18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Direction of Arrival Estimation in Sensor Arrays Using Local Series Expansion of the Received Signal

Fredrik Gustafsson*, Gustaf Hendeby*†, David Lindgren*†, George Mathai*, and Hans Habberstad†
*Dept. Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden.
Email: {fredrik, hendeby, david, geoma76}@isy.liu.se
†Dept. of Sensor & EW Systems, Swedish Defence Research Agency (FOI), SE-581 11, Linköping, Sweden.
Email: firstname.lastname@foi.se

***Abstract*—A local series expansion of a received signal is proposed for computing *direction of arrival* (DOA) in sensor arrays. The advantages compared to classical DOA estimation methods include general sensor configurations, ultra-slow sampling, small dimension of the arrays, and that it applies for both narrowband and wideband signals without prior knowledge of the signals. This makes the method well suited for DOA estimation in sensor networks where size and energy consumption have to be small. We generalize the common far-field assumption of the target to also include the near-field, which enables target tracking using a network of sensor arrays in one framework.**

I. Introduction

Direction of arrival (DOA) estimation from an array of sensors is a well studied problem in literature, see for instance [1, 6, 7], where there is a range of algorithms such as MUSIC, ESPRIT, deterministic and stochastic *maximum likelihood* (ML), etc. These are based on an assumption of narrowband and far-field signals, and the algorithms are often tailored for regular arrays (uniformly linear or circular arrays). The underlying principle is to estimate the delay of the signal between the sensors, and this delay corresponds to the DOA. Further, the rule of thumb is that the sensors in the array should be separated about half the wavelength of the received signal. Typical applications are in radio and sonar where the signal is modulated on a carrier, thus satisfying the narrowband assumption. The weak point in this classical theory is how to estimate the number of sources. Recently, significant progress has been made in tuning-free methods that also estimated the number of sources such as SPICE, see [8].

In contrast to the literature above, we consider the case of opportunistic (acoustic) signal sources, which are typically wideband. Examples include tracking motorized vehicles outdoors or talking people in a conference room. One solution is to focus the signal into a narrowband signal, for instance WINGS, see [2, 4], and thereby recasting the problem into a standard one.

We propose a new concept based on space delays rather than time delays. The derivation is based on a Taylor expansion of the signal, where the delays over space are related to a reference point through the Taylor expansion. The difference compared to classical DOA methods include:

- The time sampling can be arbitrary. Also very slow sampling, say every second or minute, makes sense in for instance sensor networks which are energy constrained.
- The space sampling, that is, separation between the sensors, must be closer than half the wavelength, preferably less than one tenth of the wavelength. This can be an advantage in sensor networks, where the sensor units have to be small. There is no lower limit on the separation in theory, but in practice the *signal-to-noise ratio* SNR determines a lower bound.
- The method applies to both wideband as well as to narrowband signals.
- The array can have arbitrary configuration and is not restricted to uniform arrays.
- In theory, the method applies to multiple sources where the number of sources can be estimated as well. However, it does not scale well since the number of parameters increases linearly with the number of sources, so the required number of array elements also increases linearly.
- The method can be directly parametrized in the source location, and hence it works equally well in the near-field as in the far-field.
- Thus, localization in an array network can be performed directly in one filter framework, rather than first letting each array estimate DOA and the applying triangulation in a second step.

For the above reason, the concept is very well suited for target tracking in array networks.

The following sections derive the method and substantiate the claims and properties above.

II. Far-Field DOA Estimation

A. Signal Model

Assume a plane propagation model where the sensors are located at (x_n, y_n) for $n = 1, 2, \dots, N$. In this section, we assume that the sensors are much closer to the origin of the coordinate system than the target. Here, a single array is considered.

