Massachusetts Institute of Technology

A Semi-Blind Method for Localization of Underwater Acoustic Sources

Amir Weiss, *Member, IEEE*, Toros Arikan, *Student Member, IEEE*, Hari Vishnu, *Senior Member, IEEE*,
Grant B. Deane, Andrew C. Singer, *Fellow, IEEE*, and Gregory W. Wornell, *Fellow, IEEE*

Abstract—Underwater acoustic localization has traditionally been challenging due to the presence of unknown environmental structure and dynamic conditions. The problem is richer still when such structure includes occlusion, which causes the loss of line-of-sight (LOS) between the acoustic source and the receivers, on which many of the existing localization algorithms rely. We develop a semi-blind passive localization method capable of accurately estimating the source's position even in the possible absence of LOS between the source and all receivers. Based on typically-available prior knowledge of the water surface and bottom, we derive a closed-form expression for the optimal estimator under a multi-ray propagation model, which is suitable for shallow-water environments and high-frequency signals. By exploiting a computationally efficient form of this estimator, our methodology makes comparatively high-resolution localization feasible. We also derive the Cramér-Rao bound for this model, which can be used to guide the placement of collections of receivers so as to optimize localization accuracy. The method improves a balance of accuracy and robustness to environmental model mismatch, relative to existing localization methods that are useful in similar settings. The method is validated with simulations and water tank experiments.

Index Terms—Localization, non-line-of-sight, underwater acoustics, matched field processing, maximum likelihood, Cramér-Rao bound, Cholesky decomposition.

I. INTRODUCTION

Underwater localization of acoustic sources is an important and challenging problem, and arises in a wide range of applications [1]–[3]. As such, it has been extensively addressed in the literature, where early work dates back to at least the mid-1970s [4]. Fruitful combinations of advanced signal processing methods and detailed underwater acoustic propagation models have led to a variety of methods for different regimes (shallow/deep water, short/long distances, etc.) [5], [6].

While an abundance of methods have been developed and proposed over the years, only a portion of these survive the ruthless test of practicality. Indeed, from a practical point of view, a good *applicable* method is one that, on the one hand exploits as much prior knowledge as possible, but on the other hand does not go too far by assuming access to unavailable information/resources. In the context of passive underwater acoustic localization, our goal in this work is to provide a robust algorithm, while judiciously balancing this trade-off.

In particular, we consider scenarios where the area of interest is characterized by shallow waters (say, up to ~ 100 m depth [7]) and relatively short distances (say, up to ~ 1 km). In this regime, under a few additional realistic assumptions (stated explicitly in the sequel), the acoustic signal propagation can be approximated by ray trace modeling [7], [8]. This allows us to *exploit* the multipath channel effect, rather than *mitigate* it. In other words, we explicitly incorporate prior knowledge on the structure of the environment, which either allows us to successfully localize using fewer resources (e.g., sensors or measurements), or to improve performance while using the same resources. Moreover, we are capable of localizing a source in the complete absence of line-of-sight (LOS) signal components, based on non-LOS (NLOS) signal reflections. Naturally, these notions have already been considered in some settings, as reviewed in what follows.

A. Related Work: Underwater Acoustic Localization

For short-range localization in shallow-water environments, straight-ray tracing is a widely-accepted approximation for acoustic signal propagation [7], [8]. In such environments, the speed of sound is (at least approximately) constant[1] and known in the relevant volume of interest. Therefore, propagation delay, namely time of arrival (TOA) or time-difference of arrival (TDOA) (e.g., [10], [11]), is usually employed as a basis for different localization methods [12]. However, since the underwater acoustic environment typically induces a rich multipath channel [13], the measured signals contain both LOS and NLOS components. Under such circumstances, the performance of TOA/TDOA-based methods usually deteriorates, possibly up to unacceptable error levels.

Since the complete multipath channel is (generally) unknown, a possible remedy is to first identify and separate the LOS components. Diamant, *et al.* propose in [14] a method for classifying the signal components as LOS and NLOS, and for subsequent range estimation based on the classified LOS components. While this approach can certainly work, it does not attempt to exploit the NLOS reflections, which contain valuable information on the unknown source location. Emokpae and Younis propose in [15] a surface-reflection-based method in an active setting, where only the surface reflections are exploited. In [16], Emokpae, *et al.* present an extended, enhanced version of this notion, where a scheme that employs both the LOS and surface-reflected NLOS components is developed to locate a lost (drifted away) node of an underwater sensor network. To use this method, all nodes in the network

---

[1]Nearly constant sound speed may be found, e.g., in very shallow waters, or shallow waters that are well-mixed [9].

are required to have a sensor array, with more than one sensor, and the waveform emitted from the lost node (i.e., source) is assumed to be known, which is not always possible and less common in passive settings. Assuming that perfect knowledge of the physical model is available, which translates into an equivalent impulse response, matched field processing (MFP) [17] is a well-known technique that makes full use of the environmental structure for enhanced localization. However, as mentioned in [18], in realistic applications model mismatch is a serious problem for MFP, on top of its heavy computational workload. Recent increasing efforts towards reducing system cost [19], [20] and computational complexity [21], [22], while exploiting environmental structure [23], motivate our current work.

B. Semi-Blind Localization: Motivation and Contributions

We propose a semi-blind localization (SBL) method that uses a spatially diverse network of receivers. Each receiver is required to have a single sensor (rather than a sensor array, as in [16]), which reduces hardware requirements, and hence the overall cost of the system.[2] The information lost by restricting the number of sensors is mitigated by leveraging available partial prior knowledge on the structure of the environment, namely the depths of the sensors and the ocean bottom. Our SBL method, developed in a nonBayesian framework, jointly estimates the associated parameters of the implied impulse response with the unknown source position, and thus can be viewed as a form of focalization [18]. However, it is more naturally related to the direct position determination approach [25], originally proposed for narrowband radio frequency signals. We show that with some carefully chosen adaptations, a similar, though generalized approach leads to our SBL method,[3] which provides a good balance between accuracy and robustness to some physical model mismatch. We demonstrate this via simulation experiments by comparing to MFP and to the TDOA method referred to as "generalized cross-correlation with phase transform" (GCC-PHAT) [26]–[28], which is well-known due to its resilience to multipath.

We note in passing that if additional knowledge of the environment is available, one may consider taking a Bayesian approach, and incorporate the available knowledge by introducing an appropriate prior distribution on (all or some of) the unknowns. In this work, we take a nonBayesian approach.

Our main contributions are the following:

- *A novel SBL method for underwater acoustic sources*: We adopt the widely-accepted straight-ray tracing approach for shallow-water to define a three-ray model, which explicitly takes into account the NLOS surface and bottom signal reflections. Consequently, on top of enhanced accuracy due to this multipath model, our method is capable of localization in the absence of LOS, due to a potential occluder, such as a vessel or pier pilings.
- *Computationally efficient direct localization*: Contrary to indirect (e.g., TDOA-based), standard localization methods (e.g., [29]–[31]), we take a different approach, in

---

[2] In an application such as the ocean-of-things [3], an optimal subset of sensors could be chosen from a larger set of sensors [24].

[3] In contrast to the previous claim in [25], that this approach "is suitable only for RF signals and not for underwater emitter location".
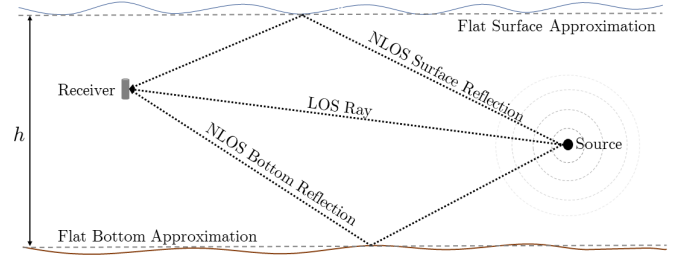

Fig. 1: A 2-dimensional illustration of the three-ray model. When the surface and bottom are approximately flat in the operational area [33], [34], this model enables NLOS-based localization in the potential absence of the LOS signal component, e.g., due to an occluder.

which our algorithm is applied directly to the observed signals. Consequently, the notion of TDOA is redundant in our framework. Specifically, we assume that the source's waveform is unknown, and in particular, we do not assume it is a pulse-type signal. We provide a computationally efficient algorithm to the resulting nonlinear optimization problem (see Section IV, Proposition 3), and demonstrate that the algorithm works well for pulse- or non-pulse-type signals in Section VI.

- *Lower bound on asymptotic performance*: We develop the Cramér-Rao lower bound (CRLB) on the mean-squared error (MSE) of any unbiased localization method for a special case of our signal model, in which our proposed solution coincides with the maximum likelihood estimate (MLE) of the source position. We demonstrate the validity of this bound with respect to ocean ambient noise, using previously collected ocean acoustic recordings [32].
- *Applicability proof of concept*: We provide a proof of concept, demonstrated on acoustic measurements collected in a well-controlled, small-scale water tank, which provides an acoustically frequency-scaled model for the shallow-water environment.

