Vector-Sensor Array Processing for Electromagnetic Source Localization

Arye Nehorai, *Fellow, IEEE,* and Eytan Paldi, *Member, IEEE*

***Abstract*— We present a new approach for localizing electromagnetic sources using sensors where the output of each is a vector consisting of the complete six electric and magnetic field components. Two types of source transmissions are considered: 1) single signal transmission (SST), and 2) dual signal transmission (DST). The model is given in terms of several parameters, including the wave direction of arrival (DOA) and state of polarization. A compact expression is derived for the Cramér–Rao bound (CRB) on the estimation errors of these parameters for the multi-source multi-vector-sensor model. Quality measures including mean-square angular error (MSAE) and covariance of vector angular error (CVAE) are introduced, and their lower bounds are derived. The advantage of using vector sensors is highlighted by explicit evaluation of the MSAE and CVAE bounds for source localization with a *single* vector sensor. A simple algorithm for estimating the source DOA with this sensor is presented along with its statistical performance analysis.**

I. Introduction

THE localization of source signals using sensor data processing has attracted significant attention lately. Most existing methods employ sensor arrays in which the output of each sensor is a scalar corresponding, for example, to the pressure in the acoustic case or to a scalar function of the electric field in the electromagnetic case. This paper (see also [1]) considers new methods for multiple electromagnetic source localization using sensors whose output is a *vector* corresponding to the complete electric and magnetic fields at the sensor. These sensors, which will be called *vector sensors,* can consist for example of two orthogonal triads of scalar sensors that measure the electric and magnetic field components. The main advantage of the vector sensors is that they make use of all available electromagnetic information and, hence, should outperform the scalar-sensor arrays in accuracy of direction of arrival (DOA) estimation. Vector sensors should also allow the use of smaller array apertures while maintaining performance. (Note that we use the term "vector sensor" for a device that measures a complete physical vector quantity.)

Section II derives the measurement model. The electromagnetic sources considered can originate from two types of transmissions: 1) single signal transmission (SST), in which a single signal message is transmitted and 2) dual signal transmission (DST), in which two separate signal messages are transmitted simultaneously (from the same source); see, for example, [2] and [3]. The interest in DST is due to the fact that it makes full use of the two spatial degrees of freedom present in a transverse electromagnetic planewave. This is particularly important in the wake of increasing demand for economical spectrum usage by existing and emerging modern communication technologies.

Section III analyzes the minimum attainable variance of unbiased DOA estimators for a general vector-sensor array model and multi-electromagnetic sources that are assumed to be stochastic and stationary. A compact expression for the corresponding Cramér–Rao bound (CRB) on the DOA estimation error that extends previous results for the scalar-sensor array case in [4] (see also [5]) is derived.

A significant property of the vector sensors is that they enable DOA (azimuth and elevation) estimation of an electromagnetic source with a *single* vector sensor and a single snapshot. This result is explicitly shown by using the CRB expression for this problem in Section IV. A bound on the associated asymptotic normalized mean-square angular error (MSAE), which is invariant to the reference coordinate system, is used for an in-depth performance study. Compact expressions for this MSAE bound provide physical insight into the SST and DST source localization problems with a single vector sensor.

The CRB matrix for an SST source in the sensor coordinate frame exhibits some nonintrinsic singularities (i.e., singularities that are not inherent in the physical model while being dependent on the choice of the reference coordinate system) and has complicated entry expressions. Therefore, we introduce a new vector angular error defined in terms of the incoming wave frame. A bound on the asymptotic normalized covariance of the vector angular error (CVAE) is derived. The relationship between the CVAE and MSAE and their bounds is presented. The CVAE matrix bound for the SST-source case is shown to be diagonal, easy to interpret, and has only intrinsic singularities.

A simple algorithm is proposed for estimating the source DOA with a single vector sensor, motivated by the Poynting vector. The algorithm is applicable to various types of sources (e.g., wide-band and non-Gaussian); it does not require minimization of a cost function and can be applied in real time. Statistical performance analysis evaluates the variance of the estimator under mild assumptions and compares it with the MSAE lower bound.

Manuscript received June 13, 1992; revised January 15, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Daniel Fuhrmann. This work was supported by the Air Force Office of Scientific Research under Grant AFOSR-90-10164, the Office of Naval Research under Grant N00014-91-J-1298, the National Science Foundation under Grant MIP-9122753, and the HTI Fellowship.

The authors are with the Department of Electrical Engineering, Yale University, New Haven, CT 06520-8284.

IEEE Log Number 9214202.

Section V extends these results to the multi-source multi-vector-sensor case with special attention to the two-source single-vector-sensor case. Section VI summarizes the main results and gives some ideas of possible extensions.

The main difference between the present paper and previous papers on source direction estimation is in our use of vector sensors with *complete* electric and magnetic data. Most previous papers dealt with scalar sensors. Other papers that considered estimation of the polarization state and source direction are [6]–[11]. Reference [6] discussed the use of subspace methods to solve this problem using diversely polarized electric sensors. References [7]–[9] devised algorithms for arrays with 2-D electric measurements. Reference [10] provided performance analysis for scalar arrays with two types of electric sensor polarizations (diversely polarized). An earlier reference [11] proposed an estimation method using a 3-D vector sensor and implemented it with magnetic sensors. All these references used only part of the electromagnetic information at the sensors, thereby reducing the observability of DOA's. In most of them, time delays between distributed sensors played an essential role in the estimation process.

For a planewave (which is typically associated with a single source in the far field), the magnitude of the electric and magnetic fields can be found from each other. Hence, it may be felt that one (complete) field is deducible from the other. However, this is not true when the source direction is unknown. Additionally, the electric and magnetic fields are orthogonal to each other and to the source DOA vector; hence, measuring both fields significantly increases the accuracy of the source DOA estimation. This is true in particular for an incoming wave that is nearly linearly polarized, as will be explicitly shown by the Cramér–Rao bound.

Our proposed use of the complete electromagnetic vector data enables source parameter estimation with a single vector sensor (even with a single snapshot) where time delays are not used at all. In fact, this is shown to be possible for at least two sources. As a result, the derived CRB expressions for this problem are applicable to wide-band sources. The source DOA parameters considered include azimuth and elevation. This paper also considers, to the best of the authors' knowledge, for the first time direction estimation to DST sources, as well as the CRB on wave ellipticity and orientation angles (which will be defined later) for SST sources using vector sensors. The MSAE and CVAE quality measures and the associated bounds are also new. Their application is not limited to electromagnetic vector-sensor processing.

II. The Measurement Model

This section presents the measurement model for the estimation problems that are considered in the later parts of the paper.

A. Single-Source Single-Vector-Sensor Model

1) Basic Assumptions: Throughout this paper, it will be assumed that the wave is traveling in a nonconductive, homogeneous, and isotropic medium. Additionally, the following will be assumed:

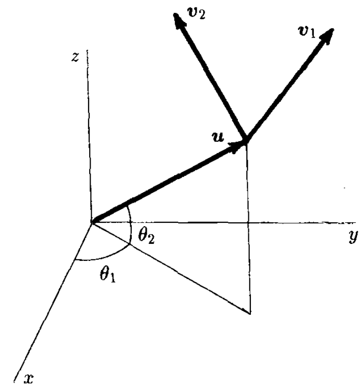


Fig. 1. The orthonormal vector triad $(\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2)$.

A1: *Planewave at the sensor*: This is equivalent to a far-field assumption (or a maximum wavelength that is much smaller than the source-to-sensor distance), a point-source assumption (i.e., the source size is much smaller than the source-to-sensor distance), and a point-like sensor (i.e., the sensor's dimensions are small compared with the minimum wavelength).

A2: *Band-limited spectrum*: The signal has a spectrum including only frequencies ω satisfying $\omega_{\min} \leq |\omega| \leq \omega_{\max}$, where $0 < \omega_{\min} < \omega_{\max} < \infty$. This assumption is satisfied in practice. The lower and upper limits on ω are also needed, respectively, for the far-field and point-like sensor assumption.

Let $\mathcal{E}(t)$ and $\mathcal{H}(t)$ be the vector-phasor representations (or complex envelopes, see, e.g., [12], [13], and Appendix A) of the electric and magnetic fields at the sensor. In addition, let $\boldsymbol{u}$ be the unit vector at the sensor pointing towards the source, i.e.

$$\boldsymbol{u} = \begin{bmatrix} \cos\theta_1 \cos\theta_2 \\ \sin\theta_1 \cos\theta_2 \\ \sin\theta_2 \end{bmatrix} \tag{2.1}$$

where θ_1 and θ_2 denote, respectively, the azimuth and elevation angles of $\boldsymbol{u}$; see Fig. 1. Thus, $\theta_1 \in [0, 2\pi)$, and $|\theta_2| \leq \pi/2$.

In Appendix A it is shown that for planewaves Maxwell's equations can be reduced to an equivalent set of two equations without any loss of information. Under the additional assumption of a band-limited signal, these two equations can be written in terms of phasors. The results are summarized in the following theorem.

Theorem 2.1: Under assumption **A1** Maxwell's equations can be reduced to an equivalent set of two equations. With the additional band-limited spectrum assumption **A2**, they can be written as

$$\boldsymbol{u} \times \mathcal{E}(t) = -\eta \mathcal{H}(t) \tag{2.2a}$$

$$\boldsymbol{u} \cdot \mathcal{E}(t) = 0 \tag{2.2b}$$

where η is the intrinsic impedance of the medium, and "$\times$" and "$\cdot$" are the cross and inner products of $\mathbb{R}^3$ applied to vectors in $\mathbb{C}^3$. (That is, if $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{C}^3$, then $\boldsymbol{v} \cdot \boldsymbol{w} = \Sigma_i v_i w_i$. This is different from the usual inner product of $\mathbb{C}^3$).

Proof: See Appendix A. (Note that $\boldsymbol{u} = -\boldsymbol{\kappa}$, where $\boldsymbol{\kappa}$ is the unit vector in the direction of the wave propagation).

Thus, under the plane and band-limited wave assumptions, the vector-phasor equations (2.2) provide all the information contained in the original Maxwell equations. This result will be used in the following to construct measurement models in which the Maxwell equations are incorporated entirely.

2) The Measurement Model: Suppose that a vector sensor measures all the six components of the electric and magnetic fields. (It is assumed that the sensor does not influence the electric and magnetic fields). The measurement model is based on the phasor representation of the measured electromagnetic data (with respect to a reference frame) at the sensor. Let $\boldsymbol{y}_E(t)$ be the measured electric field phasor vector at the sensor at time t and $\boldsymbol{e}_E(t)$ its noise component. Then, the electric part of the measurement will be

$$\boldsymbol{y}_E(t) = \mathcal{E}(t) + \boldsymbol{e}_E(t). \tag{2.3}$$

Similarly, from (2.2a), after appropriate scaling, the magnetic part of the measurement will be taken as

$$\boldsymbol{y}_H(t) = \boldsymbol{u} \times \mathcal{E}(t) + \boldsymbol{e}_H(t). \tag{2.4}$$

In addition to (2.3) and (2.4), we have the constraint (2.2b).

Define the matrix cross product operator that maps a vector $\boldsymbol{v} \in \mathbb{R}^{3\times 1}$ to $(\boldsymbol{u} \times \boldsymbol{v}) \in \mathbb{R}^{3\times 1}$ by

$$(\boldsymbol{u}\times) \triangleq \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix} \tag{2.5}$$

where u_x, u_y, u_z are the x, y, z components of the vector $\boldsymbol{u}$. With this definition, (2.3) and (2.4) can be combined to

$$\begin{bmatrix} \boldsymbol{y}_E(t) \\ \boldsymbol{y}_H(t) \end{bmatrix} = \begin{bmatrix} I_3 \\ (\boldsymbol{u}\times) \end{bmatrix} \mathcal{E}(t) + \begin{bmatrix} \boldsymbol{e}_E(t) \\ \boldsymbol{e}_H(t) \end{bmatrix} \tag{2.6}$$

where I_3 denotes the 3×3 identity matrix. For notational convenience, the dimension subscript of the identity matrix will be omitted whenever its value is clear from the context.

The constraint (2.2b) implies that the electric phasor $\mathcal{E}(t)$ can be written

$$\mathcal{E}(t) = V\boldsymbol{\xi}(t) \tag{2.7}$$

where V is a 3×2 matrix whose columns span the orthogonal complement of $\boldsymbol{u}$ and $\boldsymbol{\xi}(t) \in \mathbb{C}^{2\times 1}$. It is easy to check that the matrix

$$V = \begin{bmatrix} -\sin\theta_1 & -\cos\theta_1 \sin\theta_2 \\ \cos\theta_1 & -\sin\theta_1 \sin\theta_2 \\ 0 & \cos\theta_2 \end{bmatrix} \tag{2.8}$$

whose columns are orthonormal, satisfies this requirement. For future reference, we note that since $||\boldsymbol{u}||^2 = 1$, the columns of V, which are denoted by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$, can be constructed, for example, from the partial derivatives of $\boldsymbol{u}$ with respect to θ_1 and θ_2 and post normalization when needed. Thus

$$\boldsymbol{v}_1 = \frac{1}{\cos\theta_2} \frac{\partial \boldsymbol{u}}{\partial \theta_1} \tag{2.9a}$$

$$\boldsymbol{v}_2 = \boldsymbol{u} \times \boldsymbol{v}_1 = \frac{\partial \boldsymbol{u}}{\partial \theta_2} \tag{2.9b}$$

and $(\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2)$ is a right orthonormal triad; see Fig. 1. (Observe that the two coordinate systems shown in the figure actually have the same origin). The signal $\boldsymbol{\xi}(t)$ fully determines the components of $\mathcal{E}(t)$ in the plane where it lies, namely, the plane orthogonal to $\boldsymbol{u}$ spanned by $\boldsymbol{v}_1, \boldsymbol{v}_2$. This implies that there are two degrees of freedom present in the spatial domain (or the wave's plane) or that two independent signals can be transmitted simultaneously.

Combining (2.6) and (2.7), we now have

$$\begin{bmatrix} \boldsymbol{y}_E(t) \\ \boldsymbol{y}_H(t) \end{bmatrix} = \begin{bmatrix} I \\ (\boldsymbol{u}\times) \end{bmatrix} V\boldsymbol{\xi}(t) + \begin{bmatrix} \boldsymbol{e}_E(t) \\ \boldsymbol{e}_H(t) \end{bmatrix}. \tag{2.10}$$

This system is equivalent to (2.6) with (2.2b).

The measured signals in the sensor reference frame can be further related to the original source signal at the transmitter using the following lemma.

Lemma 2.1: Every vector $\boldsymbol{\xi} = [\xi_1, \xi_2]^T \in \mathbb{C}^{2\times 1}$ has the representation

$$\boldsymbol{\xi} = ||\boldsymbol{\xi}|| e^{i\varphi} Q\boldsymbol{w} \tag{2.11}$$

where

$$Q = \begin{bmatrix} \cos\theta_3 & \sin\theta_3 \\ -\sin\theta_3 & \cos\theta_3 \end{bmatrix} \tag{2.12a}$$

$$\boldsymbol{w} = \begin{bmatrix} \cos\theta_4 \\ i\sin\theta_4 \end{bmatrix} \tag{2.12b}$$

and where $\varphi \in (-\pi, \pi], \theta_3 \in (-\pi/2, \pi/2], \theta_4 \in [-\pi/4, \pi/4]$. Moreover, $||\boldsymbol{\xi}||, \varphi, \theta_3, \theta_4$ in (2.11) are uniquely determined if and only if $\xi_1^2 + \xi_2^2 \neq 0$.

Proof: See Appendix B.

The equality $\xi_1^2 + \xi_2^2 = 0$ holds if and only if $|\theta_4| = \pi/4$, corresponding to circular polarization (defined below). Hence, from Lemma 2.1, the representations (2.11) and (2.12) is not unique in this case, as should be expected, since the orientation angle θ_3 is ambiguous. It should be noted that the representations (2.11) and (2.12) are known and were used (see, e.g., [14]) without a proof. However, Lemma 2.1 of existence and uniqueness appears to be new. The existence and uniqueness properties are important to guarantee identifiability of parameters.

The physical interpretations of the quantities in the representations (2.11) and (2.12) are as follows:

$\|\|\boldsymbol{\xi}\|\| e^{i\varphi}$	Complex envelope of the source signal (including amplitude and phase)
$\boldsymbol{w}$	normalized overall transfer vector of the source's antenna and medium, i.e., from the source complex envelope signal to the principal axes of the received electric wave
Q	rotation matrix that performs the rotation from the principal axes of the incoming electric wave to the $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ coordinates.

Let ω_c be the reference frequency of the signal phasor representation; see Appendix A. In the narrow-band SST case, the incoming electric wave signal $\mathrm{Re}\{e^{i\omega_c t}\|\boldsymbol{\xi}(t)\|e^{i\varphi}(t)Q\boldsymbol{w}\}$ moves on a quasistationary ellipse whose semi-major and semi-minor axes' lengths are proportional, respectively, to $\cos\theta_4$ and $\sin\theta_4$; see Fig. 2 and [15]. The ellipse's eccentricity is thus determined by the magnitude of θ_4. The sign of θ_4 determines the spin sign or direction. More precisely, a positive (negative) θ_4 corresponds to a positive (negative) spin with right- (left)-handed rotation with respect to the wave propagation vector $\kappa = -\boldsymbol{u}$. As shown in Fig. 2, θ_3 is the rotation angle between the $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ coordinates and the electric ellipse axes $(\tilde{\boldsymbol{v}}_1, \tilde{\boldsymbol{v}}_2)$. The angles θ_3 and θ_4 will be referred to, respectively, as the orientation and ellipticity angles of the received electric wave ellipse. In addition to the electric ellipse, there is also a similar but perpendicular magnetic ellipse.

It should be noted that if the transfer matrix from the source to the sensor is time invariant, then so are θ_3 and θ_4.

The signal $\boldsymbol{\xi}(t)$ can carry information coded in various forms. In the following, we discuss briefly both existing forms and some motivated by the above representation.

3) Single Signal Transmission (SST) Model: Suppose that a single modulated signal is transmitted. Then, using (2.11), this is a special case of (2.10) with

$$\boldsymbol{\xi}(t) = Q\boldsymbol{w}s(t) \tag{2.13}$$

where $s(t)$ denotes the complex envelope of the (scalar) transmitted signal. Thus, the measurement model is

$$\begin{bmatrix}\boldsymbol{y}_E(t)\\ \boldsymbol{y}_H(t)\end{bmatrix} = \begin{bmatrix} I \\ (\boldsymbol{u}\times)\end{bmatrix} VQ\boldsymbol{w}s(t) + \begin{bmatrix}\boldsymbol{e}_E(t)\\ \boldsymbol{e}_H(t)\end{bmatrix}. \tag{2.14}$$

Special cases of this transmission are linear polarization with $\theta_4 = 0$ and circular polarization with $|\theta_4| = \pi/4$.

Recall that since there are two spatial degrees of freedom in a transverse electromagnetic planewave, one could, in principle, transmit two separate signals simultaneously. Thus, the SST method does not make full use of the two spatial degrees of freedom present in a transverse electromagnetic planewave.

4) Dual Signal Transmission (DST) Models: Methods of transmission in which two separate signals are transmitted simultaneously from the same source will be called *dual signal transmissions.* Various DST forms exist, and all of them can be modeled by (2.10), with $\boldsymbol{\xi}(t)$ being a linear transformation of the 2-D source signal vector.

One DST form uses two linearly polarized signals that are spatially and temporally orthogonal with an amplitude or phase modulation (see, e.g., [2], [3]). This is a special case of (2.10), where the signal $\boldsymbol{\xi}(t)$ is written in the form

$$\boldsymbol{\xi}(t) = Q\begin{bmatrix} s_1(t)\\ is_2(t)\end{bmatrix} \tag{2.15}$$

where $s_1(t)$ and $s_2(t)$ represent the complex envelopes of the transmitted signals. To guarantee unique decoding of the two signals (when θ_3 is unknown) using Lemma 2.1, they have to satisfy $s_1(t) \neq 0, s_2(t)/s_1(t) \in (-1, 1)$. (Practically, this can be achieved by using a proper electronic antenna adapter that yields a desirable overall transfer matrix.)

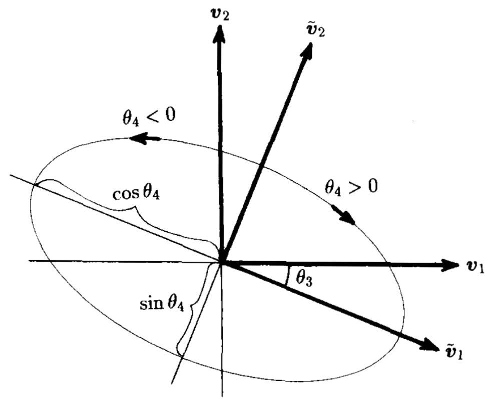


Fig. 2. Electric polarization ellipse.

