective function value. The numerical convergence is comparable but the tailored Gauss-Newton is apparently about eight times faster.

| Func. Tol. [Hz] | CPU time Varpro [ms] | CPU time Nelder-Mead [ms] | RMSE Varpro [Hz] | RMSE Nelder-Mead [Hz] |
|---|---|---|---|---|
| $10^{-4}$ | 1.0 | 8.0 | 0.23 | 0.23 |
| $10^{-7}$ | 1.4 | 11 | 0.23 | 0.23 |

most expensive part of the algorithm, regardless if it is done analytically or by finite differences. Consider, for example, the single sensor linear motion model (5c) combined with (3) and with the approximation $t_c(t) \approx z(t)/c$. The $i$th row of $\nabla G \in \mathcal{R}^{n\times p}$, $p$ being the number of nonlinear parameters, is the vector with partial derivatives,

$$\nabla_\nu G_i = \begin{bmatrix} \frac{\partial h_f}{f\partial v} \\ \frac{\partial h_f}{f\partial d} \\ \frac{\partial h_f}{f\partial t_0} \end{bmatrix} = \begin{bmatrix} \frac{-v(t_i-t_0)(2d^2+v^2(t_i-t_o)^2)}{c(d^2+v^2(t_i-t_0)^2)^{3/2}} \\ \frac{dv^2(t_i-t_0)}{c(d^2+v^2(t_i-t_0)^2)^{3/2}} \\ \frac{d^2v^2}{c(d^2+v^2(t_i-t_0)^2)^{3/2}} \end{bmatrix}, \qquad (26)$$

which is far more complex than the objective function which requires only little more computing than (5a). The many computations needed at each step of the Gauss-Newton algorithm however pay off. To get an indication of the performance, we compare the Nelder-Mead with the Varpro applied to 28 aircraft passages, see Section IV-A. Varpro is implemented in MATLAB based on the description in [14], and Nelder-Mead is the fminsearch in MATLAB. The mean CPU convergence time and the overall root mean square error is given in Table I. For this simple example, the specialized Varpro is about eight times faster. Varpro typically converges in 3 to 5 iterations (from the initial estimate described earlier), which takes around 1 ms on a Dell Latitude E6410 laptop with an Intel Core i7 M620 CPU at 2.67 GHz. This efficiency indicates in turn that it would be feasible to tailor the program for a lightweight platform like a wireless sensor network node.

### C. Frequency Estimation

In this work, frequency estimation and motion estimation are separate processing steps. The frequency estimation is here conducted in disjoint time windows centered around the time instants of the frequency samples, $t_i$, and is in essence a preprocessing step prior to the motion estimation. The division into separate steps is of course suboptimal; from estimation theoretical point of view, it is probably better to test a motion hypothesis directly against the raw data, possibly in the time domain. This one-step approach is however more complex and it is left to future work to find out how to set it up. In the frequency domain it is much easier to ignore an unknown phase and also acoustic wave distortions. It is left to future work to find out. A drawback with the one-step approach is that it would be more difficult to distribute to sensor network nodes, see the following section.

This work is focused on the motion estimation step, and only little attention will here be given to the frequency estimation. As mentioned, the frequency $y_f$ is estimated in disjoint time windows of the raw measurement signal $y$. The time windows in the current application are typically around half a second and the raw sampling rate a few kHz. Since the acoustic sources in this study are vehicles and aircraft with revolving engines or propellers that typically give periodic sound, the original model assumption is extended with harmonic components,

$$y(t) = \sum_{r=1}^{k} [A_r \cos(2\pi r y_f t + \phi_r)] + e(t), \qquad (27)$$

where $k-1$ is the number of harmonics. $A_r$, $\phi_r$, and $y_f$ are all unknown parameters, although $y_f$ is the only one used in the proposed Doppler/motion models. The random noise $e(t)$ is assumed to be Gaussian and to have a uniform spectral distribution up to the Nyquist frequency, and no power above. An assumption here is also that the frequency variation within a time window is negligible. $y_f$ is estimated by a separable nonlinear least squares technique (similar to the one described in the preceding section). Under the assumption (27), the least squares estimate is also maximum likelihood, see [15]. We propose that the numerical search is initialized by

$$\hat{y}_f^0(t_i) = \underset{f_1<u_j<f_2}{\operatorname{argmax}} \sum_{r=1}^{k} \sum_{l} w_l |F_i(u_{rj+l})|, \qquad (28)$$

where $F_i(u_j)$ is the discrete-time Fourier transform of $y$ in a time window centered at $t_i$ and at the frequency $u_j$. $w$ is a frequency window ($[w_{-1} \quad w_0 \quad w_1] = [0.5 \quad 1 \quad 0.5]$ has been used in the numerical experiments). $f_1$ and $f_2$ restrict the search to the region where the fundamental frequency has been detected. For the benefit of certain real-time applications it is worth noting that the initial estimate $\hat{y}_f^0$ is probably so good in average that the subsequent numerical search can be omitted. Omitting the second search this way reduces the computational requirements by orders of magnitude. If this is the way chosen, some technique to increase the frequency resolution of the Fourier transform is recommended, zero padding for instance.