The rest of the paper is organized as follows. The remainder of this section is devoted to an outline of our notation. In Section II we formulate the problem for the three-ray signal model depicted in Fig. 1, and the MFP solution of this model is presented in Section III. The main results, including our proposed SBL method, are presented in Section IV. In Section V we derive the respective CRLB, and present empirical simulation results that corroborate our analytical derivation in Section VI. Concluding remarks are provided in Section VIII.

C. Notation

We use x, $\boldsymbol{x}$, and $\boldsymbol{X}$ for a scalar, column vector and matrix, respectively. The superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^{-1}$ denote the transposition, complex conjugation, conjugate transposition, and inverse operators, respectively. We use $\boldsymbol{I}_K$ to denote the $K \times K$ identity matrix, and $\mathbf{O}$ for the all-zeros matrix. The pinning vector $\boldsymbol{e}_k$ denotes the k-th column of the identity matrix, with context-dependent dimension. Further, $\delta_{k\ell} \triangleq \boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{e}_\ell$ denotes the Kronecker delta of k and ℓ. $\mathbb{E}[\cdot]$ denotes expectation, $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of its (square) matrix argument, and the $\mathrm{Diag}(\cdot)$ operator forms an $M \times M$ diagonal matrix from its M-dimensional vector argument. The Kronecker product is denoted by $\otimes$. We use $\jmath$ (a dotless j) to denote $\sqrt{-1}$; $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote

the real and imaginary parts (respectively) of their complex-valued arguments. The ℓ^2 norm is denoted by $\|\cdot\|_2$, and $\text{rank}(\cdot)$ denotes the rank of its matrix argument. The symbols $\mathbb{R}$ and $\mathbb{C}$ denote the real line and complex plane, respectively. We use $\bar{\boldsymbol{x}}$ to denote the (normalized) discrete Fourier transform (DFT) of $\boldsymbol{x}$, and $\widehat{\boldsymbol{x}}$ to denote an estimate thereof. We use $\mathcal{O}(\cdot)$ to denote the standard big O notation [35].

II. PROBLEM FORMULATION

Consider L spatially-diverse, time-synchronized receivers at known locations, each consisting of a *single* omni-directional hydrophone.[4] Furthermore, consider the presence of an unknown signal in an isotropic homogeneous medium, emitted from a source whose deterministic, unknown position is denoted by the vector of coordinates $\boldsymbol{p} \in \mathbb{R}^{3\times1}$. We assume that the source is static, and is located sufficiently far from all L receivers to permit a planar wavefront (far-field) approximation in the shallow-water waveguide. Each receiver records the measured acoustic signal on a fixed observation time interval, which after sampling and baseband conversion amounts to N samples. We further assume that the area of operation can be considered as a shallow-water environment, and that the ocean floor depth in the relevant area of operation[5] is approximately constant [36]. We restrict our scope to (approximately) isovelocity environments and high frequency signals, in which the straight-ray model approximately holds. Since we focus on short ranges in shallow-water environments, we neglect nonlinear propagation effects in the waveguide.

Although the underwater acoustic channel generally gives rise to an equivalent rich multipath channel, a relatively simple, yet useful, approach is the *three-ray model*, illustrated in Fig. 1. In this approach, the modeled signal components are:

1) The direct-path LOS component;
2) The surface reflection NLOS component; and
3) The bottom reflection NLOS component.

Accordingly, the associated distances traveled by these components from the source to the ℓ-th receiver are given by [37]

$$R_{1\ell} \triangleq \|\boldsymbol{p}_\ell - \boldsymbol{p}\|_2, \qquad \text{(LOS)} \qquad (1)$$

$$R_{2\ell} \triangleq \sqrt{\rho_\ell^2 + (z_p + z_\ell)^2}, \qquad \text{(NLOS surface)} \qquad (2)$$

$$R_{3\ell} \triangleq \sqrt{\rho_\ell^2 + (2h - z_p - z_\ell)^2}, \qquad \text{(NLOS bottom)} \qquad (3)$$

where $\boldsymbol{p} \triangleq [x_p \ y_p \ z_p]^{\mathrm{T}}$; $\boldsymbol{p}_\ell \triangleq [x_\ell \ y_\ell \ z_\ell]^{\mathrm{T}}$ is the position of the ℓ-th receiver; $\rho_\ell \triangleq \sqrt{(x_p - x_\ell)^2 + (y_p - y_\ell)^2}$ is the horizontal distance between the source and the ℓ-th receiver; and h is the bottom depth in the area of interest. An illustration of the geometry associated with (1)–(3) in our coordinate system is given in Fig. 2. Therefore, assuming isovelocity, the associated time-delays of these components are

$$\tau_{r\ell}(\boldsymbol{p}) \triangleq \frac{R_{r\ell}}{c}, \ r \in \{1,2,3\}, \ \forall \ell \in \{1,\ldots,L\}, \qquad (4)$$

where c denotes the speed of sound, assumed to be known.

This model can be viewed as a third-order approximation (with respect to the delayed signal components) of the

[4] We focus on the single sensor case for convenience. However, our methodology can in principle be used when the receivers have sensor arrays.

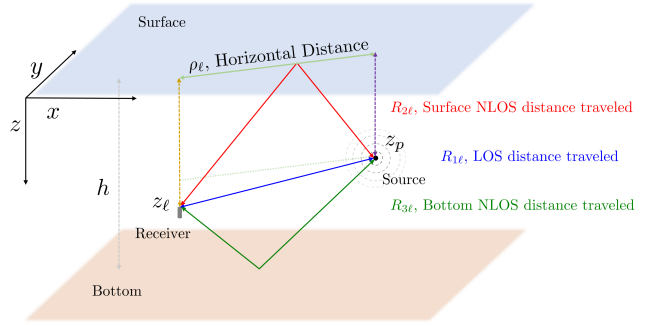[5] The smallest rectangular area encompassing the source and receivers.

Fig. 2: A 3-dimensional illustration of the geometry leading to (1)–(3).

equivalent impulse response of an acoustic channel, whose energy is concentrated in the three arrivals corresponding to the LOS component, and the surface and bottom reflections. While some unpredictable factors can give rise to additional components in the induced impulse response, the surface and bottom of the ocean are always present. Therefore, it is reasonable to incorporate these additional signal components into the model. Moreover, this simplified model allows for successful localization in the absence of (even all) LOS components in the received signals, a situation that may occur, e.g., due to the presence of potential occluders. This will be demonstrated via simulations and experiments with real data in Section VI.

A. Baseband Signal Model

Formally, and assuming the source has been detected in a given frequency band, the sampled, baseband-converted signal from the ℓ-th receiver is given by

$$x_\ell[n] = \sum_{r=1}^{3} b_{r\ell} s_{r\ell}[n] + v_\ell[n] \triangleq \boldsymbol{s}_\ell^{\mathrm{T}}[n]\boldsymbol{b}_\ell + v_\ell[n] \in \mathbb{C}, \\ \forall n \in \{1,\ldots,N\}, \ \forall \ell \in \{1,\ldots,L\}, \qquad (5)$$

where we have defined $\boldsymbol{s}_\ell[n] = [s_{1\ell}[n] \ s_{2\ell}[n] \ s_{3\ell}[n]]^{\mathrm{T}} \in \mathbb{C}^{3\times1}$, $\boldsymbol{b}_\ell = [b_{1\ell} \ b_{2\ell} \ b_{3\ell}]^{\mathrm{T}} \in \mathbb{C}^{3\times1}$, and where

1) $b_{r\ell} \in \mathbb{C}$ denotes the unknown attenuation coefficient from the source to the ℓ-th sensor associated with the r-th component (LOS or surface/bottom NLOS reflection);
2) $s_{r\ell}[n] \triangleq s(t - \tau_{r\ell}(\boldsymbol{p}))|_{t=nT_s} \in \mathbb{C}$ denotes the sampled r-th component of the unknown signal waveform at the ℓ-th sensor, where $s(t - \tau_{r\ell}(\boldsymbol{p}))$ is the analog, continuous-time waveform delayed by $\tau_{r\ell}(\boldsymbol{p})$, and T_s is the (known) sampling period; and
3) $v_\ell[n] \in \mathbb{C}$ denotes the additive noise at the ℓ-th receiver, representing the overall contributions of internal receiver noise and ambient noise, modeled as a zero-mean random process with an unknown variance $\sigma_{v_\ell}^2$.