The target emits a signal (acoustic, sonar, radio, seismic, etc.) $s(t)$, which reaches sensor n with a time shift $\tau_n(\theta)$

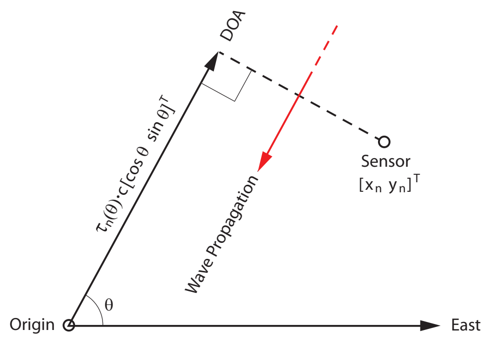


Fig. 1. Geometrical illustration of how the far-field wave DOA relates to the signal delay $\tau_n(\theta)$ in sensor n. c is the speed of wave propagation (sound speed). The projection of the sensor coordinate onto the DOA vector motivates (2). The choice of origin is not critical, but is typically defined as the middle point of the array to moderate the offset in $\tau_n(\theta)$.

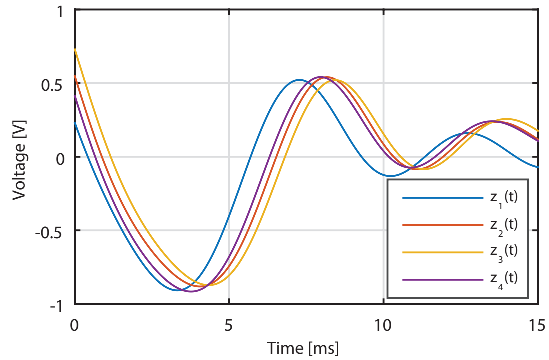


Fig. 2. Illustration of signals from the four microphone array used in the experiments, see Fig. 3. The signals are typically coherent with a detectable temporal displacement that follows from the array geometry and DOA.

relative to the origin and additive Gaussian noise $\epsilon_n(t) \sim \mathcal{N}(0, \sigma^2_{\epsilon_n})$, according to

$$z_n(t) = s\big(t + \tau_n(\theta)\big) + \epsilon_n(t), \tag{1}$$

which is one of the standard assumptions in the DOA literature. Here c is the speed of wave propagation. As illustrated in Fig. 1, the delay can be expressed as

$$\tau_n(\theta) = \frac{x_n \cos\theta + y_n \sin\theta}{c}. \tag{2}$$

Fig. 2 illustrates acoustic signals from a four microphone array.

B. Assumptions

The main assumption is that the delay differences are much smaller than the signal variation, which in terms of frequency can be stated as

$$|\tau_n(\theta) - \tau_m(\theta)| \lesssim \frac{1}{f_{\max}}, \quad \forall m, n. \tag{3}$$

Since the delay difference is bounded by the array size D, this can be stated as

$$D \lesssim \frac{c}{f_{\max}}. \tag{4}$$

As an example, a motorized vehicle has the fundamental frequency of the sound emission below 100 Hz. The shortest wavelength is thus 3 m, and the array should be much smaller than this, say 0.3 m. In our field tests, $D = 0.25$ m.

C. Taylor Expansion

If the condition in (4) is satisfied, then the signal varies smoothly over the array, and a Taylor expansion of order L catches the local behavior,

$$s(t + \tau_n(\theta)) = \sum_{i=0}^{L} \frac{\tau_n^i(\theta)}{i!} \frac{d^i}{dt^i} s(t) + \Delta_n(t). \tag{5}$$

This can be written as a linear regression

$$s(t + \tau_n(\theta)) = \mathcal{T}_n^T(\theta) S + \Delta_n(t) \tag{6}$$

where $\Delta_n(t)$ denotes the higher order terms of the Taylor expansion which in the sequel will be neglected and included in $\epsilon_n(t)$ as a consequence of the assumption (4), and

$$\mathcal{T}_n^T(\theta) = \begin{pmatrix} 1 & \tau_n(\theta) & \frac{1}{2!}\tau_n^2(\theta) & \dots & \frac{1}{L!}\tau_n^L(\theta) \end{pmatrix} \tag{7a}$$
$$S^T(t) = \begin{pmatrix} s^{(0)} & s^{(1)} & s^{(2)} & \dots & s^{(L)} \end{pmatrix} \tag{7b}$$

and $s^{(L)} = \frac{d^L}{dt^L} s(t)$. The original model (1) is thus

$$z_n(t) = \mathcal{T}_n^T(\theta) S(t) + \epsilon_n(t). \tag{8}$$

Considering the array, we thus have the following set of equations linear in the Taylor expansion parameters, but nonlinear in θ,

$$\begin{aligned} z_1(t) &= \mathcal{T}_1^T(\theta) S(t) + \epsilon_1(t) \\ z_2(t) &= \mathcal{T}_2^T(\theta) S(t) + \epsilon_2(t) \\ &\vdots \\ z_N(t) &= \mathcal{T}_N^T(\theta) S(t) + \epsilon_N(t) \end{aligned} \tag{9}$$