Another suboptimal but fast method to estimate the fundamental frequency of a harmonic series is given in [16], where an initial estimate of separate sinusoids gives the fundamental frequency on closed form. We however failed to make that approach match the robustness of (28).

An estimator that relies on multiple harmonic components is robust to a partial loss (or drop in the signal-to-noise ratio) of the frequency spectrum, even to a temporary black-out of the fundamental component. Future investigations will reveal if there are computationally efficient ways to combine the merits of the harmonic estimator with an instantaneous frequency estimator, where the assumption on negligible frequency variation within a time window is relieved, see for instance [17], [18], [19] and the references therein.

### D. Network Distributed Estimation

Although the estimation step proposed here is inherently centralized, the preprocessing frequency estimation step can, without loss of performance, be decentralized to the sensor

network nodes, provided of course that the nodes have sufficient processing power. This reduces the required network bandwidth by multiple orders of magnitude, since only the estimated frequency then needs to be transmitted, not the raw measurements.

Suboptimal approaches of distribution may be interesting to study, although we will not pursue those here. For instance, the bandwidth requirement could be even further reduced by having each node running the proposed single sensor algorithm, and send the locally estimated passage distance, speed and passage time to the fusion node. Then the fusion node may, speculatively, traverse geo-temporal hypotheses to make statements about the source trajectory, see [8], [2].

## IV. Numerical Experiments

### A. Aircraft Localization

A single microphone is deployed close to the ground. A Mooney M20J propeller-driven aircraft is flown back and forth in different directions nominally vertical above the microphone. Various flight levels produce a data set with 28 passage distances in the range 120 through 800 m. The nominal speed is 75 m/s. The aircraft carries a Topcon GRS-1 advanced global navigation satellite system (GNSS) receiver using base station corrections, by which the aircraft is tracked with sub-meter accuracy. Also, synchronized audio recordings close to the aircraft engine were conducted for reference purposes. The background noise level and wind are moderate. The speed of sound parameter is set to $c = 335$ m/s. This data set is used to evaluate the single sensor linear motion model.

The frequency is estimated in disjoint 0.5 s square windows using $k - 1 = 3$ harmonics. The sampling rate is (down-sampled to) 2 kHz. The speed $v$ and closest point of approach ($d$, $t_0$) are computed from the GNSS receiver for comparison. The proposed initial estimate is used. The least squares optimization is executed by the Varpro Gauss-Newton method.

Fig. 4 shows a typical acoustic spectrogram for an aircraft passage 220 m from the microphone. The spectrogram is focused on the fundamental frequency, over which the Doppler frequency identified through the model (3) combined with (5) is layed (white dashed line). By visual inspection, the model fits the measurement well.

The diagram in Fig. 5 gives the distance estimate errors, $\hat{d} - d$, as a function of the passage distance, $d$. Table II gives in turn the root mean square errors for all the estimated parameters. The distance errors tend to increase with increased passage distance, while this is not the case with the speed. Also the passage time estimate is stable, and surprisingly accurate. That the distance error is the parameter that is most prone to estimation errors is understood by the fact that it is basically identified by the frequency derivative, which in turn is particularly susceptible to measurement noise.

The model identification gives anticipated parameter estimates. An interesting observation is that with the approximation $t_c \approx \frac{1}{c} z$ commonly used in radar equations (that is, retardation disregarded), which is based on the assumption $v \ll c$, the root mean square distance error increases from 35 to 39 m (11%).