B. Equivalent Formulation in the Frequency Domain

Applying the normalized DFT to (5) yields the equivalent frequency-domain representation for all $\ell \in \{1,\ldots,L\}$,

$$\begin{aligned} \bar{x}_\ell[k] &= \sum_{r=1}^{3} b_{r\ell} \bar{s}[k] e^{-\jmath\omega_k \tau_{r\ell}(\boldsymbol{p})} + \bar{v}_\ell[k] \\ &\triangleq \bar{s}[k] \cdot \underbrace{\boldsymbol{d}_\ell^{\mathrm{H}}[k]\boldsymbol{b}_\ell}_{\triangleq \bar{h}_\ell[k]} + \bar{v}_\ell[k] = \bar{s}[k] \cdot \bar{h}_\ell[k] + \bar{v}_\ell[k] \in \mathbb{C}, \end{aligned} \qquad (6)$$

where we have defined

$$\boldsymbol{d}_\ell[k] \triangleq [e^{-\jmath\omega_k\tau_{1\ell}(\boldsymbol{p})} \ e^{-\jmath\omega_k\tau_{2\ell}(\boldsymbol{p})} \ e^{-\jmath\omega_k\tau_{3\ell}(\boldsymbol{p})}]^{\mathrm{H}} \in \mathbb{C}^{3\times 1},$$

$$\omega_k \triangleq \frac{2\pi(k-1)}{NT_s} \in \mathbb{R}_+, \ \forall k \in \{1,\dots,N\}.$$

For shorthand, we further define

$$\bar{\boldsymbol{x}}_\ell \triangleq [\bar{x}_\ell[1] \cdots \bar{x}_\ell[N]]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}, \overline{\boldsymbol{X}}_\ell \triangleq \mathrm{Diag}(\bar{\boldsymbol{x}}_\ell),$$
$$\bar{\boldsymbol{s}} \triangleq [\bar{s}[1] \cdots \bar{s}[N]]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}, \quad \overline{\boldsymbol{S}} \triangleq \mathrm{Diag}(\bar{\boldsymbol{s}}),$$
$$\bar{\boldsymbol{v}}_\ell \triangleq [\bar{v}_\ell[1] \cdots \bar{v}_\ell[N]]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}, \tag{7}$$
$$\boldsymbol{D}_\ell \triangleq [\boldsymbol{d}_\ell[1] \cdots \boldsymbol{d}_\ell[N]]^{\mathrm{T}} \in \mathbb{C}^{N\times 3}, \boldsymbol{H}_\ell \triangleq \mathrm{Diag}(\boldsymbol{D}_\ell \boldsymbol{b}_\ell).$$

Note that $\boldsymbol{D}_\ell$ and $\boldsymbol{H}_\ell$ are nonlinear functions of the unknown emitter position $\boldsymbol{p}$, as suggested by the definition of $\boldsymbol{d}_\ell[k]$ above and (1)–(4), though we omit this for brevity. With this notation, we may now write (6) compactly as

$$\bar{\boldsymbol{x}}_\ell = \boldsymbol{H}_\ell \bar{\boldsymbol{s}} + \bar{\boldsymbol{v}}_\ell \in \mathbb{C}^{N\times 1}, \ \forall \ell \in \{1,\dots,L\}. \tag{8}$$

Thus, the localization problem can be formulated as follows:

Problem: *Given the measurements* $\{\bar{\boldsymbol{x}}_\ell \in \mathbb{C}^{N\times 1}\}_{\ell=1}^L$ *of the signal model* (8), *localize the source, namely estimate* $\boldsymbol{p}$.

We emphasize that although we are interested solely in $\boldsymbol{p}$, the channel parameters $\{\boldsymbol{b}_\ell\}$ and the DFT coefficients $\bar{\boldsymbol{s}}$ of the emitted waveform are unknown as well.

III. THE MATCHED FIELD PROCESSING SOLUTION

The key assumption of MFP approaches is that, for a given hypothesized emitter location $\boldsymbol{p}$, the channel response $\boldsymbol{H}_\ell$ is *fully* predictable.[6] For the model (5), the attenuation coefficients $\{\boldsymbol{b}_\ell\}$ can be assumed to be given by[7]

$$b_{1\ell} = \frac{1}{R_{1\ell}}, \quad \text{(LOS attenuation)} \tag{9}$$

$$b_{2\ell} = \frac{-1}{R_{2\ell}}, \quad \text{(NLOS surface reflection attenuation)} \tag{10}$$

$$b_{3\ell} = \frac{\kappa_b}{R_{3\ell}}, \quad \text{(NLOS bottom reflection attenuation)} \tag{11}$$

for all $\ell \in \{1,\dots,L\}$, where κ_b is the bottom reflection coefficient, which (presumably) can be determined based on prior physical knowledge (e.g., assuming the bottom is sand, silt, clay, rock, etc.) and the angle of incidence, and is assumed to be known within the MFP framework for a given hypothesized emitter location $\boldsymbol{p}$. For (10), we assumed a perfectly reflecting ocean surface [8], [38], which approximately holds for calm shallow waters. Based on this knowledge, the channel responses $\{\boldsymbol{H}_\ell\}$ can be readily computed.

The MFP solution for the three-ray model, denoted for convenience as MFP3, is then given by

$$\widehat{\boldsymbol{p}}_{\mathrm{MFP3}} \triangleq \underset{\boldsymbol{p}\in\mathbb{R}^{3\times 1}}{\mathrm{argmin}} \ \underset{\bar{\boldsymbol{s}}\in\mathbb{C}^{N\times 1}}{\min} \ \widetilde{C}_{\mathrm{MFP3}}(\boldsymbol{p},\bar{\boldsymbol{s}}), \tag{12}$$

where

$$\widetilde{C}_{\mathrm{MFP3}}(\boldsymbol{p},\bar{\boldsymbol{s}}) \triangleq \sum_{\ell=1}^L \|\bar{\boldsymbol{x}}_\ell - \boldsymbol{H}_\ell(\boldsymbol{p})\bar{\boldsymbol{s}}\|_2^2, \tag{13}$$

[6]Otherwise, infeasible high-dimensional optimization is required.
[7]Ignoring the effects of volume absorption in water, which are minimal.

and here we write $\boldsymbol{H}_\ell(\boldsymbol{p})$ (rather than $\boldsymbol{H}_\ell$) to emphasize the dependence on $\boldsymbol{p}$. The simplified MFP3 solution is given by

$$\widehat{\boldsymbol{p}}_{\mathrm{MFP3}} = \underset{\boldsymbol{p}\in\mathbb{R}^{3\times 1}}{\mathrm{argmax}} \sum_{k=1}^N \frac{\left|\bar{\boldsymbol{x}}[k]^{\mathrm{H}}\bar{\boldsymbol{h}}_k(\boldsymbol{p})\right|^2}{\|\bar{\boldsymbol{h}}_k(\boldsymbol{p})\|^2}, \tag{14}$$

where we have defined, for every k-th DFT component,

$$\bar{\boldsymbol{x}}[k] \triangleq [\bar{x}_1[k] \ \cdots \ \bar{x}_L[k]]^{\mathrm{T}} \in \mathbb{C}^{L\times 1},$$
$$\bar{\boldsymbol{h}}_k(\boldsymbol{p}) \triangleq [\bar{h}_1[k] \ \cdots \ \bar{h}_L[k]]^{\mathrm{T}} \in \mathbb{C}^{L\times 1},$$

using $\bar{\boldsymbol{h}}_k(\boldsymbol{p})$ (rather than $\bar{\boldsymbol{h}}_k$) to emphasize the dependence on $\boldsymbol{p}$. For completeness of the exposition, the derivation of the simplified form (14) is given in the supplementary materials.

In (14), the channel impulse response $\bar{\boldsymbol{h}}_k(\boldsymbol{p})$ is considered to be fully known for any given hypothesized position $\boldsymbol{p}$ (via (1)–(4) and (9)–(11)). In other words, assuming perfect knowledge of $\boldsymbol{b}_1,\dots,\boldsymbol{b}_L$ means that any relevant physical parameter, such as the ocean bottom sediment coefficient κ_b in (11), is assumed to be perfectly known as well. We relax this (somewhat unrealistic) assumption in our semi-blind localization approach described next.

IV. THE PROPOSED SEMI-BLIND LOCALIZATION METHOD

As our semi-blind framework, we only assume that the bottom depth h is known, but we do not assume that we have any prior knowledge of the channel attenuation coefficients. Thus, since the waveform emitted from the source is also unknown, we may assume without loss of generality (w.l.o.g.) that $\|\bar{\boldsymbol{s}}\|_2 = 1$, viz., $\bar{\boldsymbol{s}} \in \mathcal{S}_N \triangleq \{\boldsymbol{z} \in \mathbb{C}^{N\times 1}: \|\boldsymbol{z}\|_2 = 1\}$, where $\mathcal{S}_N$ is the N-dimensional unit sphere. This assumption, which is common in similar (semi-)blind formulations (e.g., [39]), is justified due to the inherent *scaling* ambiguity in (8),

$$\alpha \in \mathbb{C}: \boldsymbol{H}_\ell\bar{\boldsymbol{s}} = \mathrm{Diag}(\boldsymbol{D}_\ell \boldsymbol{b}_\ell)\bar{\boldsymbol{s}} = \mathrm{Diag}\left(\boldsymbol{D}_\ell\underbrace{\left(\tfrac{1}{\alpha}\boldsymbol{b}_\ell\right)}_{\triangleq\widetilde{\boldsymbol{b}}_\ell}\right)\underbrace{(\alpha\bar{\boldsymbol{s}})}_{\triangleq\widetilde{\bar{\boldsymbol{s}}}}$$
$$= \underbrace{\mathrm{Diag}\left(\boldsymbol{D}_\ell\widetilde{\boldsymbol{b}}_\ell\right)}_{\triangleq\widetilde{\boldsymbol{H}}_\ell}\widetilde{\bar{\boldsymbol{s}}} = \widetilde{\boldsymbol{H}}_\ell\widetilde{\bar{\boldsymbol{s}}}, \quad \forall \ell \in \{1,\dots,L\},$$

which, granted, is immaterial to our localization problem.