Another DST form uses two circularly polarized signals with opposite spins. In this case

$$\boldsymbol{\xi}(t) = Q[\boldsymbol{w}\tilde{s}_1(t) + \overline{\boldsymbol{w}}\tilde{s}_2(t)] \tag{2.16a}$$

$$\boldsymbol{w} = (1/\sqrt{2})[1, i]^T \tag{2.16b}$$

where $\overline{\boldsymbol{w}}$ denotes the complex conjugate of $\boldsymbol{w}$. The signals $\tilde{s}_1(t), \tilde{s}_2(t)$ represent the complex envelopes of the transmitted signals. The first term on the r.h.s. of (2.16) corresponds to a signal with positive spin and circular polarization ($\theta_4 = \pi/4$), whereas the second term corresponds to a signal with negative spin and circular polarization ($\theta_4 = -\pi/4$). The uniqueness of (2.16) is guaranteed without the conditions needed for the uniqueness of (2.15).

The above-mentioned DST models can be applied to communciation problems. Assuming that $\boldsymbol{u}$ is given, it is possible to measure the signal $\boldsymbol{\xi}(t)$ and recover the original messages as follows. For (2.15), an existing method resolves the two messages using mechanical orientation of the receiver's antenna (see, e.g., [3]). Alternatively, this can be done electronically using the representation of Lemma 2.1 without the need to know the orientation angle. For (2.16), note that $\boldsymbol{\xi}(t) = \boldsymbol{w}e^{i\theta_3}\tilde{s}_1(t) + \overline{\boldsymbol{w}}e^{-i\theta_3}\tilde{s}_2(t)$, which implies the uniqueness of (2.16) and indicates that the orientation angle has been converted into a phase angle whose sign depends on the spin sign. The original signals can be directly recovered from $\boldsymbol{\xi}(t)$ up to an additive constant phase without knowledge of the orientation angle. In some cases it is of interest to estimate the orientation angle. Let W be a matrix whose columns are $\boldsymbol{w}, \overline{\boldsymbol{w}}$. For (2.16), this can be done using equal calibrating signals and then premultiplying the measurement by W^{-1} and measuring the phase difference between the two components of the result. This can also be used for real-time estimation of the angular velocity $d\theta_3/dt$.

In general, it can be stated that the advantage of the DST method is that it makes full use of the spatial degrees of freedom of transmission. However, the above DST methods need the knowledge of $\boldsymbol{u}$ and, in addition, may suffer from possible cross polarizations (see, e.g., [2]), multi-path effects, and other unknown distortions from the source to the sensor.

The use of the proposed vector sensor can motivate the design of new improved transmission forms. Here, we suggest a new dual signal transmission method that uses on-line electronic calibration in order to resolve the above problems. Similar to the previous methods it also makes full use of the spatial degrees of freedom in the system. However, it overcomes the need to know $\boldsymbol{u}$ and the overall transfer matrix from source to sensor.

Suppose the transmitted signal is $\boldsymbol{z}(t) \in \mathbb{C}^{2\times 1}$ (this signal is as it appears before reaching the source's antenna). The measured signal is

$$\begin{bmatrix} \boldsymbol{y}_E(t) \\ \boldsymbol{y}_H(t) \end{bmatrix} = C(t)\boldsymbol{z}(t) + \begin{bmatrix} \boldsymbol{e}_E(t) \\ \boldsymbol{e}_H(t) \end{bmatrix} \tag{2.17}$$

where $C(t) \in \mathbb{C}^{6\times 2}$ is the unknown source-to-sensor transfer matrix that may be slowly varying due, for example, to the source dynamics. To facilitate the identification of $\boldsymbol{z}(t)$, the transmitter can send calibrating signals, for instance, transmit $\boldsymbol{z}_1(t) = [1, 0]^T$ and $\boldsymbol{z}_2(t) = [0, 1]^T$, separately. Since these inputs are in phasor form, this means that actually constant carrier waves are transmitted. Obviously, one can then estimate the columns of $C(t)$ by averaging the received signals, which can be used later for finding the original signal $\boldsymbol{z}(t)$ by using, for example, least-squares estimation. Better estimation performance can be achieved by taking into account *a priori* information about the model.

In future research, it would be of interest to develop optimal coding methods (modulation forms) for maximum channel capacity while maintaining acceptable distortions of the decoded signals despite unknown varying channel characteristics.

Observe that actually, any combination of the variables $||\boldsymbol{\xi}||, \varphi, \theta_3$, and θ_4 can be modulated to carry information. A binary signal can be transmitted using the spin sign of the polarization ellipse (sign of θ_4). Lemma 2.1 guarantees the identifiability of these signals from $\boldsymbol{\xi}(t)$.

B. multi-source multi-vector-Sensor Model

Suppose that waves from n distant electromagnetic sources are impinging on an array of m vector sensors and that assumptions **A1** and **A2** hold for each source. To extend the model (2.10) to this scenario, we need the following additional assumptions, which imply that **A1** and **A2** hold uniformly on the array:

A3: *Planewave across the array*: In addition to **A1**, for each source, the array size d_A has to be much smaller than the source-to-array distance so that the vector $\boldsymbol{u}$ is approximately independent of the individual sensor positions.

A4: *Narrow-band signal assumption*: The maximum frequency of $\mathcal{E}(t)$, which is denoted by ω_m, satisfies $\omega_m\, d_A/c \ll 1$, where c is the velocity of wave propagation (i.e., the minimum modulating wavelength is much larger than the array size). This implies that $\mathcal{E}(t - \tau) \simeq \mathcal{E}(t)$ for all differential delays τ of the source signals between the sensors.

Note that (under the assumption $\omega_m < \omega_c$) since $\omega_m = \max\{|\omega_{\min} - \omega_c|, |\omega_{\max} - \omega_c|\}$, it follows that **A4** is satisfied if $(\omega_{\max} - \omega_{min})\, d_A/2c \ll 1$, and ω_c is chosen to be close enough to $(\omega_{\max} + \omega_{\min})/2$.

Let $\boldsymbol{y}_{EH}(t)$ and $\boldsymbol{e}_{EH}(t)$ be the $6m \times 1$-dimensional electromagnetic sensor phasor measurement and noise vectors

$$\boldsymbol{y}_{EH}(t) \triangleq [(\boldsymbol{y}_E^{(1)}(t))^T, (\boldsymbol{y}_H^{(1)}(t))^T, \cdots, (\boldsymbol{y}_E^{(m)}(t))^T, (\boldsymbol{y}_H^{(m)}(t))^T]^T \tag{2.18a}$$

$$\boldsymbol{e}_{EH}(t) \triangleq [(\boldsymbol{e}_E^{(1)}(t))^T, (\boldsymbol{e}_H^{(1)}(t))^T, \cdots, (\boldsymbol{e}_E^{(m)}(t))^T, (\boldsymbol{e}_H^{(m)}(t))^T]^T \tag{2.18b}$$

where $\boldsymbol{y}_E^{(j)}(t)$ and $\boldsymbol{y}_H^{(j)}(t)$ are, respectively, the measured phasor electric and magnetic vector fields at the jth sensor and similarly for the noise components $\boldsymbol{e}_E^{(j)}(t)$ and $\boldsymbol{e}_H^{(j)}(t)$. Then, under assumptions **A3** and **A4** and from (2.10), we find that the array measured phasor signal can be written as

$$\boldsymbol{y}_{EH}(t) = \sum_{k=1}^{n} \boldsymbol{e}_k \otimes \begin{bmatrix} I_3 \\ (\boldsymbol{u}_k \times) \end{bmatrix} V_k \boldsymbol{\xi}_k(t) + \boldsymbol{e}_{EH}(t) \tag{2.19}$$

where $\otimes$ is the Kronecker product and $\boldsymbol{e}_k$ denotes the kth column of the matrix $E \in \mathbb{C}^{m\times n}$ whose (j, k) entry is

$$E_{jk} = e^{-i\omega_c \tau_{jk}} \tag{2.20}$$

where τ_{jk} is the differential delay of the kth source signal between the jth sensor and the origin of some fixed reference coordinate system (e.g., at one of the sensors). Thus, $\tau_{jk} = -(\boldsymbol{u}_k \cdot \boldsymbol{r}_j)/c$, where $\boldsymbol{u}_k$ is the unit vector in the direction from the array of the kth source, and r_j is the position vector of the jth sensor in the reference frame. The rest of the notation in (2.19) is similar to the single-source case, cf., (2.1), (2.8), and (2.10). The vector $\boldsymbol{\xi}_k(t)$ can have either the SST or the DST form described above.

Observe that the signal-manifold matrix in (2.19) can be written as the Khatri–Rao product (see, e.g., [16], [17]) of E and a second matrix whose form depends on the source transmission type (i.e., SST or DST).

III. Cramér–Rao Bound for a Vector-Sensor Array

A. Statistical Model

Consider the problem of finding the parameter vector $\boldsymbol{\theta}$ in the following discrete-time vector-sensor array model associated with n vector sources and m vector sensors

$$\boldsymbol{y}(t) = A(\boldsymbol{\theta})\boldsymbol{x}(t) + \boldsymbol{e}(t) \qquad t = 1, 2, \cdots \tag{3.1}$$

where $\boldsymbol{y}(t) \in \mathbb{C}^{\overline{\mu}\times 1}$ are the vectors of observed sensor outputs (or snapshots), $\boldsymbol{x}(t) \in \mathbb{C}^{\overline{\nu}\times 1}$ are the unknown source signals, and $\mathbf{e}(t) \in \mathbb{C}^{\overline{\mu}\times 1}$ are the additive noise vectors. The transfer matrix $A(\boldsymbol{\theta}) \in \mathbb{C}^{\overline{\mu}\times\overline{\nu}}$ and the parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{\overline{q}\times 1}$ are given by

$$A(\boldsymbol{\theta}) = [A_1(\boldsymbol{\theta}^{(1)}) \cdots A_n(\boldsymbol{\theta}^{(n)})] \tag{3.2a}$$

$$\boldsymbol{\theta} = [(\boldsymbol{\theta}^{(1)})^T, \cdots, (\boldsymbol{\theta}^{(n)})^T]^T \tag{3.2b}$$

where $A_k(\boldsymbol{\theta}^{(k)}) \in \mathbb{C}^{\overline{\mu}\times\nu_k}$ and the parameter vector of the kth source $\boldsymbol{\theta}^{(k)} \in \mathbb{R}^{q_k\times 1}$, thus $\overline{\nu} = \Sigma_{k=1}^{n} \nu_k$, and $\overline{q} = \Sigma_{k=1}^{n} q_k$. The following notation will also be used

$$\boldsymbol{y}(t) = [(\boldsymbol{y}^{(1)}(t))^T, \cdots, (\boldsymbol{y}^{(m)}(t))^T]^T \tag{3.3a}$$

$$\boldsymbol{x}(t) = [(\boldsymbol{x}^{(1)}(t))^T, \cdots, (\boldsymbol{x}^{(n)}(t))^T]^T \tag{3.3b}$$

where $\boldsymbol{y}^{(j)}(t) \in \mathbb{C}^{\mu_j \times 1}$ is the vector measurement of the jth sensor, implying $\overline{\mu} = \Sigma_{j=1}^{m} \mu_j$, and $\boldsymbol{x}^{(k)}(t) \in \mathbb{C}^{\nu_k \times 1}$ is the vector signal of the kth source. Clearly, $\overline{\mu}$ and $\overline{\nu}$ correspond, respectively, to the total number of sensor components and source signal components.

The model (3.1) generalizes the commonly used multi-scalar-source multi-scalar-sensor one (see, e.g., [6], [18]). It will be shown later that the electromagnetic multi-vector-source multi-vector-sensor data models are special cases of (3.1) with appropriate choices of matrices.

For notational simplicity, the explicit dependence on $\boldsymbol{\theta}$ and t will be occasionally omitted.

We make the following commonly used assumptions on the model (3.1):

A5: The source signal sequence $\{\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots\}$ is a sample from a temporally uncorrelated stationary (complex) Gaussian process with zero mean and

$$E\boldsymbol{x}(t)\boldsymbol{x}^*(s) = P\delta_{t,s}$$
$$E\boldsymbol{x}(t)\boldsymbol{x}^T(s) = 0 \qquad (\text{for all } t \text{ and } s).$$

where E is the expectation operator, the superscript "$*$" denotes the conjugate transpose, and $\delta_{t,s}$ is the Kronecker delta.

A6: The noise $\boldsymbol{e}(t)$ is (complex) Gaussian distributed with zero mean and

$$E\boldsymbol{e}(t)\boldsymbol{e}^*(s) = \sigma^2 I \delta_{t,s}$$
$$E\boldsymbol{e}(t)\boldsymbol{e}^T(s) = 0 \qquad (\text{for all } t \text{ and } s).$$

It is also assumed that the signals $\boldsymbol{x}(t)$ and the noise $\boldsymbol{e}(s)$ are independent for all t and s.

A7: The matrix A has full rank $\overline{\nu} < \overline{\mu}$ (thus, A^*A is p.d.) and a continuous Jacobian $\partial A / \partial \boldsymbol{\theta}$ in some neighborhood of the true $\boldsymbol{\theta}$. The matrix $APA^* + \sigma^2 I$ is assumed to be positive definite, which implies that the probability density functions of the model are well defined in some neighborhood of the true $\boldsymbol{\theta}, P, \sigma^2$. Additionally, the matrix in braces in (3.4) is assumed to be nonsingular.

The unknown parameters in the model (3.1) include the vector $\boldsymbol{\theta}$, the signal covariance matrix P, and the noise variance σ^2. The problem of estimating $\boldsymbol{\theta}$ in (3.1) from N snapshots $\boldsymbol{y}(1), \cdots, \boldsymbol{y}(N)$ and the statistical performance of estimation methods are the main concerns of this paper.

B. The Cramér–Rao Bound

Consider the estimation of $\boldsymbol{\theta}$ in the model (3.1) under the above assumptions and with $\boldsymbol{\theta}, P, \sigma^2$ unknown. We have the following theorem.

Theorem 3.1: The Cramér–Rao lower bound on the covariance matrix of any (locally) unbiased estimator of the vector $\boldsymbol{\theta}$ in the model (3.1), under assumptions **A5–A7** with $\boldsymbol{\theta}, P, \sigma^2$ unknown and $\nu_k = \nu$ for all k, is a positive definite matrix given by

$$\text{CRB}(\boldsymbol{\theta}) = \frac{\sigma^2}{2N}\{\text{Re}\,[\text{btr}\,((1 \boxtimes U) \boxdot (D^* \Pi_c D)^{bT})]\}^{-1} \quad (3.4)$$

where

$$U = P(A^*AP + \sigma^2 I)^{-1} A^*AP \quad (3.5a)$$
$$\Pi_c = I - \Pi \quad (3.5b)$$
$$\Pi = A(A^*A)^{-1}A^* \quad (3.5c)$$
$$D = [D_1^{(1)} \cdots D_{q_1}^{(1)} \cdots D_1^{(n)} \cdots D_{q_n}^{(n)}] \quad (3.5d)$$
$$D_l^{(k)} = \frac{\partial A_k}{\partial \theta_l^{(k)}} \quad (3.5e)$$

and where 1 denotes a $\overline{q} \times \overline{q}$ matrix with all entries equal to one, and the block trace operator btr $(\cdot)$, the block Kronecker product $\boxtimes$, the block Schur-Hadamard product $\boxdot$, and the block transpose operator bT are as defined in Appendix I with blocks of dimensions $\nu \times \nu$, except for the matrix 1 that has blocks of dimensions $q_i \times q_j$.

Furthermore, the CRB in (3.4) remains the same independently of whether σ^2 is known or unknown.

Proof: See Appendix C.

Remarks: The assumption $\nu_k = \nu$ was necessary to cast the above result in a matrix form. When this condition is not satisfied, a scalar expression for the entries of the inverse CRB matrix that appears in Appendix C can be used. Observe that for $n = 1$, the block Kronecker product reduces to the usual Kronecker product. If $\nu = 1$, then all the block operators reduce to the usual scalar operators. If $\nu = 1$ and $q_k = 1, k = 1, \cdots, n$ then (3.4) reduces to the scalar case in [4] (see also [5]). In addition, note that $U = PA^*R^{-1}AP$ where

$$R = APA^* + \sigma^2 I. \quad (3.6)$$

Theorem 3.1 can be extended to include a larger class of unknown sensor noise covariance matrices (see Appendix D).

IV. MSAE, CVAE and Single-Source Single-Vector-Sensor Analysis

This section introduces the MSAE and CVAE quality measures and their bounds for source direction and orientation estimation in 3-D space. The bounds are applied to analyze the statistical performance of parameter estimation of an electromagnetic source whose covariance is unknown using a single vector sensor. Note that single-vector-sensor analysis is valid for *wide-band* sources as assumptions **A3** and **A4** are not needed.

A. The MSAE

We define the mean-square angular error, which is a quality measure that is useful for gaining physical insight into DOA (azimuth and elevation) estimation and for performance comparisons. The analysis of this subsection is not limited to electromagnetic measurements or to Gaussian data.

The angular error, say δ, corresponding to a direction error $\Delta\boldsymbol{u}$ in $\boldsymbol{u}$, can be shown to be $\delta = 2\arcsin(||\Delta\boldsymbol{u}||/2)$. Hence, $\delta^2 = ||\Delta\boldsymbol{u}||^2 + O(||\Delta\boldsymbol{u}||^4)$. Since $\Delta\boldsymbol{u} = (\partial\boldsymbol{u}/\partial\theta_1)\Delta\theta_1 + (\partial\boldsymbol{u}/\partial\theta_2)\Delta\theta_2 + O((\Delta\theta_1)^2 + (\Delta\theta_2)^2)$, where $\Delta\theta_1, \Delta\theta_2$ are the errors in θ_1 and θ_2, we have

$$\delta^2 = (\cos\theta_2 \cdot \Delta\theta_1)^2 + (\Delta\theta_2)^2 + O(|\Delta\theta_1|^3 + |\Delta\theta_2|^3). \quad (4.1)$$

We introduce the following definitions.

Definition 4.1: A model will be called *regular* if it satisfies any set of sufficient conditions for the CRB to hold (see, e.g., [19] and [20]).

Definition 4.2: The *asymptotic normalized mean-square angular error* of a direction estimator will be defined as

$$\mathrm{MSAE} \triangleq \lim_{N\to\infty} \{NE(\delta^2)\} \tag{4.2}$$

whenever this limit exists.

Definition 4.3: A direction estimator will be called *regular* if its errors satisfy $E[|\Delta\theta_1|^3 + |\Delta\theta_2|^3] = o(1/N)$, the gradient of its bias with respect to θ_1, θ_2 exists and is $o(1)$ as $N \to \infty$, and its MSAE exists. (If $|\theta_2| = \pi/2$, then θ_1 is undefined, and we can use the equivalent condition $E[\|\Delta \boldsymbol{u}\|^3] = o(1/N)$).

Equation (4.1) shows that under the assumptions that the model and estimator are regular, we have

$$E(\delta)^2 \geq [\cos^2\theta_2 \cdot \mathrm{CRB}(\theta_1) + \mathrm{CRB}(\theta_2)] + o(1/N) \quad \text{as } N \to \infty \tag{4.3}$$

where $\mathrm{CRB}(\theta_1)$ and $\mathrm{CRB}(\theta_2)$ are, respectively, the Cramér–Rao bounds for the azimuth and elevation. Using (4.3) we have the following theorem.

Theorem 4.1: For a regular model, the MSAE of any regular direction estimator is bounded from below by

$$\mathrm{MSAE}_{CR} \triangleq N[\cos^2\theta_2 \cdot \mathrm{CRB}(\theta_1) + \mathrm{CRB}(\theta_2)]. \tag{4.4}$$

Observe that MSAE_{CR} is not a function of N. Additionally, MSAE_{CR} is a tight bound if it is attained by some second-order efficient regular estimator (which is usually the maximum likelihood (ML) estimator; see e.g., [21]). For vector-sensor measurements, this bound has the desirable property of being invariant to the choice of reference coordinate frame since the information content in the data is invariant under rotational transformations. This invariance property also holds for the MSAE of an estimator if the estimate is independent of known rotational transformations of the data.

For a regular model, the bound (4.4) can be used for performance analysis of any regular direction (azimuth and elevation) finding algorithm.

B. DST Source Analysis

Assume that we wish to estimate the direction to a DST source whose covariance is unknown using a vector sensor. We will first present a statistical model for this problem as a special case of (3.1) and then investigate in detail the resulting CRB and MSAE.

The measurement model for the DST case is given in (2.10). Suppose the noise vector of (2.10) is (complex) Gaussian with zero mean and the following covariances:

$$E\begin{bmatrix} \boldsymbol{e}_E(t) \\ \boldsymbol{e}_H(t)\end{bmatrix}[\boldsymbol{e}_E^*(s), \boldsymbol{e}_H^*(s)] = \begin{bmatrix} \sigma_E^2 I_3 & 0 \\ 0 & \sigma_H^2 I_3 \end{bmatrix}\delta_{t,s}$$
$$E\begin{bmatrix} \boldsymbol{e}_E(t) \\ \boldsymbol{e}_H(t)\end{bmatrix}[\boldsymbol{e}_E^T(s), \boldsymbol{e}_H^T(s)] = 0 \qquad (\text{for all } t \text{ and } s).$$

Our assumption that the noise components are statistically independent stems from the fact that they are created separately at *different* sensor components (even if the sensor components belong to a vector sensor). Note that under assumption **A1** the measurement includes a source plane-wave component and sensor self noise.