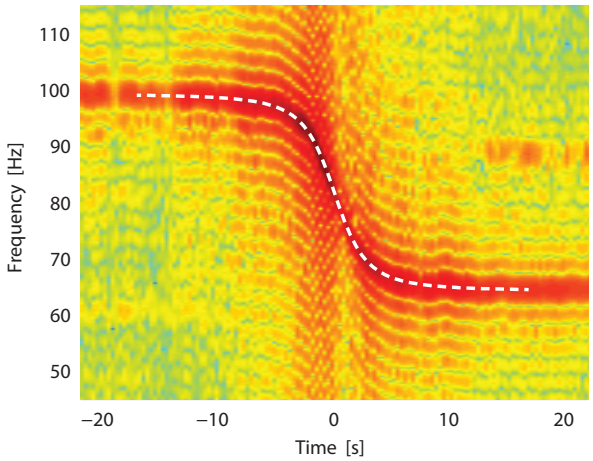


Fig. 4: Acoustic spectrogram of the Mooney M20J passing 220 m from the sensor at 73 m/s. The white dashed line is the frequency according to the identified model (3) combined with (5). $t = 0$ is the true passage time (closest point of approach) according to the GNSS receiver.

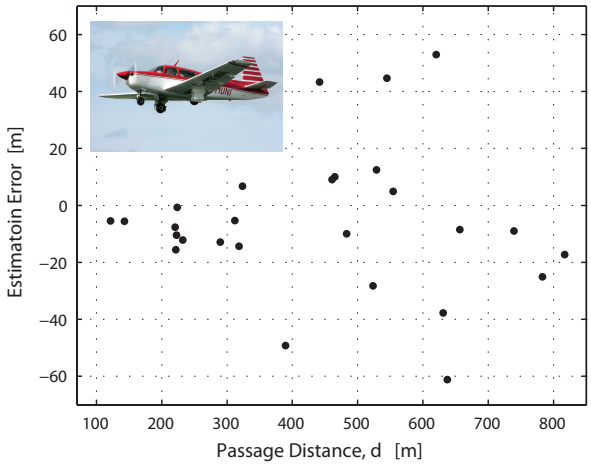


Fig. 5: Distance error versus distance for 28 passages of a Mooney M20J aircraft.

### B. ATV Localization

Three microphones are deployed close to the ground (M1, M2, M3). An all-terrain vehicle passes this microphone network, first on a straight path, then on a curved path. Also this vehicle carries the accurate GNSS receiver giving reliable reference tracks. The nominal speed is 14 m/s. The speed of sound parameter is set to $c = 335$ m/s. This data set is used to demonstrate the sensor network circular motion model. Unfortunately it does not contain enough data to make any statistical analysis of the results—this is left to future experiments.

The frequency is estimated in disjoint 0.33 s square windows using $k - 1 = 3$ harmonics. The sampling rate is (down-sampled to) 2 kHz. An initial estimate is obtained by combining parameter estimates from the single sensor model applied to each sensor using geometrical arguments. As before,

TABLE II: Root mean square estimation errors (RMSE) for 28 aircraft passages. The parameters are the passage distance $d$, the speed $v$, and the passage instant (closest point of approach) $t_0$. The speed estimate is fairly independent of the passage distance. As seen in Fig. 5 this is not the case with the distance error which appears to increase with distance.

| Parameter | $d$ [m] | $v$ [m/s] | $t_0$ [s] |
|---|---|---|---|
| RMSE | 35 | 1.2 | 0.18 |

TABLE III: True parameter values versus estimated for the all-terrain vehicle data set. The true values are the least squares fit to the GNSS reference track. RMSE is the value of the least squares objective function for the true parameter, and for the estimated parameter, respectively.

| | $v$ [m/s] | $\alpha_0$ [°] | $p_x$ [m] | $p_y$ [m] | $1/g$ [m] | RMSE [Hz] |
|---|---|---|---|---|---|---|
| Fig. 6, true | 14.6 | 139 | 255 | 811 | 1925 | 0.27 |
| Fig. 6, estimated | 15.8 | 138 | 252 | 811 | -2018 | 0.14 |
| Fig. 7, true | 13.8 | 129 | 249 | 784 | 86 | 0.32 |
| Fig. 7, estimated | 14.8 | 132 | 252 | 786 | 85 | 0.27 |

the least squares optimization is executed using the Varpro Gauss-Newton method.

Fig. 6 and 7 give a visual comparison between estimated trajectory and reference track for the straight and curved passage, respectively. A numeric comparison of parameter values is given in Table III.

The parameter errors are possibly acceptable, but yet significant. The errors are derived to rather large noise in the frequency measurements, see Fig. 8. It is noted that the data set is not ideal, since there are significant fluctuations in the engine frequency that may be interpreted as measurement noise.