Our proposed SBL solution can be viewed as the MLE of $\boldsymbol{p}$, obtained by joint estimation of *all* the unknown deterministic model parameters, under the assumption that the noise processes $\{\bar{\boldsymbol{v}}_\ell\}_{\ell=1}^L$ from all different sensors are temporally white complex normal (CN) processes, mutually statistically independent, with equal variances. In this case, the MLE of $\boldsymbol{p}$ is the solution to the nonlinear least squares problem

$$\widehat{\boldsymbol{p}}_{\mathrm{SBL}} \triangleq \underset{\boldsymbol{p}\in\mathbb{R}^{3\times 1}}{\mathrm{argmin}} \underbrace{\underset{\substack{\bar{\boldsymbol{s}}\in\mathcal{S}_N \\ \boldsymbol{B}\in\mathbb{C}^{3\times L}}}{\min} \widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p},\bar{\boldsymbol{s}},\boldsymbol{B})}_{C_{\mathrm{SBL}}(\boldsymbol{p})} \triangleq \underset{\boldsymbol{p}\in\mathbb{R}^{3\times 1}}{\mathrm{argmin}} \ C_{\mathrm{SBL}}(\boldsymbol{p}),$$

$$\tag{15}$$

where the objective function $\widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p},\bar{\boldsymbol{s}},\boldsymbol{B})$ is defined as

$$\widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p},\bar{\boldsymbol{s}},\boldsymbol{B}) \triangleq \sum_{\ell=1}^L \|\bar{\boldsymbol{x}}_\ell - \mathrm{Diag}(\boldsymbol{D}_\ell(\boldsymbol{p})\boldsymbol{b}_\ell)\bar{\boldsymbol{s}}\|_2^2, \tag{16}$$

and here we write $\boldsymbol{D}_\ell(\boldsymbol{p})$ (rather than $\boldsymbol{D}_\ell$) to emphasize the dependence on $\boldsymbol{p}$. In contrast to the MFP3 solution (12), our proposed solution (15) is due to joint estimation of all the unknown model parameters $\boldsymbol{p}, \bar{\boldsymbol{s}}$ and $\boldsymbol{B} \triangleq [\boldsymbol{b}_1 \ldots \boldsymbol{b}_L] \in \mathbb{C}^{3 \times L}$, including the channel coefficients $\boldsymbol{B}$. Thus, in our proposed approach, we do *not* assume that the channel response is fully known for a given hypothesized position $\boldsymbol{p}$ of the source.

Intuitively, this approach should lead to a more robust solution than MFP3 with respect to deviations from the channel knowledge (9)–(11), at the cost of extra computational effort. Fortunately, as we show in Section IV, by exploiting the low-dimensional structure of the data, the additional computational cost is negligible. Moreover, although (15) defines a nonlinear high-dimensional optimization problem with $3L$ additional unknowns relative to MFP3, it boils down to a 3-dimensional optimization problem, similar to (14).

Our main result is the following localization algorithm:

The SBL Estimator:
Input: $\{\bar{\boldsymbol{x}}_\ell\}_{\ell=1}^L, c, h$, 3D grid of the volume of interest.
Output: The SBL estimate, $\widehat{\boldsymbol{p}}_{\text{SBL}}$.
1. For every candidate $\boldsymbol{p}$ on the grid:
1.1. Compute the matrices $\{\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\}_{\ell=1}^L$;
1.2. Compute the Cholesky decompositions

$$\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^* \triangleq \boldsymbol{\Gamma}_\ell^{\mathrm{H}} \boldsymbol{\Gamma}_\ell \in \mathbb{C}^{3 \times 3}, \ \forall \ell \in \{1, \ldots, L\}, \quad (17)$$

and obtain the matrices $\{\boldsymbol{\Gamma}_\ell\}_{\ell=1}^L$;
1.3. Compute the matrix

$$\boldsymbol{U}(\boldsymbol{p}) \triangleq [\bar{\boldsymbol{X}}_1 \boldsymbol{D}_1^* \boldsymbol{\Gamma}_1^{-1} \cdots \bar{\boldsymbol{X}}_L \boldsymbol{D}_L^* \boldsymbol{\Gamma}_L^{-1}] \in \mathbb{C}^{N \times 3L},$$
$$(18)$$

and construct the matrix

$$\widetilde{\boldsymbol{Q}}(\boldsymbol{p}) \triangleq \boldsymbol{U}(\boldsymbol{p})^{\mathrm{H}} \boldsymbol{U}(\boldsymbol{p}) \in \mathbb{C}^{3L \times 3L}. \quad (19)$$

1.4. Compute $\lambda_{\max}\left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right)$;
2. Find $\widehat{\boldsymbol{p}}_{\text{SBL,grid}}$, the maximizer point on the grid.
3. Return $\widehat{\boldsymbol{p}}_{\text{SBL}}$, the solution of a nonlinear optimization solver (e.g., trust-region [40]) initialized by $\widehat{\boldsymbol{p}}_{\text{SBL,grid}}$.

We now provide the analysis, based on which the algorithm above is derived. For convenience, we define

$$\boldsymbol{P}_{\bar{\boldsymbol{s}}} \triangleq \mathrm{Diag}(|\bar{\boldsymbol{s}}|^2) \triangleq P_s \cdot (\boldsymbol{I}_N + \boldsymbol{\mathcal{E}}) \in \mathbb{R}_+^{N \times N}, \quad (20)$$

where $P_s \in \mathbb{R}_+$ is the average signal power (with $|\cdot|^2$ operating elementwise), and $\boldsymbol{\mathcal{E}}$ is a diagonal matrix with "small" elements, such that $\varepsilon_{\max} \triangleq |\lambda_{\max}(\boldsymbol{\mathcal{E}})| < 1$.

Proposition 1 (SBL for Spectrally Flat Waveforms). *Consider the case where $\boldsymbol{\mathcal{E}} = \mathbf{O}$, and define the data-dependent matrix,*

$$\boldsymbol{Q}(\boldsymbol{p}) \triangleq \sum_{\ell=1}^L \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}} \in \mathbb{C}^{N \times N}, \quad (21)$$

for any hypothesized source position $\boldsymbol{p}$. Then,

$$\widehat{\boldsymbol{p}}_{\text{SBL}} = \underset{\boldsymbol{p} \in \mathbb{R}^{3 \times 1}}{\mathrm{argmax}} \ \lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p})). \quad (22)$$

Proposition 1, whose proof is given in Appendix A, tells us that, for spectrally flat waveforms $\bar{\boldsymbol{s}}$, the source's position

	MFP3	SBL
Unknowns	$\boldsymbol{p}, \bar{\boldsymbol{s}}$	$\boldsymbol{p}, \bar{\boldsymbol{s}}, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$
Objective function	$\sum_{k=1}^N \frac{|\bar{\boldsymbol{x}}[k]^{\mathrm{H}} \boldsymbol{h}_k(\boldsymbol{p})|^2}{\|\boldsymbol{h}_k(\boldsymbol{p})\|^2}$	$\lambda_{\max}\left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right)$
Complexity	$\mathcal{O}(NL)$	$\mathcal{O}(NL^3)$
Required physical parameters	$h, c, \kappa_b, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$	h, c

TABLE I: Comparison of the primary attributes of traditional MFP3 and the proposed estimator, SBL, for the three-ray model.

estimator can be computed based only on $\lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p}))$. Moreover, although our model has more unknowns, (22) is obtained by (only) a 3-dimensional optimization.

The next proposition, whose proof appears in Appendix B, states that the simplified form $\lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p}))$ of the objective function can be a good approximation to (15) when $\boldsymbol{\mathcal{E}} \neq \mathbf{O}$, namely for waveforms that are not spectrally flat. In turn, this implies that (22) can be used to localize a source emitting a general waveform.

Proposition 2 (SBL for General Waveforms). *Consider the case where $\boldsymbol{\mathcal{E}}$ is not necessarily equal to $\mathbf{O}$. Then,*

$$\widehat{\boldsymbol{p}}_{\text{SBL}} = \underset{\boldsymbol{p} \in \mathbb{R}^{3 \times 1}}{\mathrm{argmax}} \ \lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p})) + \mathcal{O}(\varepsilon_{\max}).$$

It follows that whenever $\varepsilon_{\max} \ll 1$, we have

$$\widehat{\boldsymbol{p}}_{\text{SBL}} \approx \underset{\boldsymbol{p} \in \mathbb{R}^{3 \times 1}}{\mathrm{argmax}} \ \lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p})).$$

However, as we demonstrate via simulations and real data in Section VI, our proposed estimator exhibits good performance even for waveforms that are far from being spectrally flat. Thus, Proposition 2 implies that only $\lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p}))$ is required for approximately optimal localization. In particular, it suffices to use, e.g., the power method, rather than computing the complete eigenvalue decomposition of $\boldsymbol{Q}(\boldsymbol{p})$. However, the computational complexity can be reduced even more, as implied by the following proposition, whose proof is given in Appendix C.