This can conveniently be written in matrix form as

$$z(t) = \mathcal{T}^T(\theta) S(t) + \epsilon(t) \tag{10}$$

where

$$\mathcal{T}^T(\theta) = \begin{pmatrix} \mathcal{T}_1^T(\theta) \\ \mathcal{T}_2^T(\theta) \\ \vdots \\ \mathcal{T}_N^T(\theta) \end{pmatrix}. \tag{11}$$

D. DOA Estimation

The key point with the partially linear model in (10) is that the *least squares* (LS) estimate, which coincides with the *maximum likelihood* (ML) estimate for Gaussian noise $\epsilon(t)$, is conveniently computed by searching for the optimal DOA θ, where all the linear parameters can be estimated analytically. That is, the optimization only concerns a scalar parameter, independently of the number N of array elements and the order L of the Taylor expansion.

The LS estimate is by definition

$$(\hat{\theta}, \hat{S}) = \arg\min_{\theta, S} V(\theta, S) \tag{12}$$

where V denotes the LS loss function given by

$$V(\theta, S) = \|z - \mathcal{T}^T(\theta) S\|^2 \tag{13}$$

The linear sub-structure makes the estimation problem fit the *separable least square* (SLS) framework which eventually makes solving the optimization problem more computationally efficient.

Let θ be a fixed parameter. Then the estimate of S is given by

$$\hat{S}(\theta) = \arg\min_{S} V(\theta, S) = \arg\min_{S} \|z - \mathcal{T}^T(\theta) S\|^2 \tag{14}$$

For the above optimization problem, the estimate can be computed using least squares

$$\hat{S}(\theta) = \big(\mathcal{T}^T(\theta)\big)^{\dagger} z, \tag{15}$$

were † denotes the Moore-Penrose pseudo-inverse. This leads to the following estimate of θ

$$\hat{\theta} = \arg\min_{\theta} V(\theta, \hat{S}(\theta)) = \arg\min_{\theta} \|z - \mathcal{T}^T(\theta) \hat{S}(\theta)\|^2. \tag{16}$$

If θ is one or two dimensional, then the estimate of θ can be computed quite efficiently by evaluating it over a fine grid [5].

E. Multiple Signal Sources

In principle, the linearized signal model (10) can be extended to multiple signal sources straightforwardly. Denote the number of signal sources K, then (10) becomes

$$z(t) = \sum_{k=1}^{K} \mathcal{T}^T(\theta_k) S_k(t) + \epsilon(t). \tag{17}$$

To satisfy the obvious identification criteria, the number of unknowns must necessarily be less than the number of observations, so $N \geq (L+2)K$ (K sets of Taylor expansions of order L, and one extra angle parameter to each source). That is, the number of required sensor elements in the array increases linearly with the number of signal sources.

F. Design Issues

Real signals are seldom band-limited, so the maximum frequency $f_{\max}$ is not well defined. However, a low-pass filter can always be applied to all sensors. That is, $f_{\max}$ can be seen as a design parameter, just as the size D of the array. The order L of the Taylor expansion is also free to choose by the user. In practice, the design order should be as follows:

1) D is given by array construction.
2) $f_{\max}$ is selected based on source excitation to get the best possible SNR. Still, we need to satisfy $f_{\max} \ll c/D$.
3) The Taylor order L is monotonically increasing function of $Df_{\max}$, and it has to satisfy the necessary (in general also sufficient) constraint $(L+2) \leq N$, to get a unique solution modulo $180°$.

A good first try is to start with a first order Taylor expansion $L = 1$.