To relate the model (2.10) to (3.1), define a scaled measurement $\boldsymbol{y}(t) \triangleq [r\boldsymbol{y}_E^T(t), \boldsymbol{y}_H^T(t)]^T$, where $r \triangleq \sigma_H/\sigma_E$ is assumed to be known. (The results of this section actually hold also when r is unknown as will be explained later). The resulting scaled noise vector $\boldsymbol{e}(t) \triangleq [r\boldsymbol{e}_E^T(t), \boldsymbol{e}_H^T(t)]^T$ then satisfies assumption **A6** with $\sigma = \sigma_H$. Assume further that the signal $\boldsymbol{\xi}(t)$ satisfies assumption **A5** with $\boldsymbol{x}(t) = \boldsymbol{\xi}(t)$. Then, under these assumptions, the scaled version of the DST source (2.10) can be viewed as a special case of (3.1) with $m = n = 1$ and

$$A = \begin{bmatrix} rV \\ (\boldsymbol{u}\times)V \end{bmatrix} \qquad \boldsymbol{x}(t) = \boldsymbol{\xi}(t) \qquad \sigma^2 = \sigma_H^2$$
$$\boldsymbol{\theta} = [\theta_1, \theta_2]^T \tag{4.5}$$

where the unknown parameters are $\boldsymbol{\theta}, P, \sigma^2$. The parameter vector of interest is $\boldsymbol{\theta}$, whereas P and σ^2 are the so-called nuisance parameters.

The above discussion shows that the CRB expression (3.4) is applicable to the present problem with the special choice of variables in (4.5). Since $n = 1$ and $\overline{q} = 2$, we have $1 \boxtimes U = 1_2 \otimes U$, where 1_2 is a 2×2 matrix with all entries equal to one. Hence, in this case

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{\sigma^2}{2N}\{\mathrm{Re}\,[\mathrm{btr}\,((1_2 \otimes U) \boxdot (D^*\Pi_c D)^{bT})]\}^{-1}. \tag{4.6}$$

To compute the matrices U and Π_c, it is useful to note the following general properties of $\boldsymbol{u}$ and V

$$(\boldsymbol{u}\times)^* = -(\boldsymbol{u}\times) \tag{4.7a}$$
$$(\boldsymbol{u}\times)^2 = -(I - \boldsymbol{u}\boldsymbol{u}^*) \tag{4.7b}$$
$$VV^* = I - \boldsymbol{u}\boldsymbol{u}^*. \tag{4.7c}$$

From the orthonormality of $\boldsymbol{u}$ and the columns of V and using (4.5) and (4.7), we find that

$$A^*A = r^2V^*V + V^*(\boldsymbol{u}\times)^*(\boldsymbol{u}\times)V = (1 + r^2)I_2 \tag{4.8}$$

$$AA^* = \begin{bmatrix} r^2(I - \boldsymbol{u}\boldsymbol{u})^* & r(\boldsymbol{u}\times)^* \\ r(\boldsymbol{u}\times) & I - \boldsymbol{u}\boldsymbol{u}^* \end{bmatrix}. \tag{4.9}$$

Substitution of (4.8) into (3.5a) yields

$$U = P[P + \sigma_{||}^2 I_2]^{-1}P \tag{4.10}$$

where

$$\sigma_{||}^2 \triangleq \frac{\sigma_E^2 \cdot \sigma_H^2}{\sigma_E^2 + \sigma_H^2}. \tag{4.11}$$

The variance $\sigma_{||}^2$ can be viewed as an equivalent noise variance of two measurements with independent noise variances σ_E^2 and σ_H^2. (The subscript "$||$" is chosen by the analogy between the total information in these measurements (sum of the inverses of noise variances) and the conductivity of two resistors in parallel). Substitution of (4.8) and (4.9) in (3.5b) gives

$$\Pi_c = \frac{1}{1 + r^2}\begin{bmatrix} I + r^2\boldsymbol{u}\boldsymbol{u}^* & r(\boldsymbol{u}\times) \\ r(\boldsymbol{u}\times)^* & r^2 I + \boldsymbol{u}\boldsymbol{u}^* \end{bmatrix}. \tag{4.12}$$

The derivative matrix D (3.5d) specializes in this case to

$$D = \begin{bmatrix} \dfrac{\partial A}{\partial \theta_1} & \dfrac{\partial A}{\partial \theta_2} \end{bmatrix}. \tag{4.13}$$

To compute these derivatives note that

$$\frac{\partial \boldsymbol{v}_1}{\partial \theta_1} = \sin\theta_2 \cdot \boldsymbol{v}_2 - \cos\theta_2 \cdot \boldsymbol{u} \quad (4.14a)$$

$$\frac{\partial \boldsymbol{v}_1}{\partial \theta_2} = 0 \quad (4.14b)$$

$$\frac{\partial \boldsymbol{v}_2}{\partial \theta_1} = -\sin\theta_2 \cdot \boldsymbol{v}_1 \quad (4.14c)$$

$$\frac{\partial \boldsymbol{v}_2}{\partial \theta_2} = -\boldsymbol{u}. \quad (4.14d)$$

Hence, from (4.5)

$$\frac{\partial A}{\partial \theta_1} = \begin{bmatrix} r(\sin\theta_2 \cdot \boldsymbol{v}_2 - \cos\theta_2 \cdot \boldsymbol{u}) & -r\sin\theta_2 \cdot \boldsymbol{v}_1 \\ -\sin\theta_2 \cdot \boldsymbol{v}_1 & -(\sin\theta_2 \cdot \boldsymbol{v}_2 - \cos\theta_2 \cdot \boldsymbol{u}) \end{bmatrix} \quad (4.15a)$$

$$\frac{\partial A}{\partial \theta_2} = \begin{bmatrix} 0 & -r\boldsymbol{u} \\ -\boldsymbol{u} & 0 \end{bmatrix}. \quad (4.15b)$$

Substituting (4.15) into (4.13) and using (4.12), we obtain

$$D^*\Pi_c D = \begin{bmatrix} r^2\cos^2\theta_2 & 0 & 0 & r^2\cos\theta_2 \\ 0 & \cos^2\theta_2 & -\cos\theta_2 & 0 \\ 0 & -\cos\theta_2 & 1 & 0 \\ r^2\cos\theta_2 & 0 & 0 & r^2 \end{bmatrix}. \quad (4.16)$$

Hence

$$\operatorname{btr}((1_2 \otimes U) \boxdot (D^*\Pi_c D)^{bT}) = \begin{bmatrix} (r^2U_{11} + U_{22})\cos^2\theta_2 & (r^2U_{12} - U_{12}^*)\cos\theta_2 \\ (r^2U_{12}^* - U_{12})\cos\theta_2 & U_{11} + r^2U_{22} \end{bmatrix} \quad (4.17)$$

where U_{ij} denotes the (i,j) entry of U. Substituting (4.17) in (3.4), we find that the CRB for a DST source and a single vector sensor is

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{\sigma^2}{2N\Delta} \cdot \begin{bmatrix} U_{11} + r^2U_{22} & (1-r^2)\cos\theta_2 \cdot \operatorname{Re} U_{12} \\ (1-r^2)\cos\theta_2 \cdot \operatorname{Re} U_{12} & (r^2U_{11} + U_{22})\cos^2\theta_2 \end{bmatrix} \quad (4.18)$$

where

$$\Delta = [r^2(\operatorname{tr} U)^2 + (1-r^2)^2 \det(\operatorname{Re} U)]\cos^2\theta_2. \quad (4.19)$$

Recall that $\mathrm{CRB}(\boldsymbol{\theta})$ takes into account the fact that P and $\sigma = \sigma_H$ are unknown, whereas $r = \sigma_H/\sigma_E$ is known.

The case of unknown r can be analyzed using the approach of Appendix D with $\tilde{\boldsymbol{y}}(t) = [\boldsymbol{y}_E^T(t), \boldsymbol{y}_H^T(t)]^T$ and $\Gamma = \operatorname{block\,diag}\{rI_3, I_3\}$. We find that if σ_H is unknown, then the CRB (4.18) is the same whether r is known or unknown. In addition, it follows from Theorem 3.1 that if r is known, then (4.18) remains the same whether σ_H is known or unknown. Thus

$$\mathrm{CRB}(\boldsymbol{\theta})_{\sigma_H \text{ known}, r \text{ known}} = \mathrm{CRB}(\boldsymbol{\theta})_{\sigma_H \text{ unknown}, r \text{ unknown}}. \quad (4.20)$$

The l.h.s. and r.h.s. of (4.20) furnish lower and upper bounds on $\mathrm{CRB}(\boldsymbol{\theta})$ under any prior information about σ_E, σ_H. Hence, $\mathrm{CRB}(\boldsymbol{\theta})$ is independent of any prior information on σ_E, σ_H. In other words, $\boldsymbol{\theta}$ is decoupled from these noise parameters.

Substituting the (1, 1) and (2, 2) entries of the CRB matrix (4.18) into (4.4), we find that the MSAE_{CR} for the present DST problem is

$$\mathrm{MSAE}_{CR}^D = \frac{(\sigma_E^2 + \sigma_H^2)\sigma_E^2\sigma_H^2 \operatorname{tr} U}{2[\sigma_E^2\sigma_H^2(\operatorname{tr} U)^2 + (\sigma_E^2 - \sigma_H^2)^2 \det(\operatorname{Re} U)]}. \quad (4.21)$$

Observe that MSAE_{CR}^D is symmetric with respect to σ_E, σ_H, as should be expected from the Maxwell equations. MSAE_{CR}^D is not a function of $\theta_1, \theta_2, \theta_3$, as should be expected, since for vector-sensor measurements the MSAE bound is by definition invariant to the choice of coordinate system. Note that MSAE_{CR}^D is independent of whether σ_E and σ_H are known or unknown.

Similar analysis shows that when only the electric field is measured, (4.21) simplifies to

$$\mathrm{MSAE}_{CR}^D = \frac{\sigma_E^2 \operatorname{tr} U}{2\det(\operatorname{Re} U)}. \quad (4.22)$$

Expression (4.21) can be applied to various signal transmission models to obtain more explicit physical interpretations. Consider, for example, the case in which $P = \sigma_s^2 I$. Equation (2.15) shows that this may correspond to transmission of two uncorrelated orthogonal linearly polarized signals, each of which has variance σ_s^2. Alternatively, from (2.16), this may correspond with two uncorrelated circularly polarized signals with opposite spins and equal variances σ_s^2. Inserting $P = \sigma_s^2 I$ into (4.10) and (4.21), we obtain

$$\mathrm{MSAE}_{CR}^D = \frac{1+\varrho}{\varrho^2} \quad (4.23)$$

where $\varrho \triangleq \sigma_s^2/\sigma_{\|}^2$ is an effective signal-to-noise ratio (SNR); see also the comment after (4.11). Observe that when ϱ is small, (4.23) behaves as $1/\varrho^2$, whereas for large ϱ, it behaves as $1/\varrho$. When only the electric field is measured, (4.23) still holds but with $\sigma_{\|}^2 = \sigma_E^2$.

C. SST Source (DST Model) Analysis

Consider the MSAE for a single signal transmission source when the estimation is done under the assumption that the source is of a dual signal transmission type. In this case, the model (2.10) has to be used but with a signal in the form of (2.13). The signal covariance is then

$$P = \sigma_s^2 Q\boldsymbol{w}(Q\boldsymbol{w})^* \quad (4.24)$$

where $\sigma_s^2 = Es^2(t)$ and Q and $\boldsymbol{w}$ are defined in (2.12). Thus, rank $P = 1$ and P has a unit-norm eigenvector $Q\boldsymbol{w}$ with an eigenvalue σ_s^2. We use the following lemma to express the matrix U in terms of P.

Lemma 4.1: If P is an Hermitian matrix of rank = 1, and if $f(z)$ is a function defined on the spectrum of P, which satisfies $f(0) = 0$, then

$$f(P) = \frac{f(\operatorname{tr} P)}{\operatorname{tr} P} P. \quad (4.25)$$

Proof: See Appendix E.

To apply the lemma to our case, note from (4.10) that here $U = f(P)$ with $f(z) = z^2/(z + \sigma_{||}^2)$ and rank $P = 1$. Hence, since $f(0) = 0$, we have

$$U = \frac{\varrho}{1+\varrho} P \tag{4.26}$$

where

$$\varrho = \frac{\operatorname{tr} P}{\sigma_{||}^2} = \frac{\sigma_s^2}{\sigma_{||}^2}. \tag{4.27}$$

Using (4.26), we get

$$\operatorname{tr} U = \frac{\varrho}{1+\varrho} \sigma_s^2 \tag{4.28a}$$

$$\det(\operatorname{Re} U) = \left(\frac{\varrho}{1+\varrho} \operatorname{Im} P_{12}\right)^2 \tag{4.28b}$$

where the fact that $\det P = 0$ has been used. Using (4.24) and (2.12), it is found that $\operatorname{Im} P_{12} = -\sigma_s^2 \sin\theta_4 \cos\theta_4$. Using this relation and substituting (4.28a), (4.28b) into (4.21) and using the fact that $\sigma^2 = \sigma_H^2$, we find that

$$\begin{aligned} \mathrm{MSAE}_{CR}^{S} &= \frac{(1+\varrho)(\sigma_E^2 + \sigma_H^2)^2}{2\varrho^2[\sigma_E^2\sigma_H^2 + (\sigma_E^2 - \sigma_H^2)^2 \sin^2\theta_4 \cos^2\theta_4]} \\ &= \frac{(1+\varrho)(1+r^2)^2}{2\varrho^2[r^2 + (1-r^2)^2 \sin^2\theta_4 \cos^2\theta_4]} \end{aligned} \tag{4.29}$$

where MSAE_{CR}^{S} denotes the MSAE_{CR} bound for the SST problem under the DST model. (It will be shown later that the same result also holds under the SST model). Observe that MSAE_{CR}^{S} is symmetric with respect to σ_E, σ_H. It is also independent of whether σ_H and σ_E are known or unknown, as can be shown from Theorem 3.1 and Appendix D. In addition, MSAE_{CR}^{S} is not a function of $\theta_1, \theta_2, \theta_3$ since for vector-sensor measurements, the MSAE bound is invariant under rotational transformations of the reference coordinate system. On the other hand, MSAE_{CR}^{S} is influenced by the ellipticity angle θ_4 through the difference in the electric and magnetic noise variances.

Table I summarizes several special cases of the expression (4.29) for MSAE_{CR}^{S}. The elliptical polarization column corresponds to an arbitrary polarization angle $\theta_4 \in [-\pi/4, \pi/4]$. The circular and linear polarization columns are obtained, respectively, as special cases of (4.29) with $|\theta_4| = \pi/4$ and $\theta_4 = 0$. The row of precise (noise-free) electric measurement (with noisy magnetic measurements) is obtained by substituting $\sigma_E^2 = 0$ in (4.29). The row of electric measurement only is obtained by deriving the corresponding CRB and MSAE_{CR}^{S}. Alternatively, MSAE_{CR}^{S} can be found for this case by taking the limit of (4.29) as $\sigma_H^2 \to \infty$.

Observe from (4.29) that when $\sigma_H^2 \neq \sigma_E^2$, MSAE_{CR}^{S} is minimized for circular polarization and maximized for linear polarization. This result is illustrated in Fig. 3, which shows the square root of MSAE_{CR}^{S} as a function of $r = \sigma_H/\sigma_E$ for three types of polarizations ($\theta_4 = 0, \pi/12, \pi/4$). The equivalent SNR = $\sigma_s^2/\sigma_{||}^2$ is kept at one, whereas the individual electric and magnetic noise variances are varied to give the desired value of r. As r becomes larger or smaller than one, MSAE_{CR}^{S} increases more significantly for sources with polarization closer to linear.

TABLE I
MSAE BOUNDS FOR A SINGLE SIGNAL TRANSMISSION SOURCE

	Elliptical	Circular	Linear
General MSAE_{CR}^{S}	(4.29)	$\frac{2(1+\varrho)}{\varrho^2}$	$\frac{(1+\varrho)(\sigma_E^2+\sigma_H^2)}{2\varrho\sigma_s^2}$
Precise Electric Measurement	0	0	$\frac{\sigma_H^2}{2\sigma_s^2}$
Electric Measurement Only	$\frac{\sigma_E^2(\sigma_E^2+\sigma_s^2)}{2\sigma_s^4 \sin^2\theta_4 \cos^2\theta_4}$	$\frac{2\sigma_E^2(\sigma_E^2+\sigma_s^2)}{\sigma_s^4}$	∞

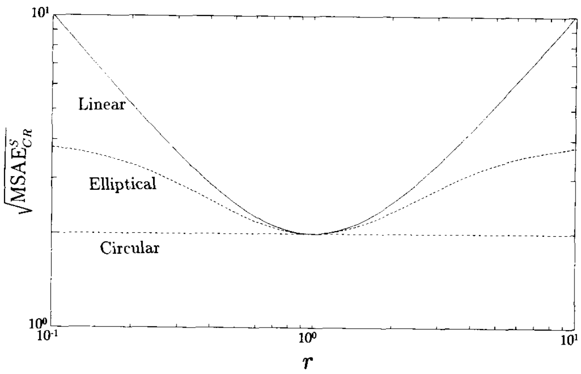


Fig. 3. Effect of change in $r = \sigma_H/\sigma_E$ on MSAE_{CR}^{S} for three types of polarizations ($\theta_4 = 0, \pi/12, \pi/4$). A single SST source, SNR = $\sigma_s^2/\sigma_{||}^2 = 1$.

When the electric (or magnetic) field is measured precisely and the source polarization is circular or elliptical, the MSAE_{CR}^{S} is zero (i.e., no angular error), whereas for linearly polarized sources, it remains positive. In the latter case, the contribution to MSAE_{CR}^{S} stems from the magnetic (or electric) noisy measurement. When only the electric (or magnetic) field is measured, MSAE_{CR}^{S} increases as the polarization changes from circular to linear. In the linear polarization case, MSAE_{CR}^{S} tends to infinity. In this case, it is impossible to uniquely identify the source direction $\boldsymbol{u}$ from the electric field only since $\boldsymbol{u}$ can then be anywhere in the plane orthogonal to the electric field vector.

The immediate conclusion is that as the source becomes closer to being linearly polarized it becomes more important to measure both the electric and magnetic fields to get good direction estimates using a single vector sensor.

These results are illustrated in Fig. 4, which shows the square root of MSAE_{CR}^{S} as a function of σ_H^2 and three polarization types ($\theta_4 = 0, \pi/12, \pi/4$). The standard deviations of the signal and electric noise are $\sigma_s = \sigma_E = 1$. The left side of the figure corresponds to (nearly) precise magnetic measurement, whereas the right side corresponds to (nearly) electric measurement only.

It is also of interest to note that the MSAE_{CR}^{S} for a circularly polarized SST source is twice the MSAE_{CR}^{D} for a DST source with two uncorrelated signals of circular polarization, opposite

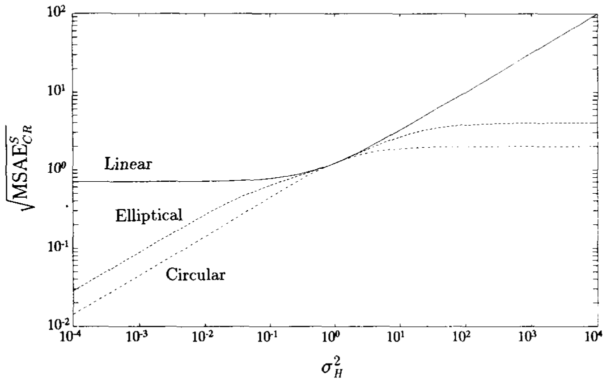


Fig. 4. Effect of change in magnitude of σ_H^2 on MSAE^S_{CR} for three types of polarizations ($\theta_4 = 0, \pi/12, \pi/4$). A single SST source, $\sigma_s = \sigma_E = 1$.

spins, and of equal power to the SST source; see (4.23). Thus, the quality of the DOA estimation is the same in these two cases. However, a similar comparison of the MSAE^S_{CR} for a linearly polarized SST source with that of a DST source with two uncorrelated linearly polarized signals (4.23) shows that the latter has in general a smaller MSAE_{CR}. This may be explained by the fact that both the electric and the magnetic fields of the DST source move in a plane rather than on a line, which makes it easier to estimate its direction.

D. SST Source (SST Model) Analysis

Suppose we wish to estimate the direction to an SST source whose variance is unknown using a single vector sensor, and the estimation is done under the correct model of an SST source. In the following, the CRB for this problem will be derived, and it will be shown that the resulting MSAE bound remains the same as when the estimation is done under the assumption of a DST source, that is, knowledge of the source type does not improve the accuracy of its direction estimate.