Proposition 3 (Efficient Computation of the SBL Objective Function). *Let $\boldsymbol{Q}(\boldsymbol{p}) \in \mathbb{C}^{N \times N}$ be defined as in (21). Then,*

$$\lambda_{\max}(\boldsymbol{Q}(\boldsymbol{p})) = \lambda_{\max}\left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right), \quad (23)$$

and the complexity of computing (23) is (only) $\mathcal{O}(NL^3)$.

Note that a naïve application of, e.g., the power method to $\boldsymbol{Q}(\boldsymbol{p})$ would cost $\mathcal{O}(N^2)$. This is already prohibitively expensive for reasonable sample sizes on the order of $N \sim 10^3$.

We emphasize that our proposed estimator implicitly optimizes over an additional $2 \cdot 3L$ unknown parameters—the channel attenuation coefficients $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$—relative to the MFP estimator of this model, while retaining the same order of computational complexity in terms of N (sample size). As an intermediate summary, a comparison of several attributes of the proposed SBL with MFP3 is given in Table I.

For the actual computation of the estimate $\widehat{\boldsymbol{p}}_{\text{SBL}}$, we propose a two-phased approach. The first phase consists of a coarse grid search over the relevant volume of interest. In the second phase, a general purpose nonlinear optimization algorithm (e.g., trust-region methods [40]) is applied, where the solution from the first phase is used for initialization.

A. Interpretation of the SBL Solution

We now provide a useful interpretation of the closed-form expression (22) of our proposed solution, based on the derivation presented in Appendix A. We begin by explaining the first step, the estimation of $\boldsymbol{b}_\ell$. From (36), when $\overline{\boldsymbol{S}}$ and $\boldsymbol{D}_\ell$ (defined in (7)) are treated as known, we see that this first step can be regarded as compensation (or, rectification) of the attenuations of each of the three signal components. It is also enlightening to see this from the noiseless case, where

$$\overline{\boldsymbol{x}}_\ell = \overline{\boldsymbol{S}}\boldsymbol{D}_\ell\boldsymbol{b}_\ell \implies \underbrace{\overline{\boldsymbol{S}}^{-1}\overline{\boldsymbol{x}}_\ell}_{\substack{\text{per-frequency} \\ \text{elementwise division}}} = \underbrace{\boldsymbol{D}_\ell\boldsymbol{b}_\ell}_{\substack{\text{per-frequency weighted} \\ \text{sum of } b_{1\ell}, b_{2\ell}, b_{3\ell}}} . \quad (24)$$

Substituting $\{\widehat{\boldsymbol{b}}_\ell\}$ (defined in (40)) into (16) yields after simplification (42)—the "$\boldsymbol{b}_\ell$'s-rectified" objective, where the rectification is based on the intermediate estimators $\{\widehat{\boldsymbol{b}}_\ell\}$, which still depend on the unknown $\overline{\boldsymbol{S}}$ and $\boldsymbol{p}$ at this phase.

Moving forward, we momentarily focus on a single (matrix) element of the sum (42). Rearranging this term, we see that

$$\overline{\boldsymbol{s}}^{\text{H}}\overline{\boldsymbol{X}}_\ell\boldsymbol{D}_\ell^*\left(\boldsymbol{D}_\ell^{\text{T}}\boldsymbol{P}_{\overline{s}}\boldsymbol{D}_\ell^*\right)^{-1}\left(\overline{\boldsymbol{X}}_\ell\boldsymbol{D}_\ell^*\right)^{\text{H}}\overline{\boldsymbol{s}} =$$
$$\overline{\boldsymbol{x}}_\ell^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\left(\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\right)^{-1}\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{x}}_\ell^*. \quad (25)$$

Again, focusing on the noiseless case to gain intuition, by substituting $\overline{\boldsymbol{x}}_\ell = \overline{\boldsymbol{S}}\boldsymbol{D}_\ell\boldsymbol{b}_\ell$, we have

$$\overline{\boldsymbol{x}}_\ell^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\left(\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\right)^{-1}\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{x}}_\ell^* =$$
$$\boldsymbol{b}_\ell^{\text{T}}\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\left(\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\right)^{-1}\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\boldsymbol{b}_\ell^* =$$
$$\boldsymbol{b}_\ell^{\text{T}}\boldsymbol{D}_\ell^{\text{T}}\overline{\boldsymbol{S}}^{\text{T}}\overline{\boldsymbol{S}}^*\boldsymbol{D}_\ell^*\boldsymbol{b}_\ell^* = \|\overline{\boldsymbol{S}}\boldsymbol{D}_\ell\boldsymbol{b}_\ell\|^2 = \sum_{k=1}^{N}\left|\overline{s}[k]\boldsymbol{d}_\ell^{\text{H}}[k]\boldsymbol{b}_\ell\right|_2^2.$$

Therefore, we interpret the maximization (42)—for a single receiver—as choosing the best set of parameters $\{\overline{\boldsymbol{s}}, \boldsymbol{p}, \boldsymbol{b}_\ell\}$, in the sense that the total energy of the received signal from the source is maximized, under the hypothesized set of parameters.

Generalizing this intuition for a signal in noise, after substituting $\overline{\boldsymbol{x}}_\ell$ into (25), for a sufficiently large N, the signal-noise cross product terms will tend to zero by virtue of the law of large numbers, since the noise DFT coefficients are uncorrelated and zero-mean.

Lastly, we generalize the intuition above from a single receiver to multiple receivers. For this, recall that (42) is in fact a *joint* maximization of the total energy of all L received signals from the same source. Therefore, it weights the L signals from different relative locations to the source while taking into account that they all contain shifted versions of *the same* waveform. This is essentially the connecting link, and the advantage in processing the data jointly (rather than individually). This joint weighting is nontrivial in the general case. However, when the source is spectrally flat, i.e., $\boldsymbol{\mathcal{E}} = \mathbf{O}$, the optimal way (in the sense of (15)) to weight and combine the data from the receivers is to form the matrix $\boldsymbol{Q}(\boldsymbol{p})$ as in (21), and to compute its maximal eigenvalue (22). A natural interpretation of the maximal eigenvalue of a semi-positive definite matrix is the energy distributed along the dominant direction (orthogonal to all others) in the space spanned by the columns of this matrix. With this interpretation, the final form of the SBL solution given in (22) is now intuitive.

V. THE CRAMÉR-RAO LOWER BOUND FOR SBL

We now analyze the localization accuracy limitations of the proposed solution in terms of the MSE,

$$\text{MSE}(\widehat{\boldsymbol{p}}, \boldsymbol{p}) \triangleq \mathbb{E}\left[\|\widehat{\boldsymbol{p}} - \boldsymbol{p}\|_2^2\right]. \quad (26)$$

Specifically, we derive the CRLB for the special case $\boldsymbol{\mathcal{E}} = \mathbf{O}$. Unlike the common approach (e.g., as in [41], [42]), wherein both the unknown source signal and noise are considered to be random, in our model only the noise is considered random. Thus, for a given waveform, the bound can be used as a tool for designing the deployment of a network of receivers, so as to maximize accuracy in regions of higher importance.

Regardless of the constant spectral level (i.e., $\boldsymbol{\mathcal{E}} = \mathbf{O}$), in our general framework $\overline{\boldsymbol{s}} \in \mathcal{S}_N$ w.l.o.g., hence in this particular case $P_s = \frac{1}{N}$. Consequently, $\overline{s}[k] = \frac{1}{\sqrt{N}}e^{\jmath\phi_s[k]}$ for all k, and the only waveform-related unknowns are the phases[8] of the DFT coefficients, denoted collectively by $\boldsymbol{\phi}_s \triangleq [\phi_s[2]\ldots\phi_s[N]] \in \mathbb{R}^{(N-1)\times 1}$.