III. Near-Field Localization and Tracking

In Sec. II, the delay $\tau_n(\theta)$ to each sensor was stated as a function of bearing in (2), which implicitly assumes a far-field target. This condition will here be relaxed for the purpose of localization and tracking in networks of sensor arrays. Although the purpose is array networks, the full framework for multi-node estimation is considered out of scope here. The extensions to multiple array should however be straight forward, once synchronization and propagation delay issues have been resolved.

A. Delay Model

Let the target position be (X, Y). Then the delay is given by

$$\tau_n(X, Y) = \frac{1}{c}\sqrt{(x_n - X)^2 + (y_n - Y)^2}. \tag{18}$$

Thus, for near-field signals, the signal model in (10) can be modified to

$$z(t) = \mathcal{T}^T(X, Y) S(t) + \epsilon(t). \tag{19}$$

This is merely a notational change of arguments of the regression matrix $\mathcal{T}(X, Y)$.

B. Localization

Similarly, the loss function gets the functional form $V\big(X, Y, \hat{S}(X, Y)\big)$. The two-dimensional minimization can be performed numerically over a grid or by gradient based optimization, or a combination.

C. Target Tracking

In target tracking, a dynamic motion model of the form

$$x(t+1) = f\big(x(t), v(t)\big) \tag{20}$$

describes the motion of the target over time. The state $x(t)$ includes the position $\big(X(t), Y(t)\big)$ at time t, and $v(t)$ denotes process noise. By letting the delay $\tau_n\big(X(t), Y(t)\big)$ be time-varying, a more or less standard nonlinear estimation problem is achieved, where (19) is the measurement equation. The explicit elimination of a large set of unknown parameters in the measurement equation (19) is, however, nonstandard.

IV. Numerical Experiment

The core algorithm that computes the LS estimate of DOA will be evaluated on real data and the results compared to a common DOA estimator based on beamforming.

An extensive field trial was performed in Lilla Gåra in 2013 [3]. Different motorized vehicles as well as an aircraft were used as test objects, and each trajectory was repeated several times. Particularly in this study, an *all terrain vehicle* (ATV) and a *motorcycle* (MC) serve as acoustic sources, and the sensor is a four microphone array, see Fig. 5 and 3. Ground truth is established by a high-performance GPS system (RTK GNSS). The scenario is depicted on an aerial image in Fig. 4.

The sound source is primarily the exhaust tube of the ATV and MC, which are located less than one meter behind the GPS. The sound propagation delay is less than a couple of meters. The DOA should thus be slightly biased, but this is compensated for in our results.

The sound is sampled with $48\,\mathrm{kHz}$ and down-sampled to $500\,\mathrm{Hz}$. A band-pass filter is applied to reduce low-frequency disturbances (non-engine sound) and with an upper frequency to satisfy $f_{\max} \ll c/D$. The examples to follow are based on using a non-causal (zero-phase) band-pass filter with pass-band 35–$175\,\mathrm{Hz}$, using a Butterworth filter order 8. In practice, a causal filter should be used, with no practical difference.

To compensate for different amplification and perhaps frequency characteristics of the microphones, the resulting signals are normalized to the same energy.

The LS loss function is computed in θ point-wise for each sample using a 2^{nd} order Taylor expansion ($L = 2$). To increase the SNR, 50 loss functions are added for a batch of 50 sound samples on the $500\,\mathrm{Hz}$ time scale. This corresponds to $0.1\,\mathrm{s}$, in which the vehicle has moved much less than a meter, thus the DOA can be seen as constant here.

A. Beamforming

The Taylor series expansion DOA estimation method will be compared to a basic reference estimator similar to beamforming, see [9]. The beamformer records the direction with, in some sense, optimal array response. The response we here define in terms of an aggregation of similarities between every pair of sensors and for several time instances $t_1, \ldots, t_n$,

$$V_{\text{BF}}(\theta) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{t=t_1}^{t_m} \Big[z_i\big(t+\tau_i(\theta)\big) - z_j\big(t+\tau_j(\theta)\big) \Big]^2, \quad (21)$$

and the beamformer output consequently as

$$\hat{\theta}_{\text{BF}} = \arg\min_{\theta} V_{\text{BF}}(\theta). \quad (22)$$

The advantage of this formulation of beamforming is that we use exactly the same expression for $\tau_i(\theta)$ as in the proposed approach to make the comparison as fair as possible. For instance, we allow the array formation to be arbitrary.