### C. Summary of Experimental Results

The impression of the aircraft experiment is that the estimation errors are reasonably low, which would be an indication of model validity. The errors are also comparable to the experimental error levels found in [3]. The observed increase in estimation error (bias) when disregarding the retardation effect further indicates that this approximation should be used cautiously. The all-terrain vehicle experiment is, due to the little data set, more a demonstration than it is a statistical validation. Anyway, it indicates that the sensor network model holds, and that it would be worth wile to collect more data for a thorough evaluation.

## V. Conclusion

Two motion models for localization of acoustic sources based on observed Doppler shift have been described. The models are accurate in the sense that they fully account for the retardation effect. The first model is tailored for a single sensor and assumes constant source speed along a straight line. The second model uses multiple sensors in a sensor network and allows a more flexible source trajectory. A state-of-the-art nonlinear least squares estimator based on Gauss-Newton variable projection techniques has been described,

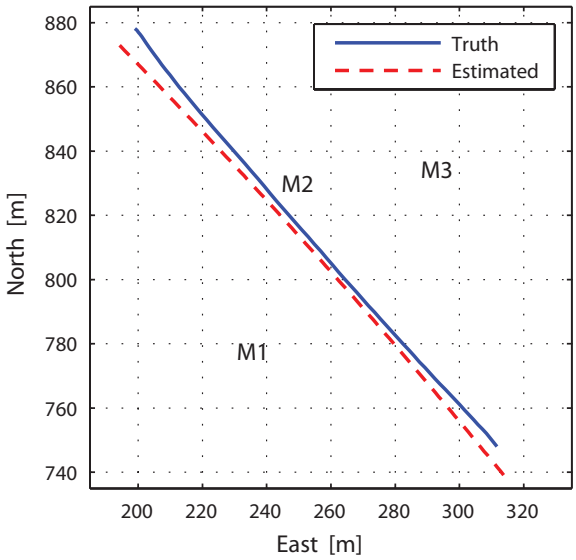


Fig. 6: Trajectory estimate of the all-terrain vehicle using the sensor network circular motion model (8); straight trajectory. M1, M2, M3 are the microphone positions. The estimate is compared to the ground truth.

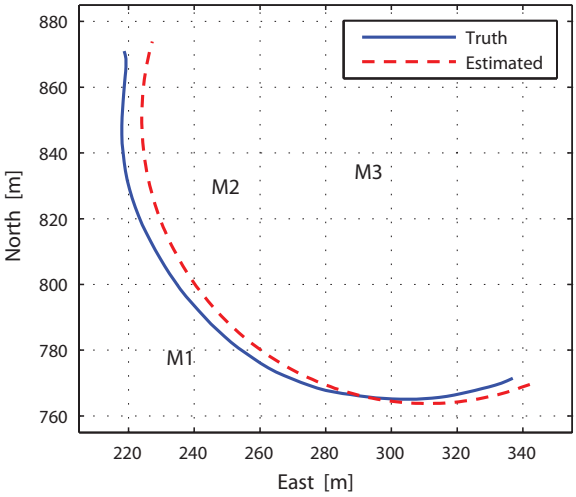


Fig. 7: Trajectory estimate of the all-terrain vehicle using the sensor network circular motion model (8); curved trajectory. M1, M2, M3 are the microphone positions. The estimate is compared to the ground truth.

which demonstrates that at least a local optimum can be reach within a few milliseconds using an office laptop from 2009. Also a robust initial estimate is proposed that uses a few simple features from the observed frequency series. A frequency estimator based on harmonic components has been briefly described. Some brief notes on algorithm network distribution are also given. The models have been demonstrated on real data from a passing aircraft and on an all-terrain vehicle.

The conclusion is that the models give accurate estimates

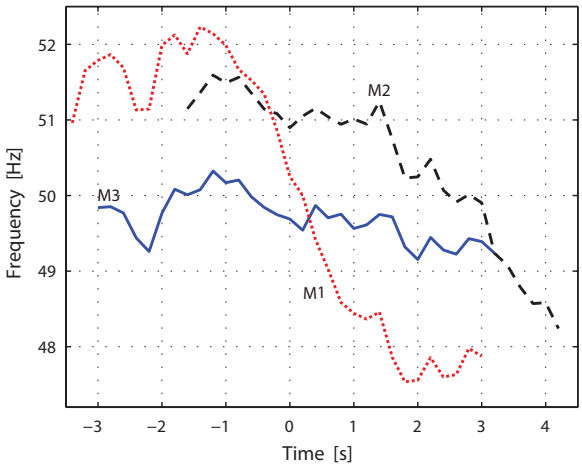


Fig. 8: Frequency estimates used as input to the model identification in Fig. 7. The fundamental source frequency is apparently around 50 Hz. The Doppler shift is clearly manifested, although the noise level is high with respect to the shift.

provided that the assumptions hold (constant speed and constant emitted frequency), and that there are computationally efficient estimators that make the algorithm practically useful, possibly even on lightweight platforms such as sensor network nodes. The techniques involved are fairly mature for practical implementations in a sensor network.