To facilitate the following derivation, we introduce a more compact representation of the measured signals. Specifically, let $\overline{\boldsymbol{x}} \triangleq [\overline{\boldsymbol{x}}_1\ldots\overline{\boldsymbol{x}}_L]^{\text{T}} \in \mathbb{C}^{NL\times 1}$. Thus, (8) reads

$$\overline{\boldsymbol{x}} = \boldsymbol{H}\overline{\boldsymbol{s}} + \overline{\boldsymbol{v}} \in \mathbb{C}^{NL\times 1}, \quad (27)$$

where $\boldsymbol{H} \triangleq \left[\boldsymbol{H}_1^{\text{T}}\ldots\boldsymbol{H}_L^{\text{T}}\right]^{\text{T}} \in \mathbb{C}^{NL\times N}$ and $\overline{\boldsymbol{v}} \triangleq [\overline{\boldsymbol{v}}_1\ldots\overline{\boldsymbol{v}}_L]^{\text{T}}$. Denoting $\boldsymbol{\sigma}_v^2 \triangleq [\sigma_{v_1}^2\ldots\sigma_{v_L}^2]^{\text{T}} \in \mathbb{R}_+^{L\times 1}$, it follows that

$$\overline{\boldsymbol{x}} \sim \mathcal{CN}\left(\boldsymbol{H}\overline{\boldsymbol{s}}, \text{Diag}(\boldsymbol{\sigma}_v^2) \otimes \boldsymbol{I}_N\right). \quad (28)$$

It is well-known that for the CN signal model $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{R})$, the Fisher information matrix (FIM) elements are given by[9] [43]

$$J[\theta_i, \theta_j] = \text{Tr}\left(\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{R}}{\partial\theta_i}\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{R}}{\partial\theta_j}\right) + 2\Re\left\{\frac{\partial\boldsymbol{\mu}^{\text{H}}}{\partial\theta_i}\boldsymbol{R}^{-1}\frac{\partial\boldsymbol{\mu}}{\partial\theta_j}\right\},$$
$$\forall i, j \in \{1, \ldots, K_\theta\},$$

where we have defined the vector of all the real-valued unknown deterministic parameters

$$\boldsymbol{\theta} \triangleq \left[\boldsymbol{p}^{\text{T}} \text{vec}(\boldsymbol{\phi}_s)^{\text{T}} \text{vec}(\Re\{\boldsymbol{B}\})^{\text{T}} \text{vec}(\Im\{\boldsymbol{B}\})^{\text{T}} \boldsymbol{\sigma}_v^2\right]^{\text{T}} \in \mathbb{R}^{K_\theta\times 1}, \quad (29)$$

with $K_\theta = 3 + (N-1) + 2\cdot 3L + L$, and $\boldsymbol{J}(\boldsymbol{\theta})$ is the FIM.

It is readily seen from (28), that in our model the mean vector and covariance matrix are functions of distinct unknown parameters. This immediately implies that

$$J[\sigma_{v_{\ell_1}}^2, \sigma_{v_{\ell_2}}^2] = N \cdot \delta_{\ell_1\ell_2}, \forall \ell_1, \ell_2 \in \{1, \ldots, L\}, \quad (30)$$
$$J[\sigma_{v_\ell}^2, \theta] = 0, \qquad \forall\theta \neq \sigma_{v_\ell}^2, \quad (31)$$

---

[8]Note that although there are N elements in $\boldsymbol{\phi}_s$, there are only $N-1$ degrees of freedom, since the (complex-valued) channel attenuation coefficients are considered unknown as well. Therefore, we assume w.l.o.g. that the first element of $\boldsymbol{\phi}_s$, considered as a reference phase, is zero.

[9]For the sake of clarity, we specifically use a different notation for the FIM's elements, with slight abuse of notation also in (30)–(32).

namely the FIM has a block diagonal structure. Furthermore, for the signal-related block, we have

$$J[\theta_i, \theta_j] = 2\Re\left\{ \frac{\partial(\boldsymbol{H}\bar{\boldsymbol{s}})^{\mathrm{H}}}{\partial\theta_i} \left(\mathrm{Diag}^{-1}(\boldsymbol{\sigma}_v^2) \otimes \boldsymbol{I}_N\right) \frac{\partial\boldsymbol{H}\bar{\boldsymbol{s}}}{\partial\theta_j} \right\},$$
$$\forall i,j \in \{1,\ldots,K_\theta\}, \ \forall\ell \in \{1,\ldots,L\}. \tag{32}$$

When $\sigma_{v_\ell}^2 = \sigma_v^2$ for all ℓ, it can be observed from (32) that the signal-related FIM block is inversely proportional to the noise variance. Hence, the associated signal-related CRLB block is inversely proportional to the signal-to-noise ratio (SNR).

It only remains to compute the derivatives of $\boldsymbol{H}\bar{\boldsymbol{s}}$ with respect to the parameters of $\boldsymbol{\theta}$, excluding $\boldsymbol{\sigma}_v^2$, which is merely technical. We defer the details of these calculations, as well as the final expressions of all the signal-related elements of the FIM to the supplementary materials, along with a Matlab implementation of this bound. Finally, the CRLB is given by[10]

$$\mathbb{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^{\mathrm{T}}\right] \succeq \boldsymbol{J}^{-1}(\boldsymbol{\theta}) \triangleq \mathrm{CRLB}(\boldsymbol{\theta}) \in \mathbb{R}^{K_\theta \times K_\theta}$$
$$\implies \mathrm{MSE}(\hat{\boldsymbol{p}}, \boldsymbol{p}) \geq \sum_{i=1}^{3}[\mathrm{CRLB}(\boldsymbol{\theta})]_{ii}, \tag{33}$$

for any unbiased estimator $\hat{\boldsymbol{\theta}}$, and the implied $\hat{\boldsymbol{p}}$ (see (29)).

VI. SIMULATION AND EXPERIMENTAL RESULTS

In this section, we consider simulation and physical experiments of source localization for different scenarios in order to corroborate our analytical derivations. First, we begin by the evaluation and visualization of the CRLB for a hybrid signal, wherein the signal-related component, namely $\boldsymbol{H}\bar{\boldsymbol{s}}$ from (27), is synthetic, and the noise-related component, namely $\bar{\boldsymbol{v}}$ from (27), is taken from previously collected ambient noise recordings from the Kauai ACOMMS ONR MURI 2011 (KAM11) experiment [32]. Second, we simulate a different scenario, wherein the receivers are deployed in a linear formation. For this setting, we evaluate the performance with respect to varying SNR, model mismatch (to assess robustness), and missing LOS components due to occluders. In these simulations, we compare our proposed method to the MFP3 solution (14) and to GCC-PHAT [27], a TDOA-based localization method, which is considered as highly robust to multipath effects. In the third experiment we compare the algorithms on data recorded from a water tank testbed.

A. Validation of the CRLB for Ocean Ambient Noise

We consider a scenario with $L = 4$ receivers, in an area with bottom depth $h = 100$ m. The locations of the receivers and the source are given in Table II. The attenuation coefficients were drawn (once, and then fixed) independently from the circularly-symmetric CN distribution, such that $\mathbb{E}\left[|b_{rl}|^2\right] = 1$, with variance 0.1^2. The speed of sound was set to $c = 1500$ m/s, and the sample size to $N = 30$. We consider the case $\boldsymbol{\mathcal{E}} = \boldsymbol{O}$, such that the waveform's DFT coefficients are $\bar{s}[k] = \frac{1}{\sqrt{N}}e^{j\phi_s[k]}$, and the phases $\{\phi_s[k]\}_{k=2}^{N}$ were drawn[11] (once, and then fixed) independently from the uniform

[10]$\boldsymbol{A} \succeq \boldsymbol{B}$ is to be interpreted to mean that $\boldsymbol{A} - \boldsymbol{B}$ is semi-positive definite.
[11]Except for the (immaterial) $\phi_s[1] = 0$, due to our semi-blind setting.

	x [m]	y [m]	z [m]
Source Position, $\boldsymbol{p}$	200.7240	100.1661	30.6374
Receiver 1, $\boldsymbol{p}_1$	150	-175	20
Receiver 2, $\boldsymbol{p}_2$	75	-225	20
Receiver 3, $\boldsymbol{p}_3$	-50	-200	20
Receiver 4, $\boldsymbol{p}_4$	-150	-150	20

TABLE II: Positions of the source and the four receivers for the setting considered in Subsection VI-A, depicted in Fig. 3 (left), with $T_s = 10^{-3}$ s. Note that the source position is not located on a (discrete) grid point.
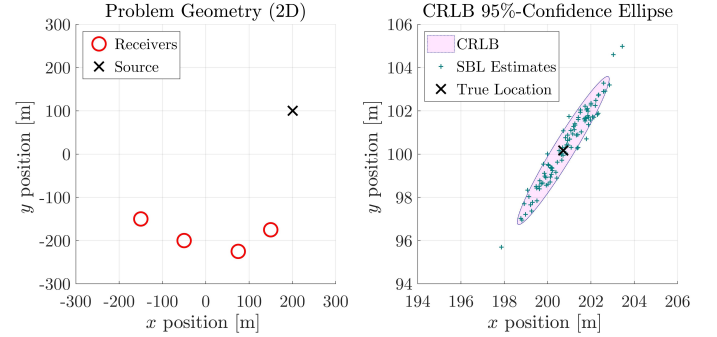