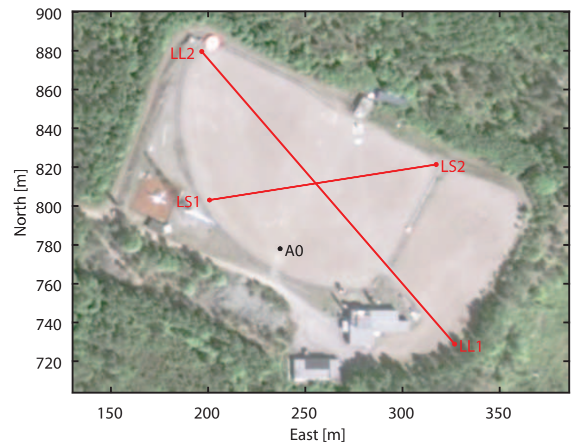


Fig. 4. Experimental setup with the array position A0 and the two vehicle tracks LL (long line) and LS (short ditto). In the experiments an acoustic target impersonated by an *all-terrain vehicle* (ATV) and a motorcycle (MC) travels the LL and LS tracks, respectively. The speeds are approximately $30\,\mathrm{km/h}$ and the sensor passage distance ca $35\,\mathrm{m}$.

In its digital implementation, the analysis of minor signal time delays requires either very high sampling rate or interpolation techniques, and is thus rather computer intensive in this naive form. However, it is here believed to be suitable as a performance reference.

The signals used in the beamformer are not down sampled, but keep the original sampling rate of $48\,\mathrm{Hz}$. To reduce the noise level, a band pass filter 35–$700\,\mathrm{Hz}$ is used as conditioner. The same batch size as above is used; $0.1\,\mathrm{s}$.

B. Results

In Fig. 6 an ATV sample of $V(\theta)$ is plotted together with $V_{\text{BF}}(\theta)$ for reference. The observed $180°$ ambiguity is a result of only using data from a single time instance which makes it impossible to tell the direction of time. This ambiguity can easily be resolved by combining data from more time instances and that way determine the sign of the signal derivatives.

Fig. 3. Microphone array used in the experiments. The microphones are symmetrically interspaced $42\,\mathrm{cm}$. Although not ideal (the middle microphone sticks up above the plane), it is assumed that all sensors (and sources) are in the same plane.

Fig. 5. One of two targets used in the experiment; an *all-terrain vehicle* (ATV). Although not apparent in this photograph, the reference GPS antenna was attached to the driver's right shoulder.

Another solution is to simply rely on prior knowledge about the target location.

Fig. 7 and 8 show the resulting DOA estimates as a function of time for the ATV departing from LL2 and LL1, respectively (see also the map in Fig. 4). The DOA estimate for the MC departing from LS2 is given in Fig. 9. The left hand side diagrams (a) give the DOA estimate compared to ground truth, and right hand side (b) the errors. In all diagrams, the linearization method is compared to beamforming.

During the *closest point of approach* (CPA), the error as illustrated in the right hand side diagrams (b) are typically in the order of a few degrees standard deviation and no bias. Initially, when the vehicle is far away, the standard deviation is larger naturally. The ATV examples show a DOA estimate bias close to LL1. This is probably caused by sound reflection in the large building seen in the lower right corner of Fig. 4. The beamformer is apparently more robust to this type of signal deficiencies.

The MC results in Fig. 9 are not suffering from reflections in the same way. The standard deviation is slightly larger though, speculatively due to a higher sound pitch, which is generally not advantageous for the linear model.

V. Conclusion

A new method for wideband *direction of arrival* (DOA) estimation has been proposed that can be used in both a near-field and far-field setup. Given a hypothesis on source direction (or position in the far-field case), the source signal is locally expanded to a Taylor series and the signal derivatives in this expansion can be solved for by ordinary linear least squares techniques. A DOA or position estimate can thus be pursued by the principle of least squares. The new method has been evaluated on a microphone array with passing motorized vehicles as acoustic sources.

The new method is operative down to very low sampling frequencies, which is an advantage in systems with hardware restrictions like battery-powered sensor nodes. We have proved that the new method works on real data, in most cases, with only slightly larger estimation error than a more traditional beamformer. However, the beamformer is anticipated to use much more computing.