To get a statistical model for the SST measurement model (2.14) as a special case of (3.1), we will make the same assumptions on the noise, and use a similar data scaling as in the above DST source case. That will again give equal noise variances in all the sensor coordinates. Assume that the signal envelope $s(t)$ satisfies assumption **A5** with $x(t) = s(t)$ in (2.14). Then, the resulting statistical model becomes a special case of (3.1) with

$$A = \begin{bmatrix} rV \\ (\boldsymbol{u}\times)V \end{bmatrix} Q\boldsymbol{w} \qquad x(t) = s(t) \qquad \sigma^2 = \sigma_H^2$$
$$\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3, \theta_4]^T. \tag{4.30}$$

The unknown parameters are $\boldsymbol{\theta}, P, \sigma^2$.

Since in this case $n = m = \nu = 1$, the CRB matrix (3.4) reduces to

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{\sigma^2}{2NU}[\mathrm{Re}\,(D^*\Pi_c D)]^{-1} \tag{4.31}$$

where $U = \varrho\sigma_s^2/(1+\varrho)$. The matrix expression (4.31) was calculated, and its entries are presented in Appendix F. The results show that the ellipticity angle θ_4 is decoupled from the rest of the parameters and that its variance is not a function of these parameters. Additionally, the parameter vector $\boldsymbol{\theta}$ is decoupled from σ_E and σ_H.

The MSAE bound for an SST source under the SST model was calculated by inserting (F.1a) and (F.1e) with a proper normalization into (4.4). The result coincides with (4.29), that is, *the MSAE bound for an SST source is the same under both the SST and the DST models.*

Observe that (F.1h) implies that the CRB variance of the orientation angle θ_3 tends to infinity as the elevation angle θ_2 approaches $\pi/2$ or $-\pi/2$. This singularity is explained by the fact that the orientation angle is a function of the azimuth (through $\boldsymbol{v}_1, \boldsymbol{v}_2$), and the latter becomes increasingly sensitive to measurement errors as the elevation angle approances the zenith or nadir. (Note that the azimuth is undefined in the zenith and nadir elevations). However, this singularity is not an intrinsic one as it depends on the chosen reference system, whereas information in the vector measurement does not.

E. CVAE and SST Source Analysis in the Wave Frame

In order to get performance results intrinsic to the SST estimation problem and thereby solve the singularity problems associated with the above model, we choose an alternative error vector that is invariant under known rotational transformations of the coordinate system. The details of the following analysis appear in Appendix G.

Denote by W the wave frame whose coordinate axes are $(\boldsymbol{u}, \tilde{\boldsymbol{v}}_1, \tilde{\boldsymbol{v}}_2)$, where $\tilde{\boldsymbol{v}}_1$ and $\tilde{\boldsymbol{v}}_2$ correspond, respectively, to the major and minor axes of the source's electric wave ellipse (see Fig. 2). For any estimator $\hat{\theta}_i, i = 1,2,3$, there is an associated estimated wave frame $\hat{W}$. Define the vector angular error $\phi_{W\hat{W}}$, which is the vector angle by which $\hat{W}$ is (right-handed) rotated about W, and by $[\phi_{W\hat{W}}]_W$ the representation of $\phi_{W\hat{W}}$ in the coordinate system W (see Appendix G). The proposed vector angular error will be $[\phi_{W\hat{W}}]_W$.

Observe that $[\phi_{W\hat{W}}]_W$ depends, by definition, only on the frames $W, \hat{W}$. Thus, for an estimator that is independent of known rotations of the data, the estimated wave frame $\hat{W}$, the vector angular error, and its covariance are independent of the sensor frame. We introduce the following definitions.

Definition 4.4: The *normalized asymptotic covariance of the vector angular error* in the wave frame will be defined as

$$\mathrm{CVAE} \triangleq \lim_{N\to\infty} \{NE([\boldsymbol{\phi}_{W\hat{W}}]_W[\boldsymbol{\phi}_{W\hat{W}}]_W^T)\} \tag{4.32}$$

whenever this limit exists.

Definition 4.5: A direction and orientation estimator will be called *regular* if its errors satisfy $E\,\Sigma_{i=1}^3 |\Delta\theta_i|^3 = o(1/N)$ and the gradient of its bias with respect to $\theta_1, \theta_2, \theta_3$ is $o(1)$ as $N \to \infty$.

Then, we have the following theorems.

Theorem 4.2: For a regular model, the CVAE of any regular direction and orientation estimator, whenever it exists, is bounded from below by

$$\mathrm{CVAE}_{CR} \triangleq N \cdot K\,\mathrm{CRB}(\theta_1, \theta_2, \theta_3)K^T \tag{4.33}$$

where

$$K = \begin{bmatrix} \sin\theta_2 & 0 & -1 \\ -\cos\theta_2\sin\theta_3 & -\cos\theta_3 & 0 \\ \cos\theta_2\cos\theta_3 & -\sin\theta_3 & 0 \end{bmatrix} \tag{4.34}$$

and $\mathrm{CRB}(\theta_1, \theta_2, \theta_3)$ is the Cramér–Rao submatrix bound for the azimuth, elevation, and orientation angles for the particular model used.

Proof: See Appendix G.

Observe that the result of Theorem 4.2 is obtained using geometrical considerations only. Hence, it is applicable to general direction and orientation estimation problems and is not limited to the SST problem only. It is dependent only on the ability to define a wave frame. For example, one can apply this theorem to a DST source with a wave frame defined by the orientation angle that diagonalizes the source signal covariance matrix.

For vector-sensor measurements, CVAE_{CR} has the desirable property of being invariant to the choice of reference coordinate frame. This invariance property holds also for the CVAE of an estimator if the estimate is independent of known rotational transformations of the data. Note that CVAE_{CR} is not a function of N.

Theorem 4.3: The MSAE and CVAE of any regular estimator are related through

$$\mathrm{MSAE} = [\mathrm{CVAE}]_{2,2} + [\mathrm{CVAE}]_{3,3}. \tag{4.35}$$

Furthermore, a similar equality holds for a regular model where the MSAE and CVAE in (4.35) are replaced by their lower bounds MSAE_{CR} and CVAE_{CR}.

Proof: See Appendix G.

In our case, $\mathrm{CRB}(\theta_1, \theta_2, \theta_3)$ is the 3×3 upper left block entry of the CRB matrix in the sensor frame given in Appendix F. Substituting this block entry into (4.33) and denoting the CVAE matrix bound for the SST problem by CVAE^S_{CR}, we have that this matrix is diagonal with nonzero entries given by

$$[\mathrm{CVAE}^S_{CR}]_{1,1} = \frac{(1+\varrho)}{2\varrho^2\cos^2 2\theta_4} \tag{4.36a}$$

$$[\mathrm{CVAE}^S_{CR}]_{2,2} = \frac{(1+\varrho)(\sigma_E^2+\sigma_H^2)}{2\varrho^2[\sigma_H^2\sin^2\theta_4 + \sigma_E^2\cos^2\theta_4]} \tag{4.36b}$$

$$[\mathrm{CVAE}^S_{CR}]_{3,3} = \frac{(1+\varrho)(\sigma_E^2+\sigma_H^2)}{2\varrho^2[\sigma_E^2\sin^2\theta_4 + \sigma_H^2\cos^2\theta_4]}. \tag{4.36c}$$

Some observations on (4.36) are summarized in the following:

- Rotation around $\boldsymbol{u}$: Nonidentifiable only for a circularly polarized signal.
- Rotation around $\tilde{\boldsymbol{v}}_1$ (electric ellipse's major axis): Nonidentifiable only for a linearly polarized signal and no magnetic measurement.
- Rotation around $\tilde{\boldsymbol{v}}_2$ (electric ellipse's minor axis): Nonidentifiable only for a linearly polarized signal and no electric measurement.
- The rotation variances around $\tilde{\boldsymbol{v}}_1$ and $\tilde{\boldsymbol{v}}_2$ are symmetric with respect to the electric and magnetic measurements.
- All the three variances in (4.36) are bounded from below by $(1+\varrho)/2\varrho^2$ (independent of the wave parameters).

The nonidentifiable (or singular) cases above are found by checking when their variances in CVAE^S_{CR} tend to infinity (see, e.g., Theorem 6.3 of [21]). The three nonidentifiable cases above should be expected as the corresponding rotations are unobservable. These singularities are intrinsic to the SST estimation problem and are independent of the reference coordinate system. The symmetry of the variances of the rotations around the major and minor axes of the ellipse with respect to the magnetic and electric measurements should be expected as their axes have a spatial angle difference of $\pi/2$.

The fact that the resulting singularities in the rotational errors are intrinsic (independent of the reference coordinate system) as well as the diagonality of the CVAE^S_{CR} bound matrix with its simple entry expressions indicate that the wave frame is a natural system in which to do the analysis.

Now, consider the augmented error vector

$$\tilde{\boldsymbol{\theta}} \triangleq [[\boldsymbol{\phi}_{W\hat{W}}]^T_W, \Delta\theta_4]^T \tag{4.37}$$

where $\Delta\theta_4$ is the estimation error in θ_4. Combining the results of Appendices F and G, we find that for regular models and estimators, the normalized asymptotic covariance of the error vector $\tilde{\boldsymbol{\theta}}$ is bounded from below by

$$B^S_{CR} \triangleq N \cdot \underline{K}\mathrm{CRB}(\boldsymbol{\theta})\underline{K}^T \tag{4.38}$$

where

$$\underline{K} = \text{block diag}\,\{K, 1\} \tag{4.39}$$

and $\mathrm{CRB}(\boldsymbol{\theta})$ is the 4×4 CRB matrix in the sensor coordinates given in Appendix F. Substituting this matrix in (4.38) and using (4.33) with the fact that CVAE^S_{CR} is diagonal, we find that B^S_{CR} is a diagonal matrix with

$$[B^S_{CR}]_{i,i} = [\mathrm{CVAE}^S_{CR}]_{i,i}, \qquad i = 1,2,3 \tag{4.40a}$$

$$[B^S_{CR}]_{4,4} = \frac{(1+\varrho)}{2\varrho^2}. \tag{4.40b}$$

F. A Cross-Product-Based DOA Estimator

We propose a simple algorithm for estimating the DOA of a single electromagnetic source using the measurements of a single vector sensor. The motivation for this algorithm stems from the average cross-product Poynting vector. Observe that $-\boldsymbol{u}$ is the unit vector in the direction of the Poynting vector given by [22]

$$\begin{aligned} S(t) &= E(t) \times H(t) = \mathrm{Re}\,\{e^{i\omega_c t}\mathcal{E}(t)\} \times \mathrm{Re}\,\{e^{i\omega_c t}\mathcal{H}(t)\} \\ &= \tfrac{1}{2}\mathrm{Re}\,\{\mathcal{E}(t) \times \overline{\mathcal{H}}(t)\} + \tfrac{1}{2}\mathrm{Re}\,\{e^{i2\omega_c t}\mathcal{E}(t) \times \mathcal{H}(t)\} \end{aligned}$$

where $\overline{\mathcal{H}}$ denotes the complex conjugate of $\mathcal{H}$. The carrier time average of the Poynting vector is defined as $\langle S\rangle_t \triangleq \frac{1}{2}\mathrm{Re}\,\{\mathcal{E}(t) \times \overline{\mathcal{H}}(t)\}$. Note that unlike $\mathcal{E}(t)$ and $\mathcal{H}(t)$, this average is not a function of ω_c. Thus, it has an intrinsic physical meaning.

At this point, we can see two possible ways for estimating $\boldsymbol{u}$:

1) Phasor time averaging of $\langle S\rangle_t$ yielding a vector denoted by $\langle S\rangle$ with the estimated $\boldsymbol{u}$ taken as the unit vector in the direction of $-\langle S\rangle$
2) estimation of $\boldsymbol{u}$ by phasor time averaging of the unit vectors in the direction of $\mathrm{Re}\,\{\mathcal{E}(t) \times \overline{\mathcal{H}}(t)\}$.

Clearly, the first way is preferable since then $\boldsymbol{u}$ is estimated after the measurement noise is reduced by the averaging process, whereas the estimated $\boldsymbol{u}$ in the second way is more sensitive to the measurement noises that may be magnified considerably.

Thus, the proposed algorithm computes

$$\hat{\boldsymbol{s}} = \frac{1}{N}\sum_{t=1}^{N} \operatorname{Re}\{\boldsymbol{y}_E(t) \times \bar{\boldsymbol{y}}_H(t)\} \tag{4.41a}$$

$$\hat{\boldsymbol{u}} = \hat{\boldsymbol{s}}/||\hat{\boldsymbol{s}}||. \tag{4.41b}$$

The statistical performance of this estimator $\hat{\boldsymbol{u}}$ is analyzed in Appendix H under the previous assumptions on $\boldsymbol{\xi}(t), \boldsymbol{e}_E(t), \boldsymbol{e}_H(t)$ except that the Gaussian assumption is omitted. The results are summarized by the following theorem.

Theorem 4.4: The estimator $\hat{\boldsymbol{u}}$ has the following properties (for both DST and SST sources):

1) If $||\boldsymbol{\xi}(t)||^2, ||\boldsymbol{e}_E(t)||, ||\boldsymbol{e}_H(t)||$ have finite first-order moments, then $\hat{\boldsymbol{u}} \to \boldsymbol{u}$ almost surely.
2) If $||\boldsymbol{\xi}(t)||^2, ||\boldsymbol{e}_E(t)||, ||\boldsymbol{e}_H(t)||$ have finite second-order moments, then $\sqrt{N}(\hat{\boldsymbol{u}} - \boldsymbol{u})$ is asymptotically normal.
3) If $||\boldsymbol{\xi}(t)||^2, ||\boldsymbol{e}_E(t)||, ||\boldsymbol{e}_H(t)||$ have finite fourth-order moments, then the MSAE is

$$\text{MSAE} = \tfrac{1}{2}\varrho^{-1}(1 + 4\varrho^{-1})(r + r^{-1})^2 \tag{4.42}$$

where $\varrho = \operatorname{tr}(P)/\sigma_{||}^2 = \text{SNR}$.
4) Under the conditions of 3), $N\delta^2$ is asymptotically χ^2 distributed with two degrees of freedom.

Proof: See Appendix H.

For the Gaussian SST case, the ratio between the MSAE of this estimator to the MSAE_{CR}^S in (4.29) is

$$\text{eff} \triangleq \frac{\text{MSAE}}{\text{MSAE}_{CR}^S} = \frac{\varrho + 4}{\varrho + 1}[1 + (r - r^{-1})^2 \sin^2\theta_4 \cos^2\theta_4]. \tag{4.43}$$

Hence, this estimator is nearly efficient if the following two conditions are met

$$\varrho \gg 1 \tag{4.44a}$$

$$r \simeq 1 \quad \text{or} \quad \theta_4 \simeq 0. \tag{4.44b}$$

Fig. 5 illustrates these results using plots of the efficiency factor (4.43) as a function of the ellipticity angle θ_4 for SNR $= \varrho = 10$ and three different values of r.

The estimator (4.41) can be improved using a weighted average of cross products between all possible pairs of real and imaginary parts of $\boldsymbol{y}_E(t)$ and $\boldsymbol{y}_H(s)$ taken at arbitrary times t and s. (Note that these cross products have directions nearly parallel to the basic estimator $\hat{\boldsymbol{u}}$ in (4.41); however, before averaging, these cross products should be premultiplied by +1 or −1 in accordance with the direction of the basic estimator $\hat{\boldsymbol{u}}$.) A similar algorithm suitable for real-time applications can also be developed in the time domain without preprocessing needed for phasor representation. It can be extended to nonstationary inputs by using a moving-average window on the data. It is of interest to find the optimal weights and the performances of these estimators.

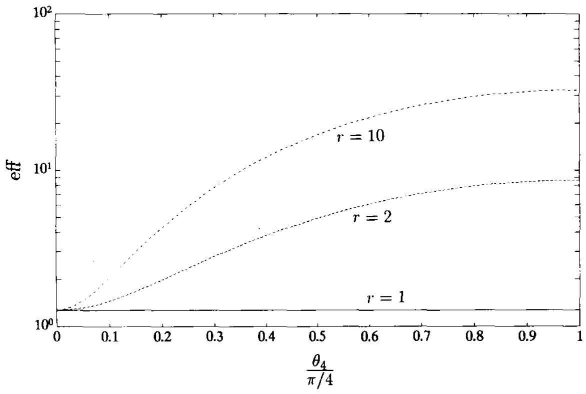


Fig. 5. Efficiency factor (4.43) of the cross-product-based direction estimator as a function of the normalized ellipticity angle for three values of $r = \sigma_H/\sigma_E$. A single source, SNR = 10.

The main advantages of the proposed cross-product-based algorithm (4.41) or one of its variants above are as follows:

- It can give a direction estimate instantly, i.e., with one time sample.
- It is simple to implement (does not require minimization of a cost function) and can be applied in real time.
- It is equally applicable to sources of various types, including SST, DST, wide-band, and non-Gaussian.
- Its MSAE is nearly optimal in the Gaussian SST case under (4.44).
- It does not depend on time delays and therefore does not require data synchronization among different sensor components.

V. Multi-source Multi-vector–Sensor Analysis

Consider the case in which we wish to estimate the directions to multiple electromagnetic sources whose covariance is unknown using an array of vector sensors. The MSAE_{CR} and CVAE_{CR} bound expressions in (4.4) and (4.33) are applicable to each of the sources in the multi-source multi-vector-sensor scenario. Suppose that the noise vector $\boldsymbol{e}_{EH}(t)$ in (2.19) is complex white Gaussian with zero mean and diagonal covariance matrix (i.e., noises from different sensors are uncorrelated) and with electric and magnetic variances σ_E^2 and σ_H^2, respectively. Suppose also that $r = \sigma_H/\sigma_E$ is known. Similar to the single-sensor case, multiply the electric measurements in (2.19) by r to obtain equal noise variances in all the sensor coordinates. The resulting models then become special cases of (3.1) as follows.

For DST signals, the block columns $A_k \in \mathbb{C}^{6m\times 2}$ and the signals $\boldsymbol{x}(t) \in \mathbb{C}^{2n\times 1}$ are

$$A_k = \boldsymbol{e}_k \otimes \begin{bmatrix} rI_3 \\ (\boldsymbol{u}_k \times) \end{bmatrix} V_k \tag{5.1a}$$

$$\boldsymbol{x}(t) = [\boldsymbol{\xi}_1^T(t), \cdots, \boldsymbol{\xi}_n^T(t)]^T]. \tag{5.1b}$$

The parameter vector of the kth source includes here its azimuth and elevation.

For the SST case, the columns $A_k \in \mathbb{C}^{6m\times 1}$ and the signals $\boldsymbol{x}(t) \in \mathbb{C}^{n\times 1}$ are

$$A_k = \boldsymbol{e}_k \otimes \begin{bmatrix} rI_3 \\ (\boldsymbol{u}_k \times) \end{bmatrix} V_k Q_k \boldsymbol{w}_k \tag{5.2a}$$

$$\boldsymbol{x}(t) = [s_1(t), \cdots, s_n(t)]^T. \tag{5.2b}$$

The parameter vector of the kth source includes here its azimuth, elevation, orientation, and ellipticity angles.

The matrices A whose (block) columns are given in (5.1a) and (5.2a) are the Khatri–Rao products (see, e.g., [16] and [17]) of the two matrices whose (block) columns are the arguments of the Kronecker products in these equations.

Mixed single and dual signal transmissions are also special cases of (3.1) with appropriate combinations of the above expressions.

A. Results for Two Sources—Single Vector Sensor

We present some observed behavior of the two-source model and a single vector sensor. It is assumed that the signal and noise vectors satisfy, respectively, assumptions **A5** and **A6**. The results are applicable to wide-band sources since a single vector sensor is used, and thus, **A3** and **A4** are not needed. In general, the analytical expressions involved are found to be intractable, and hence, the results given below are obtained by numerical evaluation, assuming r is known.

The following is a summary of results concerning the localization of *two* uncorrelated sources:

1) Identification of two sources' DOA's is generally possible with *only one* vector sensor.
2) When only the electric field is measured, the DOA's (azimuths and elevations) are nonidentifiable (singular information matrix).
3) When the electric measurement is precise, the CRB variances are generally nonzero.
4) The MSAE^S_{CR} can increase without bound with decreasing source angular separation for sources with the same ellipticity and spin direction, but remarkably, it remains bounded for sources with different ellipticities or opposite spin directions.

Properties 2 and 3 are in general different from the single-source case. Property 2 shows that it is necessary to include both the electric and magnetic measurements to estimate the direction to more than one source. Property 4 demonstrates the great advantage of using the electromagnetic vector sensor in that it allows high resolution of sources with different ellipticities or opposite spins. Note that this generally requires a very large aperture using a scalar-sensor array.

The above result on the ability to resolve two sources that are different only in their ellipticity or spin direction appears to be new. Note also the analogy to Pauli's "exclusion principle." In our case, two narrow-band SST sources are distinguishable if and only if they have different sets of parameters. The set in our case includes wavelength, direction, ellipticity, and spin sign.