Fig. 3: Validation of the predicted asymptotic performance using the CRLB for the hybrid signal, containing recordings of ocean ambient noise collected in the KAM11 experiment. Left: The 2-dimensional setting of the scenario under consideration. Right: The 95%-confidence ellipse, as predicted by the CRLB (33), with 100 superimposed estimates.
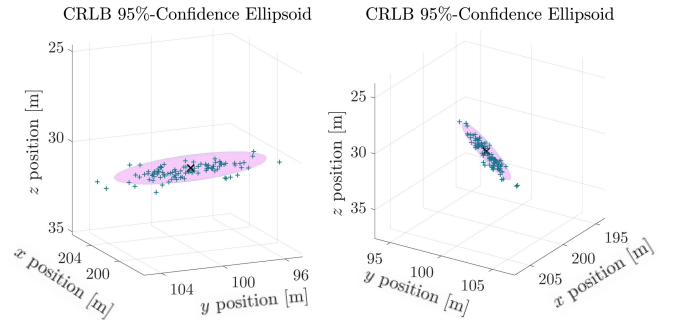

Fig. 4: Two different points of views for the (same) 3-dimensional 95%-confidence ellipsoid based on the CRLB, with 100 SBL estimates superimposed (legend as in Fig. 3, right). The CRLB accurately quantifies how the variance of the proposed solution is spread in the 3-dimensional space.

distribution $U(0, 2\pi)$. In this case, $\hat{\boldsymbol{p}}_{\mathrm{SBL}}$ is the MLE, and the CRLB accurately predicts its asymptotic variance. The received signals were generated according to (6), where the noise realization for all four sensors were taken from recordings of ocean ambient noise from the KAM11 experiment [32]. This way, we obtain a hybrid signal for this simulation, which allows us to test the validity of the bound on real ambient noise, which is potentially not CN and temporally white. Since the CRLB is informative only asymptotically (in the "small errors" regime), we set the noise variance[12] to $\sigma_{v_\ell}^2 = \sigma_v^2 = 0.1$ for all $\ell \in \{1, 2, 3, 4\}$, to have an SNR of $\|\bar{\boldsymbol{s}}\|_2^2/\sigma_v^2 = 10$ dB.

Figure 3 presents the 2-dimensional setting under consideration, and the 95% confidence ellipse computed using the CRLB (33), with superimposed estimates $\hat{\boldsymbol{p}}_{\mathrm{SBL}}$ obtained for 100

[12]We do so by first normalizing the recorded ambient noise to have unit variance, and then scale it accordingly to have the desired level of SNR.

	x [m]	y [m]	z [m]
Source Position, $\boldsymbol{p}$	100.5976	250.5837	30.1131
Receiver 1, $\boldsymbol{p}_1$	150	-250	10
Receiver 2, $\boldsymbol{p}_2$	50	-250	15
Receiver 3, $\boldsymbol{p}_3$	-50	-250	20
Receiver 4, $\boldsymbol{p}_4$	-150	-250	25
$c = 1535$ m/s, $\kappa_b = 0.85$, $h = 100$ m, $N = 100$, $T_s = 0.001$ s			

TABLE III: The setting considered in Subsection VI-B.

different noise recordings. Despite the model mismatch with respect to the noise distribution, a good fit is seen between the empirical results and the predicted theoretical accuracy due to the CRLB. Figure 4 reflects the same fit in the 3-dimensional space. This not only agrees with our analytical derivation of the bound, but also provides an empirical justification for our stochastic noise model. In this regard, we note further that the hybrid signals we use allow us to essentially isolate the (potential) noise-related model mismatch effects, and test our proposed solution with respect to deviations of this sort only.

B. Comparison with GCC-PHAT and MFP3

We now compare the proposed SBL method with MFP3 and the GCC-PHAT localization methods. In this simulated experiment, we consider the setup depicted in Fig. 5, namely a linear deployment of the receivers. Such a deployment is conceivable for naval defense purposes near the shoreline, or harbor monitoring [44]. The positions of the source and the receivers, and all relevant system and environmental parameters are given in Table III. The source's DFT coefficients $\bar{s}[k]$, as well as the noise realizations $\{\bar{v}_\ell[k]\}$, were drawn independently from the standard CN distribution in each trial. The noise variance of the ℓ-th receiver is set as $\sigma_{v_\ell}^2 = \sigma_v^2 \cdot \|\boldsymbol{b}_\ell\|^2$, and the SNR is defined here as $\mathbb{E}\left[|\bar{s}[k]|^2\right]/\sigma_v^2 = 1/\sigma_v^2$. All empirical results presented in this subsection are based on averaging 10^4 independent trials.

We first compare the localization accuracies of the methods for different SNRs. Figure 6 presents the root mean squared (RMS) miss distance, i.e., the square root of (26), vs. the SNR for each method. A 2D slice at the receiver's depth of the objective functions of each of the algorithms for a typical realization at 5 dB SNR is given in Fig. 5. For MFP3, we show the performance obtained with perfect knowledge of the channel ("Perfect Model"), i.e., when $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$ and κ_b are known exactly; and when this perfect knowledge is accurate except for the phases of $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$ ("Imperfect Model"), which are drawn independently from $U(0, 2\pi)$. As observed, although GCC-PHAT improves when the SNR increases, it essentially cannot cope well—in a 3-dimensional space optimization—with the addition of the surface and bottom reflections. It is also seen the MFP3 is highly sensitive to deviations from the assumed channel response. In contrast, such deviations are completely transparent to SBL, as it considers these parameters as unknown, and implicitly optimizes over them jointly with all the other unknowns (see (36), Appendix A). The robustness at the moderate cost in performance relative to MFP3 is evident. Note that MFP3's superior performance is guaranteed only asymptotically, in agreement with the results in Fig. 6.

Next, we compare the performances of the three different methods with respect to perturbations in the expected channel attenuations, as prescribed by the physical model (9)–(11). This form of model mismatch is likely to occur in practice due to nonidealities[13] (e.g., inaccurate prior knowledge of κ_b). Formally, we model these deviations by generating the channel attenuation coefficients as

$$b_{r\ell}(\epsilon) = (1 - \epsilon \cdot \gamma_{r\ell})b_{r\ell} \cdot e^{j2\pi\epsilon \cdot \varphi_{r\ell}},$$
$$r \in \{1, 2, 3\}, \ \ell \in \{1, 2, 3, 4\}, \tag{34}$$

where $\{\gamma_{r\ell} \sim U(0, 0.5)\}$ and $\{\varphi_{r\ell} \sim U(0, 1)\}$ are independent. In (34), $\epsilon \in [0, 1]$ is a parameter controlling the deviation from the physical model, where $\epsilon = 0$ corresponds to no deviation from (9)–(11).

Figure 7 presents the RMS miss distance vs. ϵ. As expected, we observe an overall accuracy-robustness superiority of SBL relative to the competing algorithms. While MFP3 is superior when perfect knowledge of the channel parameters $\kappa_b, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_L$ is available, SBL is inherently indifferent to deviations from their ideal physical values. GCC-PHAT is also robust to such deviations, but completely ignores (by design) the multipath channel, and therefore cannot exploit additional signal components, such as surface and bottom reflections.

In the last simulation for this setup, we model the effect of a potential occluder between some of the receivers and the source. Specifically, for the second and third receivers (i.e., at $\boldsymbol{p}_2$ and $\boldsymbol{p}_3$), we introduce an attenuation coefficient $\beta \in [0, 1]$ to the LOS components, such that

$$b_{1\ell}(\beta) = \beta \cdot b_{1\ell} \cdot e^{j2\pi(1-\beta)\varphi_{1\ell}}, \ell \in \{2, 3\}. \tag{35}$$

When $\beta = 1$, there is no occlusion, and when $\beta = 0$, the LOS components of receivers $\ell = 2, 3$ are completely lost. The phase perturbation models the interaction with the occluder.

Figure 8 presents the RMS miss distance vs. β. It is observed that the accuracy obtained by GCC-PHAT is on the same order of the distances (~ 25m) corresponding to the time delays between the LOS and NLOS components. This level of accuracy is stable, but is not satisfactory for an SNR level of 10 dB (here, $\sigma_v^2 = 0.1$). It is also seen that in the absence of a modeling error (i.e., $\beta = 1$), MFP3 attains the highest accuracy. However, deviations from the ideal signal model, in the form of occluded LOS components of two receivers, inflict a severe performance deterioration.

As we have demonstrated in Fig. 5c and Fig. 6, when MFP3 is actually mis-matched, it could perform even worse than what is presented in Fig. 8. Still, even in this setting, our method exhibits the best accuracy-stability trade-off.

C. Experimental Results

We now demonstrate the performance of our proposed method on acoustic data acquired in our water tank testbed—the high frequency autonomous acoustic tank. This system, presented in Fig. 9, is (roughly) of size 25 cm $\times 32$ cm $\times 15$ cm, and enables us to create a controlled and challenging setting for frequency-scaled underwater localization.