In future research we will delve deeper into the computational requirements compared to other methods. It will also be studied how the new method can be used in a sensor network with multiple arrays, where the far-field assumption has been abandoned. A natural extension is also to introduce tracking to better cope with source movements.

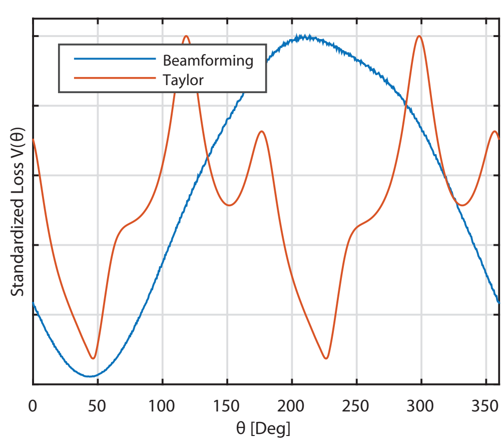


Fig. 6. Qualitative comparison between the loss function of the Taylor linearization method (13) and the beamforming loss function (22).

Acknowledgment

This work is supported by the Swedish Foundation for Strategic Research (SSF) through the Linköping University project Cooperative Localization (Cooploc), and also supported by projects at the Swedish Defence Research Agency (FOI) that are funded by the Swedish Armed Forces. These contributions are hereby gratefully acknowledged.

References

[1] M. R. Azimi-Sadjadi, A. Pezeshki, L. L. Scharf, and M. E. Hohil. Wideband DOA estimation algorithms for multiple target detection and tracking using unattended acoustic sensors. In *Proc. SPIE*, 2004.

[2] Y. Bucris, I. Cohen, and M. Doron. Bayesian focusing for coherent wideband beamforming. *Audio, Speech, and Language Processing, IEEE Transactions on*, 20(4):1282–1296, May 2012. ISSN 1558-7916. doi: 10.1109/TASL.2011.2175384.

[3] V. Deleskog, H. Habberstad, G. Hendeby, and L. David. Acoustic and visual sensor network measurements in lilla gåra 2012. Technical Report FOI Memo 4833, Swedish Defence Research Institute (FOI), 2014. URL http://www.foi.se/en/Our-Services/Reports/.

[4] M. A. Doron and A. Nevet. Robust wavefield interpolation for adaptive wideband beamforming. *Signal Processing*, 88:1579 – 1594, 2008. ISSN 0165-1684. URL https://login.e.bibl.liu.se/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S0165168408000121&site=eds-live.

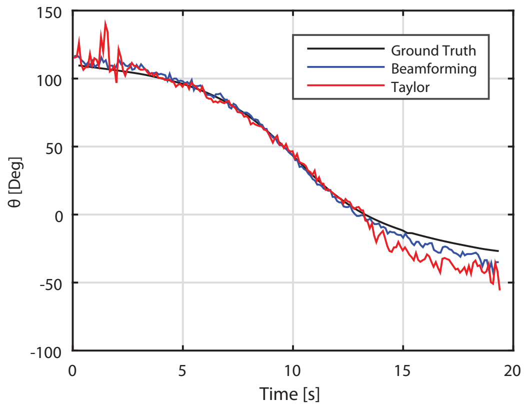
[5] F. Gustafsson. *Statistical Sensor Fusion*. Studentlitteratur, 2010.

[6] H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *Signal Processing Magazine, IEEE*, 13(4):67–94, Jul 1996. ISSN 1053-5888. doi: 10.1109/79.526899.

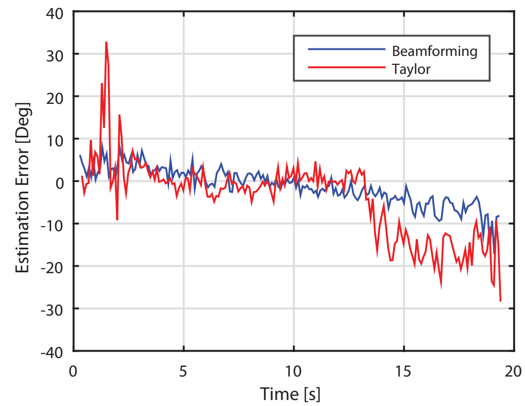
[7] E. Ozkan, M. B. Guldogan, U. Orguner, and F. Gustafsson. Ground multiple target tracking with a network of acoustic sensor arrays using phd and cphd filters. In *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, Jul 2011.