Figs. 6 and 7 illustrate the resolution of two uncorrelated equal power SST sources with a single electromagnetic vector sensor. The figures show the square root of the MSAE^S_{CR} of one of the sources for a variety of spin directions, ellipticities, and orientation angles as a function of the separation angle between the sources. (The MSAE^S_{CR} values of the two sources are found to be equal in all the following cases.) The covariances of the signals and noise are normalized such that $P = I_2, \sigma_E = \sigma_H = 1$. The azimuth angle of the first source and the elevation angles of the two sources are kept constant $(\theta_1^{(1)} = \theta_2^{(1)} = \theta_2^{(2)} = 0)$. The second source's azimuth is varied to give the desired separation angle $\Delta\theta_1 \triangleq \theta_1^{(2)}$. In Fig. 6, the cases shown are of same spin directions $(\theta_4^{(1)} = \theta_4^{(2)} = \pi/12)$ and opposite spin directions $(\theta_4^{(1)} = -\theta_4^{(2)} = \pi/12)$, same orientation angles $(\theta_3^{(1)} = \theta_3^{(2)} = \pi/4)$, and different orientation angles $(\theta_3^{(1)} = -\theta_3^{(2)} = \pi/4)$. The figure shows that the resolution of the two sources with a single vector sensor is remarkably good when the sources have opposite spin directions. In particular, the MSAE^S_{CR} remains bounded even for zero separation angle and equal orientation angles. On the other hand, the resolution is not as significant when the two sources have different orientation angles but equal ellipticity angles (then, for example, the MSAE^S_{CR} tends to infinity for zero separation angle). In Fig. 7, the orientation angles of the sources is the same $(\theta_3^{(1)} = \theta_3^{(2)} = \pi/4)$, and the polarization of the first source is kept linear $(\theta_4^{(1)} = 0)$, whereas the ellipticity angle of the second source is varied $(|\theta_4^{(2)}| = \pi/12, \pi/6, \pi/4)$ to illustrate the remarkable resolvability due to different ellipticities. It can be seen that the MSAE^S_{CR} remains bounded here even for zero separation angle.

Thus, Figs. 6 and 7 show that with one vector sensor, it is possible to resolve extremely well two uncorrelated SST sources that have only different spin directions or different ellipticities (these sources can have the same direction of arrival and the same orientation angle). This demonstrates a great advantage of the vector sensor over scalar-sensor arrays in that the latter require large array apertures to resolve sources with small separation angle.

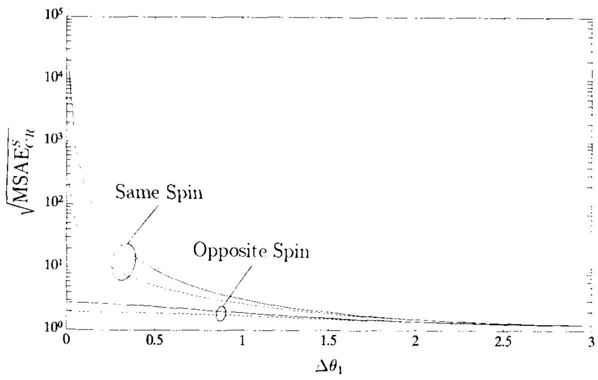


Fig. 6. MSAE^S_{CR} for two uncorrelated equal-power SST sources and a single vector sensor as a function of the source angular separation. Upper two curves: same spin directions $(\theta_4^{(1)} = \theta_4^{(2)} = \pi/12)$; lower two curves: opposite spin directions $(\theta_4^{(1)} = -\theta_4^{(2)} = \pi/12)$; solid curves: same orientation angles $(\theta_3^{(1)} = \theta_3^{(2)} = \pi/4)$; dashed curves: different orientation angles $(\theta_3^{(1)} = -\theta_3^{(2)} = \pi/4)$; remaining parameters are $\theta_1^{(1)} = \theta_2^{(1)} = \theta_2^{(2)} = 0, \Delta\theta_1 \triangleq \theta_1^{(2)}, P = I_2, \sigma_E = \sigma_H = 1$.

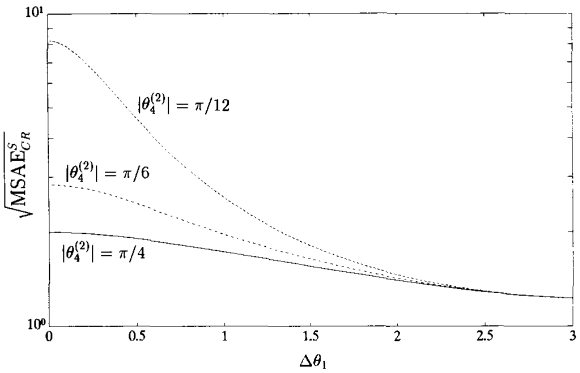


Fig. 7. MSAE_{CR}^S for two uncorrelated equal-power SST sources and a single vector sensor as a function of the source angular separation. Sources are with the same orientation angles ($\theta_3^{(1)} = \theta_3^{(2)} = \pi/4$) and different ellipticity angles ($\theta_4^{(1)} = 0$ and $\theta_4^{(2)}$ as shown in the figure). Remaining parameters are as in Fig. 6.

B. Resolution Capacity and Degrees of Freedom

The maximum number of sources whose directions can be found (the so-called resolution capacity) using a single vector sensor is of great interest.

The results above have shown that it is possible to estimate the DOA (azimuth and elevation) of at least two sources using the instantaneous electromagnetic phasor measurement of a vector sensor. For a single source and and electric measurement only, the electric phasor $\mathcal{E}$ defines an ellipse that determines a plane whose binormal determines the direction to the source. Alternatively, the relationship $\boldsymbol{u} \cdot \mathcal{E} = 0$ gives two equations in its real and imaginary parts that can be solved for θ_1 and θ_2 if the source is not linearly polarized. Note that with scalar sensors, it is necessary to use a plane array for the same purpose.

A preliminary heuristic discussion on the resolution capacity of a single sensor can also be based on the fact that the total number of scalar measurements (or degrees of freedom) of the instantaneous electric and magnetic phasors $\mathcal{E}$ and $\mathcal{H}$ is 12. Since each DOA has two parameters, this gives an upper bound of six sources for DOA estimation per vector sensor for one time sample. If also the complex signal of each source has to be estimated, the upper bound is reduced to three SST sources or two DST sources per vector sensor. The determination of the exact resolution capacity for each case and for the multi-vector-sensor case is left for future research.

VI. Concluding Remarks

A new approach for the localization of electromagnetic sources using vector sensors has been presented. We summarize some of the main results of the paper and give an outlook to their possible extensions.

Models: New models that include the complete electromagnetic data at each sensor have been introduced. Furthermore, new signal models and vector angular error models in the wave frame have been proposed. The wave frame model provides simple performance expressions that are easy to interpret and have only intrinsic singularities. Extensions of the proposed models may include additional structures for specific applications.

Cramér–Rao Bounds and Quality Measures: A compact expression for the CRB for multi-vector-source multi-vector-sensor processing has been derived. The derivation gave rise to new block matrix operators. New quality measures in 3-D space, such as the MSAE for direction estimation and CVAE for direction and orientation estimation, have been defined. Explicit bounds on the MSAE and CVAE, having the desirable property of being invariant to the choice of the reference coordinate frame, have been derived and can be used for performance analysis. These bounds are not limited to electromagnetic vector-sensor processing. Performance comparisons of vector-sensor processing with scalar-sensor counterparts are of interest.

Identifiability: The derived bounds were used to show that the fusion of magnetic and electric data at a single vector sensor increases the number of identifiable sources (or resolution capacity) in 3-D space from one source in the electric data case to at least two sources in the electromagnetic case. For a single signal transmission source, in order to get good direction estimates, the fusion of the complete data becomes more important as the polarization gets closer to linear. Finding the number of identifiable sources per sensor in a general vector-sensor array is of interest. This can be done, for example, by extending the results in [23].

Resolution: Source resolution using vector sensors is inherently different from scalar sensors, where the latter case is characterized by the classical Rayleigh principle. For example, it was shown that a single vector sensor can be used to resolve two sources in 3-D space. In particular, a vector sensor exhibits remarkable resolvability when the sources have opposite spin directions or different ellipticity angles. This is very different from the scalar-sensor array case in which a plane array with large aperture is required to achieve the same goal. Analytical results on source resolution using vector-sensor arrays and comparisons with their scalar counterparts are of interest.

Algorithms: A simple algorithm has been proposed and analyzed for finding the direction to a single source using a single vector sensor based on the cross-product operation. It is of interest to analyze the performance of the aforementioned variants of this algorithm and to extend them to more general source scenarios (e.g., larger number of sources). It is also of interest to develop new algorithms for the vector-sensor array case.

Communication: The main considerations in communication are transmission of signals over channels with limited bandwidth and their recovery at the sensor. Vector sensors naturally fit these considerations as they have maximum observability to incoming signals, and they can be used to recover DST signals. Future goals will include development of optimum signal estimation algorithms, communication forms, and coding design with vector sensors.

Implementations: The proposed methods should be implemented and tested with real data.

Sensor Development: The use of complete electromagnetic data seems to be virtually nonexistant in the literature on source localization. It is hoped that the results of this research will motivate the systematic development of high-quality electromagnetic vector sensors. A recent reference on this topic can be found in [24].

Other Extensions: The vector-sensor concept can be extended to other areas and open new possibilities. An example of this can be found in [25] for the acoustic case.

APPENDIX A
PROOF OF THEOREM 2.1

We prove first that, for plane waves, Maxwell's equations can be reduced to a system of two equations without any loss of information.

Under the assumption of nonconductive, homogeneous, and isotropic medium, the Maxwell equations become (see e.g., [22])

$$\nabla \times E = -\mu \frac{\partial H}{\partial t} \tag{A.1a}$$

$$\nabla \times H = \varepsilon \frac{\partial E}{\partial t} \tag{A.1b}$$

$$\nabla \cdot E = \varepsilon^{-1} \rho \tag{A.1c}$$

$$\nabla \cdot H = 0 \tag{A.1d}$$

where E and H are, respectively, the electric and magnetic fields and ε and μ are the permittivity and permeability of the medium.

Let c be the velocity of wave propagation in the medium and $\boldsymbol{\kappa}$ be the unit vector in the direction of the wave propagation. Then, under assumption **A1** we have

$$E(\boldsymbol{r}, t) = E(0, t-\tau), \qquad H(\boldsymbol{r}, t) = H(0, t-\tau) \tag{A.2}$$

where $\tau = (\boldsymbol{\kappa} \cdot \boldsymbol{r})/c$ and $\boldsymbol{r}$ is the location vector relative to the reference coordinate frame; thus τ is constant in a plane. Equation (A.2) is equivalent to a constant delay of wave in a plane. Let $\tilde{E} \triangleq E(0,t)$ and $\tilde{H}(t) \triangleq H(0,t)$. Then $E(\boldsymbol{r},t) = \tilde{E}(t-\tau)$ and $H(\boldsymbol{r},t) = \tilde{H}(t-\tau)$.

Equation (A.2) shows that for plane waves the operator ∇ is equivalent to $-(\boldsymbol{\kappa}/c)(\partial/\partial t)$, hence (A.1) can now be written

$$-\frac{\boldsymbol{\kappa}}{c} \times \dot{\tilde{E}} = -\mu \dot{\tilde{H}} \tag{A.3a}$$

$$-\frac{\boldsymbol{\kappa}}{c} \times \dot{\tilde{H}} = \varepsilon \dot{\tilde{E}} \tag{A.3b}$$

$$-\frac{\boldsymbol{\kappa}}{c} \cdot \dot{\tilde{E}} = \varepsilon^{-1} \rho \tag{A.3c}$$

$$-\frac{\boldsymbol{\kappa}}{c} \cdot \dot{\tilde{H}} = 0 \tag{A.3d}$$

where $\dot{\tilde{E}} \triangleq d\tilde{E}/dt$ and $\dot{\tilde{H}} \triangleq d\tilde{H}/dt$.

Using (A.3b) and (A.3d), we get, respectively, $\dot{\tilde{E}} = -(\varepsilon c)^{-1} \boldsymbol{\kappa} \times \dot{\tilde{H}}$ and $\boldsymbol{\kappa} \cdot \dot{\tilde{H}} = 0$, hence $(\dot{\tilde{E}}, \dot{\tilde{H}}, \boldsymbol{\kappa})$ is a right orthogonal triad. Using (A.3c), it then follows that $\rho = 0$. In addition, from (A.3b) and (A.3a) we obtain

$$\dot{\tilde{E}} = -(\varepsilon c)^{-1} \boldsymbol{\kappa} \times \dot{\tilde{H}} = -(\varepsilon \mu c^2)^{-1} \boldsymbol{\kappa} \times (\boldsymbol{\kappa} \times \dot{\tilde{E}}) = (\varepsilon \mu c^2)^{-1} \dot{\tilde{E}} \tag{A.4}$$

where the last equality follows from the orthogonality of $\boldsymbol{\kappa}$ and $\dot{\tilde{E}}$. Therefore, under the assumption that $\tilde{E}$ varies with time, we have $c = (\varepsilon\mu)^{-1/2}$. Denoting $\eta \triangleq (\mu/\varepsilon)^{1/2}$, i.e., η is the so-called intrinsic impedance of the medium, (A.3) may now be written

$$\boldsymbol{\kappa} \times \dot{\tilde{E}} = \eta \dot{\tilde{H}} \tag{A.5a}$$

$$\boldsymbol{\kappa} \times \eta \dot{\tilde{H}} = -\dot{\tilde{E}} \tag{A.5b}$$

$$\boldsymbol{\kappa} \cdot \dot{\tilde{E}} = 0 \tag{A.5c}$$

$$\boldsymbol{\kappa} \cdot \dot{\tilde{H}} = 0. \tag{A.5d}$$

Clearly, (A.5a) and (A.5c) are equivalent to (A.5b) and (A.5d). Thus, under the assumptions of a plane wave at the sensor and that $\tilde{E}$ varies with time, it follows that the Maxwell equations are equivalent to either of these sets of two equations, i.e., they can be reduced to either set without loss of information.

Now consider the reduction of the Maxwell equations under the additional assumption A2 of band-limited waves.

Let $\omega_c > 0$ and let $E_+(t)$ be the signal corresponding to the positive frequency component of $\tilde{E}(t)$. The phasor representation or the complex envelope (see e.g. [12], [13]) of $\tilde{E}(t)$ with respect to ω_c is defined by

$$\mathcal{E}(t) \triangleq e^{-i\omega_c t} \cdot 2E_+(t). \tag{A.6}$$

Since $\tilde{E}(t)$ is real it can be recovered from $\mathcal{E}(t)$ using $\tilde{E}(t) = \mathrm{Re}\{e^{i\omega_c t}\mathcal{E}(t)\}$ where $\mathrm{Re}\{\cdot\}$ denotes the real part of its argument. A sufficient condition for the uniqueness of $\mathcal{E}(t)$ in the representation of $\tilde{E}(t)$ is that the spectrum of $\mathcal{E}(t)$ is contained in $(-\omega_c, \infty)$.

Recall that for plane waves it is sufficient to consider only (A.5a) and (A.5c). From (A.5a) it is clear that

$$\boldsymbol{\kappa} \times \tilde{E}(t) = \eta \tilde{H}(t) + \text{const} \tag{A.7}$$

where the constant term is zero from assumption **A1** or **A2**. Hence, $\tilde{E}(t)$ and $\tilde{H}(t)$ contain the same spectral components. Therefore, applying the phasor operator as in (A.6) to both sides of (A.7) and applying the same approach to (A.5c), we get the phasor representations (2.2a), (2.2b), which completes the proof of the theorem. ∎

APPENDIX B
PROOF OF LEMMA 2.1

For notational convenience it will be useful to rewrite the lemma with a slight change of notation.

Lemma 2.1: Every vector $\boldsymbol{\xi} = [\xi_1, \xi_2]^T =\in \mathbb{C}^2$ has the representation

$$\boldsymbol{\xi} = re^{i\varphi} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta \\ i\sin\beta \end{bmatrix} \tag{B.1}$$

where $r \geq 0, \varphi \in (-\pi, \pi], \alpha \in (-\pi/2, \pi/2], \beta \in [-\pi/4, \pi/4]$. Moreover, $r, \varphi, \alpha, \beta$ in (B.1) are uniquely determined if and only if $\xi_1^2 + \xi_2^2 \neq 0$.

Proof: Let $\xi_1 + i\xi_2 = r_1 e^{i\varphi_1}, \xi_1 - i\xi_2 = r_2 e^{i\varphi_2}$, for some $r_1, r_2 \geq 0$ and $\varphi_1, \varphi_2 \in (-\pi, \pi]$. Take $r \geq 0$ and $\beta_1 \in [0, \pi/2]$ such that $r_1 + ir_2 = \sqrt{2} r e^{i\beta_1}$, and define $\beta \triangleq \beta_1 - \pi/4 \in [-\pi/4, \pi/4]$. Define $\varphi_3 \in (-\pi, \pi]$ to be the argument of $e^{i(\varphi_2 - \varphi_1)}$, and take $\alpha \triangleq \frac{1}{2}\varphi_3 \in (-\pi/2, \pi/2]$ and define $\varphi \in (-\pi, \pi]$ to be the argument of $e^{i(\varphi_1 + \alpha)}$.

Then, $\xi_1 + i\xi_2 = r_1 e^{i\varphi_1} = \sqrt{2}r\cos\beta_1 \cdot e^{i(\varphi-\alpha)} = \sqrt{2}r\cos(\beta + \pi/4)e^{i(\varphi-\alpha)} = re^{i(\varphi-\alpha)}(\cos\beta - \sin\beta)$. Similarly, $\xi_1 - i\xi_2 = r_2 e^{i\varphi_2} = \sqrt{2}r\sin\beta_1 \cdot e^{i(\varphi_1+\varphi_3)} = \sqrt{2}r\sin(\beta+\pi/4)e^{i(\varphi_1+2\alpha)} = re^{i(\varphi+\alpha)}(\cos\beta+\sin\beta)$.

Thus

$$\begin{aligned} re^{i\varphi}\begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}\begin{bmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}\cos\beta\\ i\sin\beta\end{bmatrix} &= re^{i\varphi}\begin{bmatrix}e^{-i\alpha}(\cos\beta-\sin\beta)\\ e^{i\alpha}(\cos\beta+\sin\beta)\end{bmatrix}\\ &= \begin{bmatrix}\xi_1+i\xi_2\\ \xi_1-i\xi_2\end{bmatrix} = \begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}\begin{bmatrix}\xi_1\\ \xi_2\end{bmatrix} \end{aligned} \tag{B.2}$$

which gives (B.1).

To prove the uniqueness part, suppose that $\xi_1^2 + \xi_2^2 \neq 0$ and (B.1) holds for some $r, \varphi, \alpha, \beta$. Thus $\xi_1 + i\xi_2 \neq 0$ and $\xi_1 - i\xi_2 \neq 0$.

Using (B.1) we have in the same manner

$$\begin{bmatrix}\xi_1+i\xi_2\\ \xi_1-i\xi_2\end{bmatrix} = r\begin{bmatrix}e^{i(\varphi-\alpha)}(\cos\beta-\sin\beta)\\ e^{i(\varphi+\alpha)}(\cos\beta+\sin\beta)\end{bmatrix}. \tag{B.3}$$

Thus, $\xi_1 + i\xi_2 = r(\cos\beta - \sin\beta)e^{i(\varphi-\alpha)} = r_1e^{i\varphi_1}$ and $\xi_1 - i\xi_2 = r(\cos\beta+\sin\beta)e^{i(\varphi+\alpha)} = r_2e^{i\varphi_2}$. Hence, $r(\cos\beta - \sin\beta) = r_1 > 0$ and $r(\cos\beta + \sin\beta) = r_2 > 0$ and also $e^{i(\varphi-\alpha)} = e^{i\varphi_1}, e^{i(\varphi+\alpha)} = e^{i\varphi_2}$. We have $r_1^2 + r_2^2 = r^2[(\cos\beta-\sin\beta)^2 + (\cos\beta+\sin\beta)^2] = 2r^2$, so r is uniquely determined. Also, $\cos\beta = (r_1+r_2)/2r, \sin\beta = (r_2-r_1)/2r$, so β is uniquely determined in $[-\pi/4, \pi/4]$. In addition, $e^{i(\varphi_2-\varphi_1)} = e^{2i\alpha}$, so α is uniquely determined in $(-\pi/2, \pi/2]$. Also $e^{i\varphi} = e^{i(\varphi_1+\alpha)}$, so φ is uniquely determined.

To prove the "only if" part, observe that if $\xi_1^2 + \xi_2^2 = 0$, there are three cases to be checked. Case 1: $\xi_1 + i\xi_2 = 0$ and $\xi_1 - i\xi_2 \neq 0$. In this case, $r_1 = 0, \beta = \pi/4$, and $\varphi - \alpha$ is not determined. Case 2: $\xi_1 + i\xi_2 \neq 0$ and $\xi_1 - i\xi_2 = 0$. In this case, $r_2 = 0, \beta = -\pi/4$, and $\varphi + \alpha$ is not determined. Case 3: $\xi_1 + i\xi_2 = \xi_1 - i\xi_2 = 0$. In this case, $r = 0$ and β, φ, α are not determined.