Future work will include analytical and simulated error analysis, acquisition of more data for evaluation, and also investigation of efficient instantaneous frequency estimation of harmonic frequency components.

## Acknowledgment

This work is part of the project Collaborative Localization at Linköping University, which is funded by the Swedish Foundation for Strategic Research (SSF) and also part of the Centre for Advanced Sensors, Multisensors and Sensor Networks, FOCUS, at the Swedish Defence Research Agency, FOI, funded by Swedish Governmental Agency for Innovation Systems (VINNOVA). We gratefully acknowledge the support from SSF and VINNOVA which made this work possible.

## References

[1] R. Webster, "An exact trajectory solution from doppler shift measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-18, no. 2, pp. 249–252, march 1982.

[2] Y. Chan and J. Towers, "Sequential localization of a radiating source by doppler-shifted frequency measurements," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 28, no. 4, pp. 1084–1090, oct 1992.

[3] B. G. Ferguson, "A ground-based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller-driven aircraft," *Journal of Acoustical Society of America*, vol. 92, no. 3, pp. 1403–1407, 1991.

[4] J. Statman and E. Rodemich, "Parameter estimation based on doppler frequency shifts," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-23, no. 1, pp. 31–39, jan. 1987.

[5] M. A. Richards, *Fundamentals of radar Signal Processing*. McGraw-Hill, 2005.

[6] Y. Chan, "A 1-D search solution for localization from frequency measurements," *Oceanic Engineering, IEEE Journal of*, vol. 19, no. 3, pp. 431–437, jul 1994.

[7] E. Weinstein and N. Levanon, "Passive array tracking of a continuous wave transmitting projectile," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-16, no. 5, pp. 721–726, sept. 1980.

[8] D. Torney, "Localization and observability of aircraft via doppler shifts," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 3, pp. 1163–1168, july 2007.

[9] I. Shames, A. Bishop, M. Smith, and B. Anderson, "Analysis of target velocity and position estimation via doppler-shift measurements," in *2011 Australian Control Conference*, Melbourne/Australia, November 2011, pp. 507–512.

[10] M. B. Guldogan, D. Lindgren, F. Gustafsson, H. Habberstad, and U. Orguner, "Multiple target tracking with Gaussian mixture PHD filter using passive acoustic Doppler-only measurements," in *Information Fusion (FUSION), 2012 15th International Conference on*. IEEE, 2012, pp. 2600–2607.

[11] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer, 1999.

[12] G. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on Numerical Analysis*, vol. 10, no. 2, pp. 413–432, 1973.

[13] L. Kaufman, "A variable projection method for solving separable nonlinear least squares problems," *BIT Numerical Mathematics*, vol. 15, pp. 49–57, 1975.

[14] D. Gay and L. Kaufman, "Tradeoffs in algorithms for separable nonlinear least squares," in *Proc. Congress on Computations in Applied mathematics*, R. Vichnevetsky and J. J. H. Miller, Eds., Dublin, Ireland, 1991, pp. 157–8.

[15] P. Stoica and R. Moses, *Introduction to spectral analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.

[16] H. Li, P. Stoica, and J. Li, "Computationally efficient parameter estimation for harmonic sinusoidal signals," *Signal Processing*, no. 80, pp. 1937–1944, 2000.

[17] I. Djurović and L. Stanković, "An algorithm for the Wigner distribution based instantaneous frequency estimation in a high noise environment," *Signal Processing*, vol. 84, no. 3, pp. 631–643, 2004.

[18] B. G. Ferguson and B. G. Quinn, "Application of the short-time fourier transform and the Wigner–Ville distribution to the acoustic localization of aircraft," *Journal of Acoustical Society of America*, vol. 96, no. 2, pp. 821–827, 1994.

[19] D. C. Reid, A. M. Zoubir, and B. Boashash, "Aircraft flight parameter estimation based on passive acoustic techniques using the polynomial Wigner–Ville distribution," *Journal of the Acoustical Society of America*, vol. 102, no. 1, pp. 207–223, 1997.