[8] P. Stoica, P. Babu, and J. Li. Spice: A sparse covariance-based estimation method for array processing. *Signal Processing, IEEE Transactions on*, 59(2):629–638, Feb 2011. ISSN 1053-587X. doi: 10.1109/TSP.2010.2090525.

[9] B. V. Veen and K. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, pages 4–24, 1988.

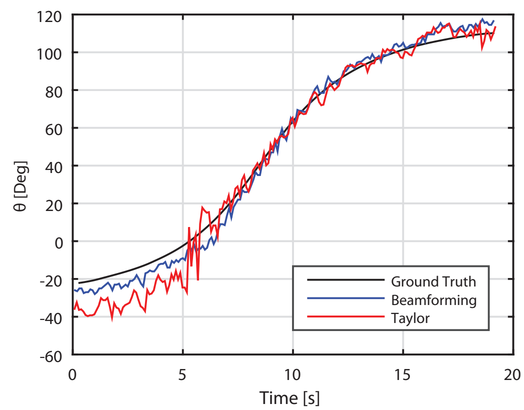


(a) Estimated DOA

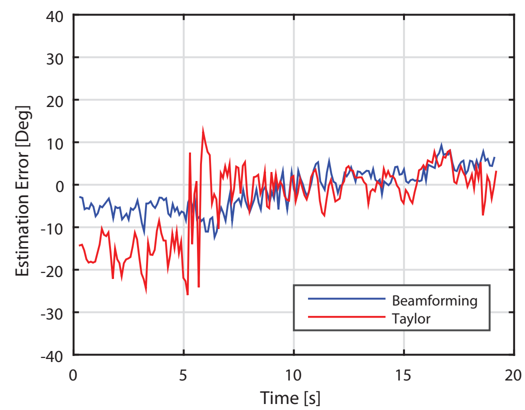


(b) Estimation Error

Fig. 7. DOA estimation results for the ATV going from LL2 to LL1, see Fig. 4. The estimate from the proposed Taylor based method (red) is compared to beamforming (blue) and ground truth (black).

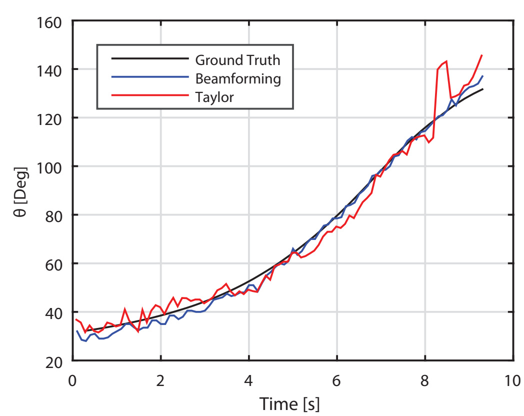


(a) Estimated DOA

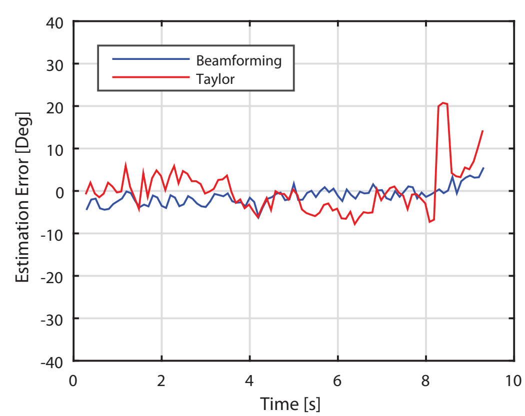


(b) Estimation Error

Fig. 8. DOA estimation results for the ATV going from LL1 to LL2, thus opposite heading compared to the experiment above.



(a) Estimated DOA



(b) Estimation Error

Fig. 9. DOA estimation results for the MC going from LS2 to LS1, see Fig. 4.