Appendix C
Proof of Theorem 3.1

To find the CRB of $\boldsymbol{\theta}$ when $\boldsymbol{\theta}, P, \sigma^2$ are unknown in (3.1), we use the notion of the concentrated log-likelihood function. The normalized (i.e., multiplied by $-1/N$) log likelihood function for (3.1) is given (up to an additive constant independent of $\boldsymbol{\theta}, P, \sigma^2$) by

$$L(\boldsymbol{\theta}, \sigma^2, P) \triangleq \ln\det R + \operatorname{tr}[R^{-1}\hat{R}] \tag{C.1}$$

where R is defined in (3.6) and

$$\hat{R} = \frac{1}{N}\sum_{t=1}^{N} y(t)y^*(t). \tag{C.2}$$

Under the restrictions that $\det(A^*A) > 0, R > 0$ (where P is assumed to be Hermitian), the unique global minimizer of $L(\boldsymbol{\theta}, P, \sigma^2)$ with respect to P and σ^2 is given by

$$\tilde{P} = (A^*A)^{-1}A^*\hat{R}A(A^*A)^{-1} - \tilde{\sigma}^2(A^*A)^{-1} \tag{C.3a}$$

$$\tilde{\sigma}^2 = \frac{1}{\overline{\mu} - \overline{\nu}}\operatorname{tr}[\Pi_c\hat{R}] \tag{C.3b}$$

(see e.g., [26]–[29]). The normalized log-likelihood function concentrated with respect to P and σ^2 is given by

$$F(\boldsymbol{\theta}) \triangleq L(\boldsymbol{\theta}, \tilde{\sigma}^2, \tilde{P}) = \ln\det[\Pi\hat{R}\Pi + \tilde{\sigma}^2\Pi_c] + \overline{\mu}. \tag{C.4}$$

Next, to avoid certain technical assumptions related only to the method of proof used in [30] (e.g., existence of Hessian of the log-likelihood function), we need the following two lemmas.

Lemma C.1: Under the assumptions of Theorem 3.1 and the additional assumption that $A(\boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$, the inverse Cramér–Rao bound matrix for $\boldsymbol{\theta}$, when $\boldsymbol{\theta}, P, \sigma^2$ are unknown in (3.1), is given (elementwise) by

$$[\mathrm{CRB}^{-1}(\boldsymbol{\theta})]_{rs} = \frac{2N}{\sigma^2}\operatorname{Re}\left\{\operatorname{tr}\left[U\frac{\partial A^*}{\partial\theta_s}\Pi_c\frac{\partial A}{\partial\theta_r}\right]\right\} \tag{C.5}$$

where U and Π_c are defined in (3.5) and $1 \leq r, s \leq \overline{q}$. (An implicit assumption is that the matrix whose entries appear on the r.h.s. of (C.5) is nonsingular.)

Proof: It is known (see [30]) that under the above assumptions the following limit

$$F_0''(\boldsymbol{\theta}) \triangleq \lim_{N\to\infty}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}F(\boldsymbol{\theta}) \tag{C.6}$$

exists almost surely and is a deterministic p.s.d. matrix (the primes symbolically denote derivatives with respect to $\boldsymbol{\theta}$). From [4], the (r, s) entry of $F_0''(\boldsymbol{\theta})$ is given by N^{-1} times the r.h.s. of (C.5) (thus by the above implicit assumption $F_0''(\boldsymbol{\theta})$ is nonsingular). Using (C.3), the *linear* parametrization (3.6) of R by P, σ^2 is identifiable, thus it has a full-rank matrix representation and by [21, Lemma 6.1, p. 601] it has a positive-definite information matrix (w.r.t. P, σ^2), and by [30] the CRB for $\boldsymbol{\theta}$ when $\boldsymbol{\theta}, P, \sigma^2$ are unknown is given, for a fixed N, through

$$\mathrm{CRB}^{-1}(\boldsymbol{\theta}) = N \cdot F_0''(\boldsymbol{\theta}). \tag{C.7}$$

This gives the result. ∎

Lemma C.2: Lemma C.1 remains true without the assumption that $A(\boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$.

Proof: We use the fact that from a well-known formula (see e.g., [31]) that holds under the assumptions of Theorem 3.1, the dependence of the Cramér–Rao bound on $\boldsymbol{\theta}$ is only through $A(\boldsymbol{\theta})$ and its Jacobian $\partial A(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ at the true parameter vector. Let $\tilde{A}(\boldsymbol{\theta})$ be the linear interpolation of $A(\boldsymbol{\theta})$ and its Jacobian $\partial A(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ at the true $\boldsymbol{\theta}$ (i.e., $\tilde{A}(\boldsymbol{\theta})$ is the first-order Taylor expansion of $A(\boldsymbol{\theta})$ at the true $\boldsymbol{\theta}$). Using the above fact, the two models given by $A(\boldsymbol{\theta})$ and $\tilde{A}(\boldsymbol{\theta})$ have the same CRB. From Lemma C.1 (applied to $\tilde{A}(\boldsymbol{\theta})$), $\mathrm{CRB}(\boldsymbol{\theta})$ is given by a modified (C.5) in which $A(\boldsymbol{\theta})$ is replaced by $\tilde{A}(\boldsymbol{\theta})$. The result then follows from the fact that the dependence of this last expression on $\boldsymbol{\theta}$ is only through $\tilde{A}(\boldsymbol{\theta})$ and its Jacobian at the true $\boldsymbol{\theta}$.

Proof of the Theorem: Let $\theta_l^{(i)}$ denotes the lth entry of $\boldsymbol{\theta}^{(i)}, 1 \leq i \leq n$ and $1 \leq l \leq q_i$. Observe that

$$\frac{\partial A}{\partial\theta_l^{(i)}} = [\underbrace{0\cdots 0}_{\sum_{k=1}^{i-1}\nu_k}\underbrace{D_l^{(i)}}_{\nu_i}\underbrace{0\cdots 0}_{\sum_{k=i+1}^{n}\nu_k}] \tag{C.8}$$

where

$$D_l^{(i)} \triangleq \frac{\partial A_i}{\partial \theta_l^{(i)}}. \tag{C.9}$$

Using Lemma C.1, we extend the definition (C.6) for $F_0''(\boldsymbol{\theta})$ by redefining it as N^{-1} times the matrix whose entries appear on the r.h.s. of (C.5). As a result $F_0''(\boldsymbol{\theta})$ is still defined even if the Hessian of $F(\boldsymbol{\theta})$ does not exist. From Lemma C.2, it remains to cast $F_0''(\boldsymbol{\theta})$ in a form that implies (3.4), which in turn implies (by assumption **A7**) that the implicit assumption of Lemma C.1 holds.

Let $[F_0'']_{\langle ij\rangle} \in \mathbb{C}^{q_i \times q_j}$ denote the (i,j) block entry of $F_0''(\boldsymbol{\theta})$. Then, the (l,p) entry of this matrix is

$$\begin{aligned}[F_0'']_{\langle ij\rangle lp} &= \frac{2}{\sigma^2}\operatorname{Re}\left[\operatorname{tr}\left[U\frac{\partial A^*}{\partial \theta_p^{(j)}}\Pi_c\frac{\partial A}{\partial \theta_l^{(i)}}\right]\right\} \\ &= \frac{2}{\sigma^2}\operatorname{Re}\{\operatorname{tr}[U_{\langle ij\rangle}D_p^{(j)*}\Pi_c D_l^{(i)}]\} \end{aligned}\tag{C.10}$$

where $1 \le p \le q_j$ and $U_{\langle ij\rangle} \in \mathbb{C}^{\nu_i \times \nu_j}$ denotes the (i,j) block entry of U with $1 \le i,j \le n$. Observe that $D_l^{(i)}$ is the $(\Sigma_{k=1}^{i-1} q_k + l)$th block column of D, cf. (3.5d), or notationally

$$D_l^{(i)} = D_{\Sigma_{k=1}^{i-1} q_k + l}. \tag{C.11}$$

Hence

$$\begin{aligned}[F_0'']_{\langle ij\rangle lp} = \frac{2}{\sigma^2}\operatorname{Re}\{&\operatorname{tr}[U_{\langle ij\rangle}(D_{\Sigma_{k=1}^{j-1} q_k+p})^* \\ &\cdot \Pi_c(D_{\Sigma_{k=1}^{i-1} q_k + l})]\}.\end{aligned} \tag{C.12}$$

Assume now that $\nu_i = \nu$ for all i, then collecting terms we have that

$$[F_0'']_{\langle ij\rangle} = \operatorname{btr}[(1_{\langle ij\rangle} \otimes U_{\langle ij\rangle}) \boxdot ((D^*\Pi_c D)_{\langle ji\rangle})^{bT}] \tag{C.13}$$

where $\otimes$ is the Kronecker product, and the block trace operator btr, the block Schur-Hadamard product $\boxdot$ and the block transpose operator bT are as defined in Appendix I, with blocks of dimension $\nu \times \nu$, except for $1_{\langle ij\rangle}$, which is a $q_i \times q_j$ matrix with all entries equal to one. Note that the matrices $(1_{\langle ij\rangle} \otimes U_{\langle ij\rangle})$ and $(D^*\Pi_c D)_{\langle ji\rangle}$ are of dimensions $q_i\nu \times q_j\nu$. Collecting the above block results and using the block Kronecker product definition of Appendix I, we obtain the desired matrix expression (3.4). The proof of the last statement of the theorem (i.e., $\operatorname{CRB}(\boldsymbol{\theta})$ remains the same independently of whether σ^2 is known or unknown) follows in the same manner for the case in which σ^2 is known. ■

Appendix D
Extension to Unknown Sensor Noise Covariance

The results of Theorem 3.1 can be extended to a larger class of unknown (structured) sensor noise covariance matrices as follows. Suppose that the model is

$$\tilde{\boldsymbol{y}}(t) = \tilde{A}(\boldsymbol{\theta})\boldsymbol{x}(t) + \tilde{e}(t) \tag{D.1}$$

where the assumptions are as above, except that the noise covariance is

$$E\tilde{e}(t)\tilde{\boldsymbol{e}}^*(s) = \Sigma\delta_{t,s} \tag{D.2}$$

and Σ is a positive definite matrix. Assume that there exists a matrix $\Gamma \in \mathbb{C}^{m\times m}$ such that

$$\Gamma\Sigma\Gamma^* = \sigma^2 I_m \tag{D.3}$$

for some unknown parameter σ^2, and Γ is parametrized by an unknown parameter vector γ that is assumed to have the local identifiability property (see e.g., [21]). From the Gaussian assumption, the negative log-likelihood function of $\tilde{\boldsymbol{y}}(t)$, up to additive and multiplicative constants is

$$\begin{aligned}\tilde{L}(\boldsymbol{\theta},\boldsymbol{\gamma},\sigma^2,P) &\triangleq \operatorname{tr}[\hat{\tilde{R}}(\tilde{A}P\tilde{A}^* + \Sigma)^{-1}] \\ &\quad + \ln\det[\tilde{A}P\tilde{A}^* + \Sigma] \\ &= \operatorname{tr}[\hat{R}(APA^* + \sigma^2 I)^{-1}] \\ &\quad + \ln\det[APA^* + \sigma^2 I] - \ln\det[\Gamma\Gamma^*]\end{aligned} \tag{D.4}$$

where $\hat{\tilde{R}} = (1/N)\Sigma_{t=1}^N \tilde{\boldsymbol{y}}(t)\tilde{\boldsymbol{y}}^*(t)$ and

$$A = \Gamma\tilde{A} \qquad \hat{R} = \Gamma\hat{\tilde{R}}\Gamma^*. \tag{D.5}$$

The first two terms on r.h.s. of (D.4) have a similar form to the corresponding ones in (C.1). The last term in (D.4) is a function of $\boldsymbol{\gamma}$ only. Hence, we can concentrate (D.4) with respect to P and σ^2 as is done in [26]–[29] for $\Sigma = \sigma^2 I$. The result is

$$\tilde{F}(\boldsymbol{\theta},\boldsymbol{\gamma}) = \ln\det[\Pi\hat{R}\Pi + \hat{\sigma}^2\Pi_c] - \ln\det[\Gamma\Gamma^*] \tag{D.6}$$

where the variables in the first term on the r.h.s of (D.6) are the same as for $\Sigma = \sigma^2 I$ but with A and $\hat{R}$ given in (D.5). Observe that all the terms (in particular, $\hat{R}$) in (D.6) depend on the unknown vector $\boldsymbol{\gamma}$. When $\boldsymbol{\gamma}$ is known, Theorem 3.1 can be applied to find $\operatorname{CRB}(\boldsymbol{\theta})$. When $\boldsymbol{\gamma}$ is unknown, the joint CRB of $\boldsymbol{\theta},\boldsymbol{\gamma}$ is given through the deterministic a.s. limit of the Hessian of the r.h.s. of (D.6), see [30] and Appendix C. (The submatrix of $[\operatorname{CRB}(\boldsymbol{\theta},\boldsymbol{\gamma})]^{-1}$ corresponding to $\boldsymbol{\theta}$ remains the same as for the case with $\boldsymbol{\gamma}$ known.)

Appendix E
Proof of Lemma 4.1

For any Hermitian matrix P with unit-norm eigenvectors $\boldsymbol{v}_i$ and eigenvalues λ_i, we have

$$f(P) = \sum_i f(\lambda_i)\boldsymbol{v}_i\boldsymbol{v}_i^* \tag{E.1}$$

for any f defined on the spectrum of P. In our case, rank $P = 1$ and hence P has only a single nonzero eigenvalue $\lambda_1 = \operatorname{tr} P$ with an eigenvector $\boldsymbol{v}_1$. With $f(0) = 0$ it is found that

$$f(P) = f(\lambda_1)\boldsymbol{v}_1\boldsymbol{v}_1^* + \sum_{i>1} f(\lambda_i)\boldsymbol{v}_i\boldsymbol{v}_i^* = f(\lambda_1)\boldsymbol{v}_1\boldsymbol{v}_1^*. \tag{E.2}$$

Using $P = \Sigma_i \lambda_i\boldsymbol{v}_i\boldsymbol{v}_i^* = \lambda_1\boldsymbol{v}_1\boldsymbol{v}_1^*$ we get

$$f(P) = \frac{f(\lambda_1)}{\lambda_1}P. \tag{E.3}$$

But since $\lambda_1 = \operatorname{tr} P$ we get the desired result (4.25). ■

APPENDIX F
CRB FOR AN SST SOURCE (SST MODEL)

This appendix presents the entries of $\mathrm{CRB}(\boldsymbol{\theta})$ for the special case of an SST source, SST model, and a single vector sensor, see (4.30). These entries were calculated using the symbolic software program MACSYMA. For notational simplicity, the entries of a normalized matrix C defined through $\mathrm{CRB}(\boldsymbol{\theta}) = [(1+\varrho)(1+r^2)/2N\varrho^2]C$ are presented. Also, since this matrix is symmetric only the upper triangular entries are shown in the equation at the bottom of the page.

APPENDIX G
PROOFS OF THEOREMS 4.2 AND 4.3

A. Background

For ease of reference, it is useful to review some basic principles on rotation of coordinate systems in $\mathbb{R}^3$ needed for the proofs. For more details, see, for example, [32], [33].

Fact 1: To every nonsingular matrix $A \in \mathbb{R}^{3\times 3}$, there is a corresponding coordinate system, denoted also by A, such that every $\boldsymbol{v} \in \mathbb{R}^{3\times 1}$ has its representation in A given by $[\boldsymbol{v}]_A = A^{-1}\boldsymbol{v}$. If $A, B \in \mathbb{R}^{3\times 3}$ are nonsingular, then $[\boldsymbol{v}]_A = C_B^A[\boldsymbol{v}]_B$ where $C_B^A \triangleq A^{-1}B$ serves as a translation matrix from vector representation in B to that in A. Note that the columns of C_B^A are the columns of B represented in A, i.e., $C_B^A = [B]_A$. For nonsingular $A, B, C, \in \mathbb{R}^{3\times 3}$, we have the composition rule $C_A^C = C_B^C C_A^B$.

Definition: Given nonsingular $A, B \in \mathbb{R}^{3\times 3}$, we say that the system A *can be rotated* to B if C_A^B is orthogonal and $\det C_A^B = 1$. The rotation of $\boldsymbol{v} \in \mathbb{R}^{3\times 1}$ by a *vector angle* $\boldsymbol{\phi} \in \mathbb{R}^{3\times 1}$ is defined as follows: If $\boldsymbol{\phi} = 0$, then the rotation of $\boldsymbol{v}$ is $\boldsymbol{v}$. If $\boldsymbol{\phi} \neq 0$ then $\boldsymbol{v}$ is right-handed rotated around the axis $\hat{\boldsymbol{\phi}} \triangleq |\boldsymbol{\phi}|^{-1}\boldsymbol{\phi}$ by an angle $|\boldsymbol{\phi}|$. We say that B is a rotation of A by a vector angle ϕ if the columns of B are the rotations of the corresponding columns of A by ϕ. For $\boldsymbol{\phi}, \boldsymbol{v} \in \mathbb{R}^{3\times 1}$, the matrix form of the cross-product operator $\boldsymbol{v} \to \boldsymbol{\phi} \times \boldsymbol{v}$ is defined by $(\boldsymbol{\phi}\times)$.

Fact 2: .If A can be rotated to B, then B is a rotation of A by a vector angle ϕ_{AB} such that $|\boldsymbol{\phi}_{AB}| \leq \pi$ and

$$C_B^A = e^{([\boldsymbol{\phi}_{AB}]_A\times)}. \tag{G.1}$$

It follows that $[\boldsymbol{\phi}_{AB}]_A = C_B^A[\phi_{AB}]_B = [\boldsymbol{\phi}_{AB}]_B$ and $\mathrm{tr}(C_B^A) = 1 + 2\cos|\boldsymbol{\phi}_{AB}|$. Moreover, if $\mathrm{tr}(C_B^A) > -1$, then $|\boldsymbol{\phi}_{AB}| < \pi$ and $\boldsymbol{\phi}_{AB}$ is uniquely determined, and if $\mathrm{tr}(C_B^A) = -1$ then $|\boldsymbol{\phi}_{AB}| = \pi$ and $\boldsymbol{\phi}_{AB}$ is determined up to a multiplication by -1. Thus, the matrix $[\boldsymbol{\phi}_{AB}]_A([\boldsymbol{\phi}_{AB}]_A)^T$ is uniquely determined (by A and B or C_B^A) in all cases.

B. Proof of Theorem 4.2

Denote by I the sensor reference frame, by A the frame $(\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2)$, and by W the wave frame $(\boldsymbol{u}, \tilde{\boldsymbol{v}}_1, \tilde{\boldsymbol{v}}_2)$. Let C_A^I be the translation matrix from a vector representation in A to that in I, and similarly let C_W^A be the translation matrix from W to A. We have

$$C_A^I = [\boldsymbol{u} \vdots \boldsymbol{v}_1 \vdots \boldsymbol{v}_2] = \begin{bmatrix} \cos\theta_1\cos\theta_2 & -\sin\theta_1 & -\cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 & \cos\theta_1 & -\sin\theta_1\sin\theta_2 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{bmatrix} \tag{G.2a}$$

$$C_{1,1} = \frac{(1-r^2)\{(1-2\cos^2\theta_3)\sin^2\theta_4 + \cos^2\theta_3\} + r^2}{[(1-r^2)^2\sin^2\theta_4\cos^2\theta_4 + r^2]\cos^2\theta_2} \tag{F.1a}$$

$$C_{1,2} = -\frac{(1-r^2)(1-2\sin^2\theta_4)\sin\theta_3\cos\theta_3}{[(1-r^2)^2\sin^2\theta_4\cos^2\theta_4 + r^2]\cos\theta_2} \tag{F.1b}$$

$$C_{1,3} = \frac{\{(1-r^2)[(1-2\cos^2\theta_3)\sin\theta_2\sin^2\theta_4 + \cos^2\theta_3] + r^2\}\sin\theta_2}{[(1-r^2)^2\sin^2\theta_4\cos^2\theta_4 + r^2]\cos^2\theta_2} \tag{F.1c}$$

$$C_{1,4} = 0 \tag{F.1d}$$

$$C_{2,2} = -\frac{(1-r^2)\{(1-2\cos^2\theta_3)\sin^2\theta_4 + \cos^2\theta_3\} - 1}{(1-r^2)^2\sin^2\theta_4\cos^2\theta_4 + r^2} \tag{F.1e}$$

$$C_{2,3} = -\frac{(1-r^2)(1-2\sin^2\theta_4)\sin\theta_3\cos\theta_3\sin\theta_2}{[(1-r^2)^2\sin^2\theta_4\cos^2\theta_4 + r^2]\cos\theta_2} \tag{F.1f}$$

$$C_{2,4} = 0 \tag{F.1g}$$

$$\begin{aligned} C_{3,3} &= \frac{(A\cos^2\theta_4 - B)\sin^2\theta_4 - (1-r^4)\sin^2\theta_2\cos^2\theta_3 - r^4\sin^2\theta_2 - r^2}{\{(1-r^2)[-4(1-r^4)\sin^2\theta_4\cos^2\theta_4 + (1-5r^2)]\sin^2\theta_4\cos^2\theta_4 + (r^4+r^2)\}\cos^2\theta_2} \\ A &= -4(1-r^4)\sin^2\theta_2\cos^2\theta_3 - (1-r^2)^2\cos^2\theta_2 + 4(1+r^2)\sin^2\theta_2 \\ B &= (1-r^4)\sin^2\theta_2(2\cos^2\theta_3 - 1)(1 + 4\cos^4\theta_4) \end{aligned} \tag{F.1h}$$

$$C_{3,4} = 0 \tag{F.1i}$$

$$C_{4,4} = \frac{1}{1+r^2}. \tag{F.1j}$$

$$C_W^A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_3 & \sin\theta_3 \\ 0 & -\sin\theta_3 & \cos\theta_3 \end{bmatrix}. \tag{G.2b}$$

Then $C_W^I = C_A^I C_W^A$ is the translation matrix from W to I. Since C_W^I is unitary and its determinant equals to one, the system I can be rotated to W.

Let $\hat{\theta}_i = \theta_i + \Delta\theta_i, i = 1, 2, 3$, thus $\Delta\theta_i$ are the estimation errors, and let $\hat{W}$ be the coordinate frame associated with these estimates.

Using first-order Taylor expansion, we get

$$C_{\hat{W}}^I = C_W^I + \sum_{i=1}^{3} \frac{\partial}{\partial\theta_i}(C_W^I)\Delta\theta_i + O(|\Delta\boldsymbol{\theta}|^2) \tag{G.3}$$

where $\Delta\boldsymbol{\theta} \triangleq [\Delta\theta_1, \Delta\theta_2, \Delta\theta_3]^T$. Thus

$$C_{\hat{W}}^W = C_I^W C_{\hat{W}}^I = I + \sum_{i=1}^{3} C_I^W \frac{\partial}{\partial\theta_i}(C_W^I)\Delta\theta_i + O(|\Delta\boldsymbol{\theta}|^2). \tag{G.4}$$

On the other hand

$$C_{\hat{W}}^W = e^{([\boldsymbol{\phi}_{W\hat{W}}]_W\times)} = I + ([\boldsymbol{\phi}_{W\hat{W}}]_W\times) + O(|\boldsymbol{\phi}_{W\hat{W}}|^2). \tag{G.5}$$

Define the components of the vector angular error in the wave frame by $[\boldsymbol{\phi}_{W\hat{W}}]_W = [\phi_1, \phi_2, \phi_3]^T$, then

$$C_{\hat{W}}^W = I + \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix} + O(|\boldsymbol{\phi}_{W\hat{W}}|^2). \tag{G.6}$$

Using (G.6), (G.4) it then follows that

$$\begin{aligned} [\boldsymbol{\phi}_{W\hat{W}}]_W &= \begin{bmatrix} (C_{\hat{W}}^W)_{3,2} \\ (C_{\hat{W}}^W)_{1,3} \\ (C_{\hat{W}}^W)_{2,1} \end{bmatrix} + O(|\boldsymbol{\phi}_{W\hat{W}}|^2) \\ &= K\Delta\boldsymbol{\theta} + O(|\Delta\boldsymbol{\theta}|^2) + O(|\boldsymbol{\phi}_{W\hat{W}}|^2) \end{aligned} \tag{G.7}$$

where K is given in (G.8) (see bottom of page). That is

$$K_{ij} = \left(C_I^W \frac{\partial}{\partial\theta_j}(C_W^I)\right)_{i-1,i+1}, \qquad 1 \le i, j \le 3 \tag{G.9}$$

where $i+1$ is a positive cyclic shift of the index $i \in \{1, 2, 3\}$ and $i - 1$ is a negative cyclic shift of i (e.g., $1 - 1 = 3$, $1 + 1 = 2$, $3 + 1 = 1$).

Using (G.2a), (G.2b), and (G.9), we find that K is given by (4.34). Also, from Fact 2 and (G.4) we have

$$\operatorname{tr}(C_{\hat{W}}^W) = 1 + 2\cos|\boldsymbol{\phi}_{W\hat{W}}| = 3 + O(|\Delta\boldsymbol{\theta}|^2) \tag{G.10}$$

where we used the fact that

$$\operatorname{tr}[C_I^W(\partial/\partial\theta_i)(C_W^I)] = (1/2)\operatorname{tr}[(\partial/\partial\theta_i)(C_I^W C_W^I)] = 0.$$

Hence, $(2\sin(|\boldsymbol{\phi}_{W\hat{W}}|/2))^2 = 2\cos|\boldsymbol{\phi}_{W\hat{W}}| - 2 = O(|\Delta\boldsymbol{\theta}|^2)$ and the inequality $\sin(|\boldsymbol{\phi}_{W\hat{W}}|/2) \ge (1/\pi)|\boldsymbol{\phi}_{W\hat{W}}|$ implies

$$|\boldsymbol{\phi}_{W\hat{W}}| = O(|\Delta\boldsymbol{\theta}|). \tag{G.11}$$

Thus, using (G.7) and (G.11)

$$E([\boldsymbol{\phi}_{W\hat{W}}]_W[\boldsymbol{\phi}_{W\hat{W}}]_W^T) = KE(\Delta\boldsymbol{\theta}(\Delta\boldsymbol{\theta})^T)K^T + O(E|\Delta\boldsymbol{\theta}|^3). \tag{G.12}$$

Now assume that the model and estimator are regular, see definitions 4.1 and 4.5. It follows that

$$E([\boldsymbol{\phi}_{W\hat{W}}]_W[\boldsymbol{\phi}_{W\hat{W}}]_W^T) \ge K\mathrm{CRB}(\theta_1, \theta_2, \theta_3)K^T + o(1/N). \tag{G.13}$$

Multiply both sides by N and let $N \to \infty$; the result completes the proof of Theorem 4.2. ■

C. Proof of Theorem 4.3

To prove (4.35), recall the definition of δ as the angular error between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}}$. Since $[\delta^2 - 2(1-\cos\delta)]/(1-\cos\delta)^2$ has a continuous extension to $[0, \pi]$, we have for $\delta \in [0, \pi]$

$$\begin{aligned} \delta^2 &= 2(1-\cos\delta) + O((1-\cos\delta)^2) \\ &= 2(1 - \boldsymbol{u}\cdot\hat{\boldsymbol{u}}) + O((1 - \boldsymbol{u}\cdot\hat{\boldsymbol{u}})^2). \end{aligned} \tag{G.14}$$

From (G.5)

$$\begin{aligned} C_{\hat{W}}^W &= I + ([\phi_{W\hat{W}}]_W\times) + \tfrac{1}{2}([\phi_{W\hat{W}}]_W\times)^2 \\ &\quad + \tfrac{1}{6}([\phi_{W\hat{W}}]_W\times)^3 + O(|\phi_{W\hat{W}}|^4). \end{aligned} \tag{G.15}$$

Observe that the entries of $C_{\hat{W}}^W$ are the inner products of the orthonormal bases of W and $\hat{W}$. Also, for any vector angle $\boldsymbol{\phi}$ the matrices $(\boldsymbol{\phi}\times)$ and $(\boldsymbol{\phi}\times)^3 = -|\boldsymbol{\phi}|^2(\boldsymbol{\phi}\times)$ are skew-symmetric (hence with zero diagonal entries). Using these facts we find from (G.15)

$$\begin{aligned} 2(1 - \boldsymbol{u}\cdot\hat{\boldsymbol{u}}) &= 2(1 - [C_{\hat{W}}^W]_{1,1}) \\ &= -[([\boldsymbol{\phi}_{W\hat{W}}]_W\times)^2]_{1,1} + O(|\boldsymbol{\phi}_{W\hat{W}}|^4) \\ &= \phi_2^2 + \phi_3^2 + O(|\boldsymbol{\phi}_{W\hat{W}}|^4). \end{aligned} \tag{G.16}$$

Using (G.14) and (G.16) we now get

$$\delta^2 = \phi_2^2 + \phi_3^2 + O(|\boldsymbol{\phi}_{W\hat{W}}|^4). \tag{G.17}$$

Formula (G.17) gives the representation of δ in terms of the tilt angles ϕ_2, ϕ_3. Note that δ obeys a Pythagorean relationship

$$K = \begin{bmatrix} \left(C_I^W \frac{\partial}{\partial\theta_1}(C_W^I)\right)_{3,2} & \left(C_I^W \frac{\partial}{\partial\theta_2}(C_W^I)\right)_{3,2} & \left(C_I^W \frac{\partial}{\partial\theta_3}(C_W^I)\right)_{3,2} \\ \left(C_I^W \frac{\partial}{\partial\theta_1}(C_W^I)\right)_{1,3} & \left(C_I^W \frac{\partial}{\partial\theta_2}(C_W^I)\right)_{1,3} & \left(C_I^W \frac{\partial}{\partial\theta_3}(C_W^I)\right)_{1,3} \\ \left(C_I^W \frac{\partial}{\partial\theta_1}(C_W^I)\right)_{2,1} & \left(C_I^W \frac{\partial}{\partial\theta_2}(C_W^I)\right)_{2,1} & \left(C_I^W \frac{\partial}{\partial\theta_3}(C_W^I)\right)_{2,1} \end{bmatrix} \tag{G.8}$$

for infinitesimal ϕ_2, ϕ_3. From the estimator's regularity and since $|\boldsymbol{\phi}_{W\hat{W}}| \le \pi$, it follows that $\mathrm{E}[|\boldsymbol{\phi}_{W\hat{W}}|^4] = o(1/N)$. Multiply both sides of (G.17) by N, take their expectation, and let $N \to \infty$. Then, from the definitions (4.2), (4.32) and using (G.17), we find the desired result (4.35).

To prove that for a regular model (4.35) holds also when replacing the MSAE and CVAE by their lower bounds MSAE_{CR} and CVAE_{CR}, it is sufficient to prove this for an equivalent linear model (as in the proof of Lemma C.2) for which these bounds are asymptotically attained by the ML estimate. Hence, applying (4.35) to this estimate, we find that this relationship holds also for the lower bounds.

APPENDIX H
PERFORMANCE ANALYSIS OF THE CROSS-PRODUCT-BASED ESTIMATOR

To prove Theorem 4.4, we need the following two lemmas.

Lemma H.1: If $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ are i.i.d. random vectors in $\mathbb{R}^d$, with $\mathrm{E}\boldsymbol{x}_1 = 0$, $\mathrm{E}||\boldsymbol{x}_1||^4 < \infty$, then $\mathrm{E}||\Sigma_i \boldsymbol{x}_i||^p = O(n^{p/2})$ for any $p \in [0, 4]$.

Proof: Put $\boldsymbol{s} = \Sigma_i \boldsymbol{x}_i$. Clearly $||\boldsymbol{s}||^4 = (\boldsymbol{s}^T\boldsymbol{s})^2 = (\Sigma_{i,j} \boldsymbol{x}_i^T\boldsymbol{x}_j)^2 = \Sigma_{i,j,k,l} \boldsymbol{x}_i^T\boldsymbol{x}_j\boldsymbol{x}_k^T\boldsymbol{x}_l$. From the i.i.d. property we have

$$\begin{aligned}\mathrm{E}||s||^4 = &\mathrm{E}\Bigg[\sum_{i=j=k=l} \boldsymbol{x}_i^T\boldsymbol{x}_j\boldsymbol{x}_k^T\boldsymbol{x}_l + \sum_{i=j\neq k=l} \boldsymbol{x}_i^T\boldsymbol{x}_j\boldsymbol{x}_k^T\boldsymbol{x}_l \\ &+ \sum_{i=k\neq j=l} \boldsymbol{x}_i^T\boldsymbol{x}_j\boldsymbol{x}_k^T\boldsymbol{x}_l + \sum_{i=l\neq j=k} \boldsymbol{x}_i^T\boldsymbol{x}_j\boldsymbol{x}_k^T\boldsymbol{x}_l\Bigg] \\ &\le \mathrm{E}\Bigg[\sum_i ||\boldsymbol{x}_i||^4 + \sum_{i\neq k} ||\boldsymbol{x}_i||^2||\boldsymbol{x}_k||^2 \\ &+ \sum_{i\neq j} ||\boldsymbol{x}_i||^2||\boldsymbol{x}_j||^2 \\ &+ \sum_{i\neq j} ||\boldsymbol{x}_i||^2||\boldsymbol{x}_j||^2\Bigg] \\ &= n\mathrm{E}||\boldsymbol{x}_1||^4 + 3n(n-1)(\mathrm{E}||\boldsymbol{x}_1||^2)^2 = O(n^2).\end{aligned}$$

Since $p/4 \in [0, 1]$, the function $f(x) = x^{p/4}$ is concave for $x \ge 0$. Thus, by Jensen's theorem, $\mathrm{E}||\boldsymbol{s}||^p = \mathrm{E}f(||\boldsymbol{s}||^4) \le f(\mathrm{E}||\boldsymbol{s}||^4) = (O(n^2))^{p/4} = O(n^{p/2})$. ■

Lemma H.2: Assume that $\boldsymbol{x}_1, \cdots \boldsymbol{x}_n$ are i.i.d. random vectors in $\mathbb{R}^d$ with $\mathrm{E}\boldsymbol{x}_1 = 0$, $\mathrm{E}||\boldsymbol{x}_1||^2 < \infty$, and let $\boldsymbol{u}$ be a unit vector in $\mathbb{R}^d$. Let $\boldsymbol{y} \triangleq (1/n)\Sigma_i \boldsymbol{x}_i$, denote by $\hat{\boldsymbol{u}}$ the unit vector in the direction of $\boldsymbol{u} + \boldsymbol{y}$, and by α the angle between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}}$. Then

1) $\hat{\boldsymbol{u}} = \boldsymbol{u} + (I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y} + o_P(n^{-1/2})$ and
2) when also $\mathrm{E}||\boldsymbol{x}_1||^4 < \infty$, then $\alpha^2 = \boldsymbol{y}^T(I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y} + o_P(n^{-1})$ and

$$n\mathrm{E}\alpha^2 = \operatorname{tr}[(I - \boldsymbol{u}\boldsymbol{u}^T)\operatorname{cov}(\boldsymbol{x}_1)] + O(n^{-1/2}).$$

Proof: We have $\hat{\boldsymbol{u}} = (\boldsymbol{u} + \boldsymbol{y})/||\boldsymbol{u} + \boldsymbol{y}||$ and $||\boldsymbol{u} + \boldsymbol{y}||^2 = 1 + 2\boldsymbol{u}^T\boldsymbol{y} + ||\boldsymbol{y}||^2$. If $||\boldsymbol{y}|| \le 1/2$, we have $2\boldsymbol{u}^T\boldsymbol{y} + ||\boldsymbol{y}||^2 \ge ||\boldsymbol{y}||^2 - 2||\boldsymbol{y}|| \ge -3/4$.

From the expansion $(1+x)^{-1/2} = 1 - (1/2)x + O(x^2), (x \ge -3/4)$, we obtain

$$||\boldsymbol{u} + \boldsymbol{y}||^{-1} = 1 - \boldsymbol{u}^T\boldsymbol{y} + O(||\boldsymbol{y}||^2), \qquad (||\boldsymbol{y}|| \le 1/2).$$

Thus

$$\hat{\boldsymbol{u}} = ||\boldsymbol{u} + \boldsymbol{y}||^{-1}(\boldsymbol{u} + \boldsymbol{y}) = \boldsymbol{u} + (I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y} + O(||\boldsymbol{y}||^2). \tag{H.1}$$

Note that if $||\boldsymbol{y}|| > 1/2$, then $||\hat{\boldsymbol{u}} - \boldsymbol{u} - (I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y}|| \le 2 + ||\boldsymbol{y}|| \le 2(2||\boldsymbol{y}||)^2 + 2||\boldsymbol{y}||^2 = O(||\boldsymbol{y}||^2)$.

Since $\alpha = 2\sin^{-1}((1/2)||\hat{\boldsymbol{u}} - \boldsymbol{u}||) = ||\hat{\boldsymbol{u}} - \boldsymbol{u}|| + O(||\hat{\boldsymbol{u}} - \boldsymbol{u}||^3)$, we have

$$\begin{aligned}\alpha^2 &= ||\hat{\boldsymbol{u}} - \boldsymbol{u}||^2 + O(||\hat{\boldsymbol{u}} - \boldsymbol{u}||^4) \\ &= \boldsymbol{y}^T(I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y} + O(||\boldsymbol{y}||^3).\end{aligned} \tag{H.2}$$

Note that if $||\boldsymbol{y}|| > 1/2$, then $|\alpha^2 - \boldsymbol{y}^T(I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{y}| \le \pi^2 + ||\boldsymbol{y}||^2 \le \pi^2(2||\boldsymbol{y}||)^3 + 2||\boldsymbol{y}||^3 = O(||\boldsymbol{y}||)^3$.

1) Since $\mathrm{E}||\boldsymbol{y}||^2 = n^{-1}\mathrm{E}||\boldsymbol{x}_1||^2$, we have from Markov's inequality $||\boldsymbol{y}||^2 = o_P(n^{-1/2})$ and (H.1) gives the result.
2) By Lemma H.1, we have $\mathrm{E}||\boldsymbol{y}||^3 = O(n^{-3/2})$. Thus, by Markov's inequality $||\boldsymbol{y}||^3 = o_P(n^{-1})$ and (H.2) gives the result.

Proof of Theorem 4.4: Using the measurement model (2.6), the estimator computes $\hat{\boldsymbol{s}} = N^{-1}\Sigma_{t=1}^N \operatorname{Re}\{\boldsymbol{y}_E(t) \times \overline{\boldsymbol{y}}_H(t)\}$ and then $\hat{\boldsymbol{u}} = \hat{\boldsymbol{s}}/||\hat{\boldsymbol{s}}||$. To examine the properties of this estimator, we compute from (2.3), (2.4)

$$\begin{aligned}\boldsymbol{y}_E(t) \times \overline{\boldsymbol{y}}_H(t) &= \{\mathcal{E}(t) + \boldsymbol{e}_E(t)\} \times \{\boldsymbol{u} \times \overline{\mathcal{E}}(t) + \overline{\boldsymbol{e}}_H(t)\} \\ &= \mathcal{E}(t) \times (\boldsymbol{u} \times \overline{\mathcal{E}}(t)) + \boldsymbol{e}_E(t) \times (\boldsymbol{u} \times \overline{\mathcal{E}}(t)) \\ &\quad + \mathcal{E}(t) \times \overline{\boldsymbol{e}}_H(t) + \boldsymbol{e}_E(t) \times \overline{\boldsymbol{e}}_H(t).\end{aligned} \tag{H.3}$$

The computation can be done most easily in the A frame (i.e., $(\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2)$, see Appendix G). From (2.7) we have $[\mathcal{E}(t)]_A = [0, \xi_1(t), \xi_2(t)]^T$. Thus

$$\begin{aligned}[\mathcal{E}(t) \times (\boldsymbol{u} \times \overline{\mathcal{E}}(t))]_A &= \begin{bmatrix}0\\ \xi_1\\ \xi_2\end{bmatrix} \times \left(\begin{bmatrix}1\\ 0\\ 0\end{bmatrix} \times \begin{bmatrix}0\\ \overline{\xi}_1\\ \overline{\xi}_2\end{bmatrix}\right) \\ &= \begin{bmatrix}0\\ \xi_1\\ \xi_2\end{bmatrix} \times \begin{bmatrix}0\\ -\overline{\xi}_2\\ \overline{\xi}_1\end{bmatrix} = \begin{bmatrix}||\boldsymbol{\xi}||^2\\ 0\\ 0\end{bmatrix}.\end{aligned} \tag{H.4}$$

That is, $\mathcal{E}(t) \times (\boldsymbol{u} \times \overline{\mathcal{E}}(t)) = ||\mathcal{E}(t)||^2\boldsymbol{u}$.

Next, denote

$$[\boldsymbol{e}_E]_A \triangleq [e_{Eu}, e_{Ev_1}, e_{Ev_2}]^T, \ [\boldsymbol{e}_H]_A \triangleq [e_{Hu}, e_{Hv_1}, e_{Hv_2}]^T$$

where the time dependence is omitted for notational simplicity. A similar computation gives

$$[\boldsymbol{e}_E(t) \times (\boldsymbol{u} \times \overline{\mathcal{E}}(t))]_A = \begin{bmatrix}e_{Ev_1}\overline{\xi}_1 + e_{Ev_2}\overline{\xi}_2\\ -e_{Eu}\overline{\xi}_1\\ -e_{Eu}\overline{\xi}_2\end{bmatrix} \tag{H.5a}$$

$$[\mathcal{E}(t) \times \overline{\boldsymbol{e}}_H(t)]_A = \begin{bmatrix}\xi_1\overline{e}_{Hv_2} - \xi_2\overline{e}_{Hv_1}\\ \xi_2\overline{e}_{Hu}\\ -\xi_1\overline{e}_{Hu}\end{bmatrix} \tag{H.5b}$$

$$[\boldsymbol{e}_E(t) \times \overline{\boldsymbol{e}}_H(t)]_A = \begin{bmatrix} e_{Ev_1}\overline{e}_{Hv_2} - e_{Ev_2}\overline{e}_{Hv_1} \\ e_{Ev_2}\overline{e}_{Hu} - e_{Eu}\overline{e}_{Hv_2} \\ e_{Eu}\overline{e}_{Hv_1} - e_{Ev_1}\overline{e}_{Hu} \end{bmatrix}. \tag{H.5c}$$

Combining these results, we have

$$\begin{aligned}[\boldsymbol{y}_E(t) \times \overline{\boldsymbol{y}}_H(t)]_A = &\begin{bmatrix} \|\boldsymbol{\xi}(t)\|^2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} e_{Ev_1}\overline{\xi}_1 + e_{Ev_2}\overline{\xi}_2 + \xi_1\overline{e}_{Hv_2} - \xi_2\overline{e}_{Hv_1} \\ -e_{Eu}\overline{\xi}_1 + \xi_2\overline{e}_{Hu} \\ -e_{Eu}\overline{\xi}_2 - \xi_1\overline{e}_{Hu} \end{bmatrix} \\ &+ \begin{bmatrix} e_{Ev_1}\overline{e}_{Hv_2} - e_{Ev_2}\overline{e}_{Hv_1} \\ e_{Ev_2}\overline{e}_{Hu} - e_{Eu}\overline{e}_{Hv_2} \\ e_{Eu}\overline{e}_{Hv_1} - e_{Ev_1}\overline{e}_{Hu} \end{bmatrix}.\end{aligned} \tag{H.6}$$

1) Since $\hat{\boldsymbol{s}} = N^{-1}\sum_{t=1}^N \mathrm{Re}\{\boldsymbol{y}_E(t) \times \overline{\boldsymbol{y}}_H(t)\}$ is an average of i.i.d. random vectors that have finite expectation (since $\mathrm{E}[\|\boldsymbol{\xi}(t)\|^2] = \mathrm{tr}(P) < \infty$), we have by Kolmogorov's strong law of large numbers [34, Theorem 5.4.2] $\hat{\boldsymbol{s}} \xrightarrow{a.s.} \mathrm{E}[\|\boldsymbol{\xi}(t)\|^2]\boldsymbol{u} = \mathrm{tr}(P)\cdot\boldsymbol{u}$. Thus
$$\hat{\boldsymbol{u}} = \frac{\hat{\boldsymbol{s}}}{\|\hat{\boldsymbol{s}}\|} \xrightarrow{a.s.} \frac{\mathrm{tr}(P)\cdot\boldsymbol{u}}{\mathrm{tr}(P)} = \boldsymbol{u}, \qquad (N \to \infty).$$

2) Denote $\Delta\boldsymbol{s}(t) \triangleq \mathrm{Re}\{\boldsymbol{y}_E(t) \times \overline{\boldsymbol{y}}_H(t)\} - \mathrm{tr}(P)\cdot\boldsymbol{u}$. We have under the assumption that $\|\boldsymbol{\xi}(t)\|^2, \|\boldsymbol{e}_E(t)\|, \|\boldsymbol{e}_H(t)\|$ are independent and have finite second-order moments, that $\{\Delta\boldsymbol{s}(t), t \geq 1\}$ are i.i.d. with zero mean and covariance given by the equation at the bottom of the page. Denote $\boldsymbol{z}_N \triangleq N^{-1/2}\sum_{t=1}^N \Delta\boldsymbol{s}(t)$. Since $\mathrm{E}\|\Delta\boldsymbol{s}(t)\|^2 < \infty$, we have by the central limit theorem [35, ch. 2, Theorem 9.6] that $\boldsymbol{z}_N$ converges in distribution to $\mathcal{N}(0, \mathrm{cov}(\Delta\boldsymbol{s}(t)))$. Denote $\tilde{\boldsymbol{s}} \triangleq (\mathrm{tr}\,P)^{-1}\hat{\boldsymbol{s}} = \boldsymbol{u} + N^{-1/2}(\mathrm{tr}\,P)^{-1}\boldsymbol{z}_N$. Thus, $\hat{\boldsymbol{u}} = \hat{\boldsymbol{s}}/\|\hat{\boldsymbol{s}}\| = \tilde{\boldsymbol{s}}/\|\tilde{\boldsymbol{s}}\|$. Since $\mathrm{E}\|\Delta\boldsymbol{s}(t)\|^2 < \infty$, the stochastic expansion of Lemma H.2 gives $\hat{\boldsymbol{u}} = \boldsymbol{u} + N^{-1/2}(\mathrm{tr}\,P)^{-1}[I - \boldsymbol{u}\boldsymbol{u}^T]\boldsymbol{z}_N + o_P(N^{-1/2})$. Since $N^{1/2}(\hat{\boldsymbol{u}} - \boldsymbol{u}) - (\mathrm{tr}\,P)^{-1}[I - \boldsymbol{u}\boldsymbol{u}^T]\boldsymbol{z}_N$ converges in probability to zero, we have by Slutsky's theorem, [34, Theorem 4.4.6] that the scaled estimation error of $\boldsymbol{u}$ (i.e., $N^{1/2}(\hat{\boldsymbol{u}} - \boldsymbol{u})$) converges in distribution to the limit distribution of $(\mathrm{tr}\,P)^{-1}[I - \boldsymbol{u}\boldsymbol{u}^T]\boldsymbol{z}_N$, which is $\mathcal{N}(0, (\mathrm{tr}\,P)^{-2}(I - \boldsymbol{u}\boldsymbol{u}^T)\mathrm{cov}(\Delta\boldsymbol{s}(t))(I - \boldsymbol{u}\boldsymbol{u}^T))$.

3) If $\|\boldsymbol{\xi}(t)\|^2, \|\boldsymbol{e}_E(t)\|, \|\boldsymbol{e}_H(t)\|$ have finite fourth-order moments, then $\Delta\boldsymbol{s}(t)$ are i.i.d. random vectors with zero mean and finite fourth-order moment. Let δ be the angle between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}} = (\boldsymbol{u} + N^{-1}\sum_{t=1}^N (\mathrm{tr}\,P)^{-1}\Delta\boldsymbol{s}(t))/\|\boldsymbol{u} + N^{-1}\sum_{t=1}^N (\mathrm{tr}\,P)^{-1}\Delta\boldsymbol{s}(t)\|$. Thus, by Lemma H.2 we have (with $\alpha = \delta$)
$$\begin{aligned}\mathrm{MSAE} &= \lim_{N\to\infty} N\mathrm{E}\delta^2 \\ &= \mathrm{tr}\,[(I - \boldsymbol{u}\boldsymbol{u}^T)\mathrm{cov}((\mathrm{tr}\,P)^{-1}\Delta\boldsymbol{s}(t))].\end{aligned}$$

To get an explicit expression for the MSAE, we exploit its invariance to the reference coordinate frame and work in the A frame. We have
$$\begin{aligned}&\mathrm{tr}\,[(I - \boldsymbol{u}\boldsymbol{u}^T)\mathrm{cov}(\Delta\boldsymbol{s}(t))]_A \\ &= \mathrm{tr}\left\{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathrm{E}[\Delta\boldsymbol{s}(t)\Delta\boldsymbol{s}(t)^T]_A\right\} \\ &= \tfrac{1}{2}(\sigma_E^2 + \sigma_H^2)\,\mathrm{tr}\,(P) + 2\sigma_E^2\sigma_H^2.\end{aligned}$$

Thus $\mathrm{MSAE} = (\mathrm{tr}\,P)^{-2}\mathrm{tr}\,[(I - \boldsymbol{u}\boldsymbol{u}^T)\mathrm{cov}(\Delta\boldsymbol{s}(t))]_A = (1/2)(\sigma_E^2 + \sigma_H^2)(\mathrm{tr}\,P)^{-1} + 2\sigma_E^2\sigma_H^2(\mathrm{tr}\,P)^{-2}$ which gives the result.

4) From Lemma H.2 and Slutsky's theorem, it follows that $N\delta^2$ and $(\mathrm{tr}\,P)^{-2}\boldsymbol{z}_N^T(I - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{z}_N$ converge in law to the same limiting distribution. This gives the result. ■

$$\begin{aligned}E[\Delta\boldsymbol{s}(t)\Delta\boldsymbol{s}(t)^T]_A = &\,\mathrm{cov}\begin{bmatrix} \|\boldsymbol{\xi}(t)\|^2 \\ 0 \\ 0 \end{bmatrix} \\ &+ \mathrm{cov}\left\{\mathrm{Re}\begin{bmatrix} e_{Ev_1}\overline{\xi}_1 + e_{Ev_2}\overline{\xi}_2 + \xi_1\overline{e}_{Hv_2} - \xi_2\overline{e}_{Hv_1} \\ -e_{Eu}\overline{\xi}_1 + \xi_2\overline{e}_{Hu} \\ -e_{Eu}\overline{\xi}_2 - \xi_1\overline{e}_{Hu} \end{bmatrix}\right\} \\ &+ \mathrm{cov}\left\{\mathrm{Re}\begin{bmatrix} e_{Ev_1}\overline{e}_{Hv_2} - e_{Ev_2}\overline{e}_{Hv_1} \\ e_{Ev_2}\overline{e}_{Hu} + e_{Eu}\overline{e}_{Hv_2} \\ e_{Eu}\overline{e}_{Hv_1} - e_{Ev_1}\overline{e}_{Hu} \end{bmatrix}\right\} \\ = &\,\mathrm{var}(\|\boldsymbol{\xi}(t)\|^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \tfrac{1}{2}\begin{bmatrix} (\sigma_E^2 + \sigma_H^2)\,\mathrm{tr}\,(P) & 0 & 0 \\ 0 & \sigma_E^2 P_{11} + \sigma_H^2 P_{22} & (\sigma_E^2 - \sigma_H^2)\,\mathrm{Re}\,(P_{12}) \\ 0 & (\sigma_E^2 - \sigma_H^2)\,\mathrm{Re}\,(P_{12}) & \sigma_E^2 P_{22} + \sigma_H^2 P_{11} \end{bmatrix} \\ &+ \sigma_E^2\sigma_H^2 I_3\end{aligned}$$

APPENDIX I
DEFINITIONS OF SOME BLOCK MATRIX OPERATORS

This appendix defines several block matrix operators found to be useful in this paper. The following notation will be used for a blockwise partitioned matrix A

$$A = \begin{bmatrix} A_{\langle 11\rangle} & \cdots & A_{\langle 1n\rangle} \\ \vdots & & \vdots \\ A_{\langle m1\rangle} & \cdots & A_{\langle mn\rangle} \end{bmatrix} \tag{I.1}$$

with the block entries $A_{\langle ij\rangle}$ of dimensions $\mu_i \times \nu_j$. Define $\overline{\mu} \triangleq \Sigma_{i=1}^m \mu_i, \overline{\nu} \triangleq \Sigma_{j=1}^n \nu_j$, so A is a $\overline{\mu} \times \overline{\nu}$ matrix. Since the block entries may not be of the same size, this is sometimes called an unbalanced partitioning. The following definitions will be considered.

Definition: Block transpose. Let A be an $m\mu \times n\nu$ blockwise partitioned matrix, with blocks $A_{\langle ij\rangle}$ of equal dimensions $\mu \times \nu$. Then the block transpose A^{bT} is an $n\mu \times m\nu$ matrix defined through

$$(A^{bT})_{\langle ij\rangle} = A_{\langle ji\rangle}. \tag{I.2}$$

Definition: Block Kronecker product. Let A be a blockwise partitioned matrix of dimension $\overline{\mu} \times \overline{\nu}$, with block entries $A_{\langle ij\rangle}$ of dimensions $\mu_i \times \nu_j$, and let B be a blockwise partitioned matrix of dimensions $\overline{\eta} \times \overline{\rho}$, with block entries $B_{\langle ij\rangle}$ of dimensions $\eta_i \times \rho_j$. Also, $\overline{\mu} = \Sigma_{i=1}^m \mu_i, \overline{\nu} = \Sigma_{j=1}^n \nu_j, \overline{\eta} = \Sigma_{i=1}^m \eta_i, \overline{\rho} = \Sigma_{j=1}^n \rho_j$. Then the block Kronecker product $A \boxtimes B$ is an $(\Sigma_{i=1}^m \mu_i\eta_i \times \Sigma_{j=1}^n \nu_j\rho_j)$ matrix defined through

$$(A \boxtimes B)_{\langle ij\rangle} = A_{\langle ij\rangle} \otimes B_{\langle ij\rangle} \tag{I.3}$$

i.e., the (i,j) block entry of $A \boxtimes B$ is $A_{\langle ij\rangle} \otimes B_{\langle ij\rangle}$ of dimension $\mu_i\eta_i \times \nu_j\rho_j$. Note that this definition generalizes the so-called unbalanced block Kronecker product in [36], (24).

Definition: Block Schur-Hadamard product. Let A be an $m\mu \times n\nu$ matrix consisting of blocks $A_{\langle ij\rangle}$ of dimensions $\mu \times \nu$, and let B be an $m\nu \times n\eta$ matrix consisting of blocks $B_{\langle ij\rangle}$ of dimensions $\nu \times \eta$. Then, the block Schur-Hadamard product $A \boxdot B$ is an $m\mu \times n\eta$ matrix defined through

$$(A \boxdot B)_{\langle ij\rangle} = A_{\langle ij\rangle} B_{\langle ij\rangle}. \tag{I.4}$$

Thus, each block of the product is a *usual* product of a pair of blocks and is of dimension $\mu \times \eta$.

Definition: Block trace operator: Let A be an $m\mu \times n\mu$ matrix consisting of blocks $A_{\langle ij\rangle}$ of dimensions $\mu \times \mu$. Then the block trace matrix operator $\mathrm{btr}\,[A]$ is an $m \times m$ matrix defined by

$$(\mathrm{btr}\,[A])_{ij} = \mathrm{tr}\, A_{\langle ij\rangle}. \tag{I.5}$$

ACKNOWLEDGMENT

The authors are grateful to Prof. I. Y. Bar-Itzhack from the Department of Aeronautical Engineering, Technion, Israel, for bringing [33] to their attention.

This paper is dedicated to the memory of our physics teacher, Isaac Paldi.

REFERENCES

[1] A. Nehorai and E. Paldi, "Vector sensor processing for electromagnetic source localization," in *Proc. 25th Asilomar Conf. Signals Syst. Comput.* (Pacific Grove, CA), Nov. 1991, pp. 566–572.
[2] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.
[3] B. E. Keiser, *Broadband Coding, Modulation, and Transmission Engineering*. Englewood Cliffs, NJ: Prentice Hall, 1989.
[4] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 1783–1795, Oct. 1990.
[5] B. Ottersten, M. Viberg, and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Trans. Signal Processing*, vol. 40, pp. 590–600, Mar. 1992.
[6] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. Dissertation, Stanford Univ., Stanford, CA, Nov. 1981.
[7] E. R. Ferrara, Jr. and T. M. Parks, "Direction finding with an array of antennas having diverse polarization," *IEEE Trans. Antennas Propagat.*, vol. AP-31, pp. 231–236, Mar. 1983.
[8] I. Ziskind and M. Wax, "Maximum likelihood localization of diversely polarized sources by simulated annealing," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 1111–1114, July 1990.
[9] J. Li and R. T. Compton, Jr., "Angle and polarization estimation using ESPRIT with a polarization sensitive array," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 1376–1383, Sept. 1991.
[10] A. J. Weiss and B. Friedlander, "Performance analysis of diversely polarized antenna arrays," *IEEE Trans. Signal Processing*, vol. 39, pp. 1589–1603, July 1991.
[11] J. D. Means, "Use of three-dimensional covariance matrix in analyzing the polarization properties of plane waves," *J. Geophys. Res.*, vol. 77, pp. 5551–5559, Oct. 1972.
[12] J. Dugundji, "Envelopes and pre-envelopes of real waveforms," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 53–57, Mar. 1958.
[13] S. O. Rice, "Envelopes of narrow-band signals," *Proc. IEEE*, vol. 70, pp. 692–699, July 1982.
[14] D. Giuli, "Polarization diversity in radars," *Proc. IEEE*, vol. 74, pp. 245–269, Feb. 1986.
[15] M. Born and E. Wolf, Eds., *Principles of Optics*. Oxford: Pergamon, 1980, 6th ed.
[16] C. G. Khatri and C. R. Rao, "Solution to some functional equations and their applications to characterization of probability distribution," *Sankhyā Ser. A*, vol. 30, pp. 167–180, 1968.
[17] C. R. Rao and S. K. Mitra, *Generalized Inverse of Matrices and its Applications*. New York: Wiley, 1971.
[18] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 37, pp. 720–741, May 1989.
[19] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation: Asymptotic Theory*. New York: Springer-Verlag, 1981.
[20] E. Paldi and A. Nehorai, "A generalized Cramér-Rao bound," in preparation.
[21] P. E. Caines, *Linear Stochastic Systems*. New York: Wiley, 1988.
[22] J. D. Jackson, *Classical Electrodynamics*. New York: Wiley, 1975, 2nd ed.
[23] A. Nehorai, D. Starer, and P. Stoica, "Direction-of-arrival estimation in applications with multipath and few snapshots," *Circ. Syst. Signal Processing*, vol. 10, no. 3, pp. 327–342, 1991.
[24] M. Kanda and D. A. Hill, "A three-loop method for determining the radiation characteristics of an electrically small source," *IEEE Trans. Electromag. Compat.*, vol. EMC-34, pp. 1–3, Feb. 1992.
[25] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," to be published in *IEEE Trans. Signal Processing*. (A short version appeared in *Proc. 26th Asilomar Conf. Signals, Syst. Comput.* (Pacific Grove, CA), Oct. 1992, pp. 192–198.)
[26] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol. 11, pp. 329–337, 1986.
[27] A. G. Jaffer, "Maximum likelihood direction finding of stochastic sources: A separable solution," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Processing* (New York, NY), Apr. 1988, pp. 2893–2896.
[28] Y. Bresler, "Maximum likelihood estimation of a linearly structured covariance with application to antenna array processing," in *Proc. 4th ASSP Workshop Spectrum Estimation Modeling* (Minneapolis, MN), Aug. 1988, pp. 172–175.
[29] P. Stoica and A. Nehorai, "On the concentrated stochastic likelihood function in array signal processing," Rep. 9214, Cent. for Syst. Sci., Yale Univ., New Haven, CT, Nov. 1992.

[30] B. Hochwald and A. Nehorai, "Concentrated Cramér–Rao bound expressions," to be published in *IEEE Trans. Inform. Theory*, Mar. 1994.

[31] W. J. Bangs and P. M. Schultheiss, "Space-time processing for optimal parameter estimation," in *Proc. Signal Processing NATO Adv. Study Inst.* (New York: Academic), 1973, pp. 577–590.

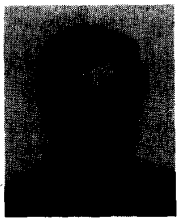
[32] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Welsey, 1980, 2nd ed.

[33] M. D. Shuster, "A survey of attitude representations," to be published in the *J. Astronaut. Sci.*, vol. 41, no. 4, pp. 439–517, Oct.-Dec. 1993.

[34] K. L. Chung, *A Course in Probability Theory*, 2nd ed. London: Academic, 1974.

[35] R. Durrett, *Probability: Theory and Examples*. Pacific Grove, CA: Wadsworth Brooks/Cole, 1991.

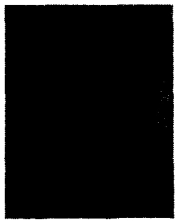
[36] R. H. Koning, "Block Kronecker products and the vecb operator," *Linear Algebra Applicat.*, vol. 149, pp. 165–184, 1991.



Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion–Israel Institute of Technology, in 1976 and 1979, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1983.

From 1983 to 1984, he was a Research Associate at Stanford University. From 1984 to 1985, he was a Research Engineer at Systems Control Technology, Inc., Palo Alto, CA. Since 1985, he has been with the Department of Electrical Engineering and the Applied Mathematics Program at Yale University, New Haven, CT, where he is an Associate Professor. He held Visiting Professorships at Uppsala University, Sweden, and the Technion, Israel, and a Visiting Consulting position at the Defense Science Organization of Singapore. His areas of interest are sensor array processing, physical modeling, adaptive filtering, system identification, and biomedical engineering.

Dr. Nehorai is an Associate Editor of the journal *Circuits, Systems and Signal Processing* and was an Associate Editor of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. He served on the Statistical Signal and Array Processing Technical Committee and the Education Committee in the IEEE Signal Processing Society. He is the Chairman of the Connecticut IEEE Signal Processing chapter. He was corecipient, with P. Stoica, of the 1989 IEEE Signal Processing Society's Senior Paper Award for the paper "MUSIC, Maximum Likelihood, and Cramér–Rao Bound." He is a member of Sigma Xi.



Eytan Paldi (S'87–M'94) received the B.Sc. and M.Sc. degrees in mathematics and the B.Sc. and D.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1977, 1984, 1978, and 1990, respectively.

From 1979 to 1986, he was a Mathematician and Software Engineer with Elbit Computers Inc., working on research and development in avionics. From 1986 to 1990, he was a teaching assistant at the Technion. During 1990, he was a consultant for Elbit Computers, Inc. During 1991, he was an HTI Post-Doctoral Fellow with the Department of Electrical Engineering at Yale University, where he is currently a Post-Doctoral Associate. His current interests are in stochastic modeling, identification and adaptive filtering, optimization, polynomial root clustering, and analytic number theory.

Dr. Paldi was awarded several Gutwirth Fellowships, the Jury Prize, and the Kennedy–Leigh Prize.