---

[13]Within the three-ray model. Of course, in practice there are more modeling mismatch factors due to the simplified three-ray model. The effects of some of these nonidealities will be evaluated in the next experiment, where we apply our method to real data, and demonstrate successful localization.
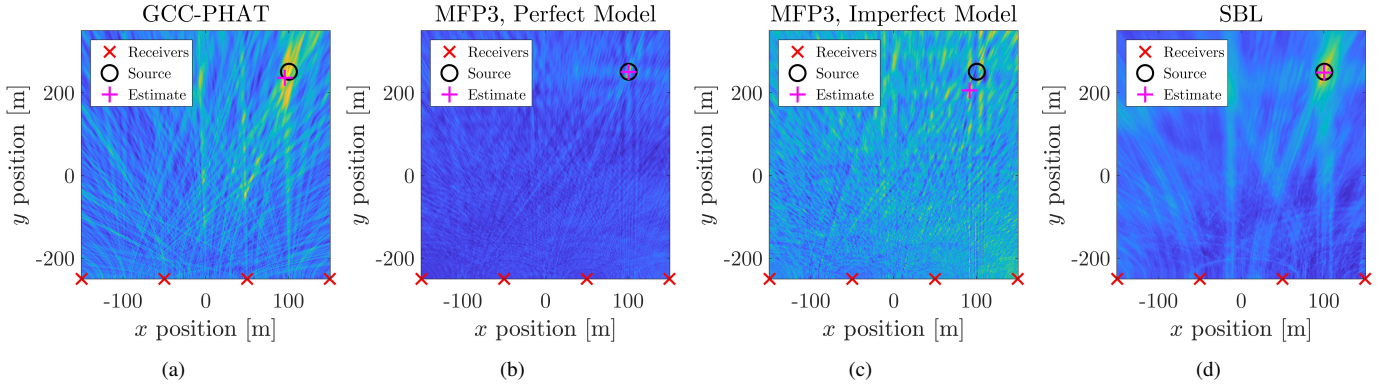
Fig. 5: 2-dimensional slices at the source depth ($z = z_p$) of the objective function for typical realizations at 5 dB SNR. (a) GCC-PHAT (b) MFP3 with perfect model (c) MFP3 with unknown phases of $\boldsymbol{b}_1 \ldots, \boldsymbol{b}_L$ (d) SBL. Evidently, GCC-PHAT and MFP3 without perfect knowledge are considerably more fragile than the proposed method, which is similar to MFP3 with perfect knowledge with respect to stability, at the cost of a higher variance.
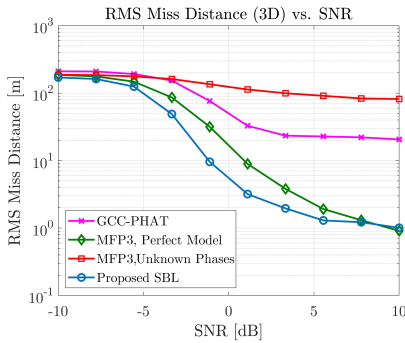

Fig. 6: RMS miss distance vs. SNR, for $T = 100$. At very low SNR, all methods perform poorly. As the SNR increases, GCC-PHAT improves only moderately, while SBL improves significantly.
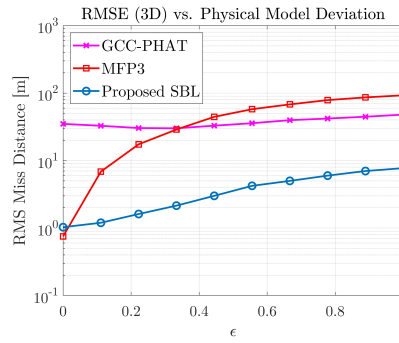
Fig. 7: RMS miss distance vs. ϵ, quantifying a deviation from the physical model, for $T = 100$ and $\sigma_{v_\ell}^2 = 0.1$. MFP3 is sensitive to such model deviations, while GCC-PHAT and SBL are robust.
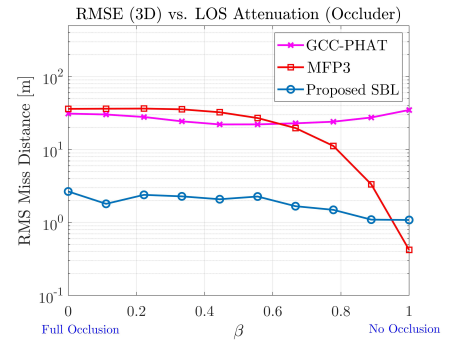
Fig. 8: RMS miss distance vs. β, the LOS attenuation coefficient of receivers $\ell = 2, 3$, for $T = 100$ and $\sigma_{v_\ell}^2 = 0.1$. Our proposed SBL offer the best accuracy-stability balance out of the three methods.
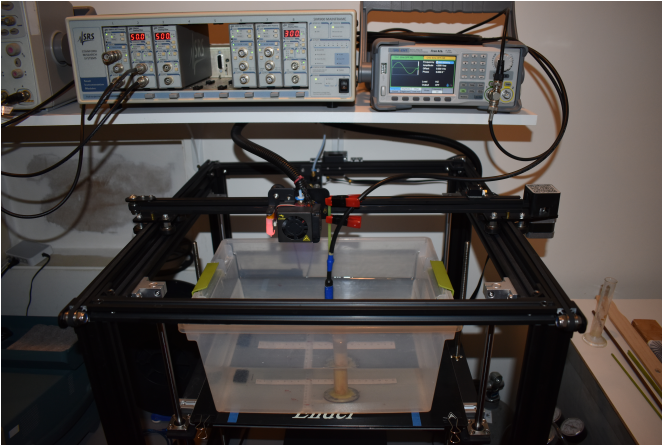

Fig. 9: A picture of our water tank. Our testbed provides high frequency ($200 - 400$ kHz) noisy acoustic data from a complex reverberant environment.

Although the water tank environment is only a scale model of a shallow-water environment, it nevertheless poses a challenging scenario. In addition to the modeled bottom and surface reflections, the water tank has four additional sides, that are reflective boundaries. These thin plastic boundaries are highly reflective, so that the test environment is highly reverberant, giving rise to a rich multipath channel. In particular, the magnitudes of the unmodeled reflections are comparable to the modeled ones in our three-ray model.

In this experiment, the source is transmitting a Gaussian

pulse at a carrier frequency of 280 kHz, and the speed of sound in the water tank is $c = 1485$ m/s. To maintain consistency across different trials, the source and receivers were set at the same depth. In this case, spatial diversity in the depth-direction is limited, hence we assume here that the source's depth is known, and approach this 2-dimensional problem.

The received signals were sampled at 2 GHz. Before applying the localization method, the signals were decimated by a factor of 2×10^3, to obtain 1 MHz bandwidth signals. For each setting, in which the source and receivers were static, the observation interval was 0.5 ms long. More technical details are given in the supplementary materials.

Figs. 10, 11 and 12 present the objective functions of GCC-PHAT, MFP3 and the proposed method, respectively, for three different source locations. Here, the search area is a 30×30 cm^2 square, centered around the first receiver. Note that we intentionally did not align the search area with the one dictated by the boundaries of the tank, since we assume that such prior knowledge is unavailable as in a real problem setting. As was observed in the simulations, it is seen that in the presence of strong multipath, GCC-PHAT suffers from the worst performance degradation, and MFP3 is the most accurate, best exploiting the environmental prior knowledge. SBL is less accurate than MFP3, but still provides reasonable estimates in the vicinity of the source's true location.

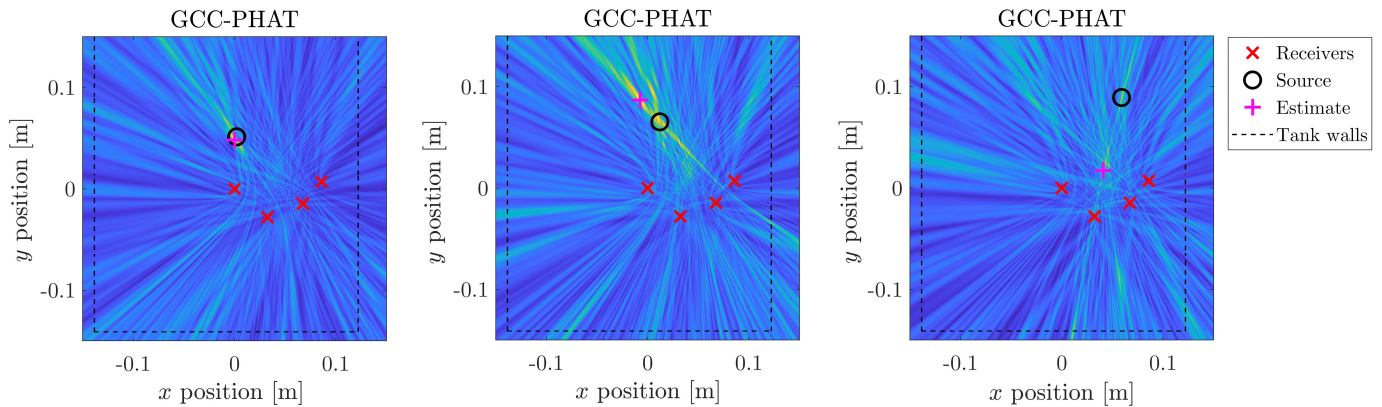Next, we repeat the experiment but now with the presence of

Fig. 10: GCC-PHAT experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9.
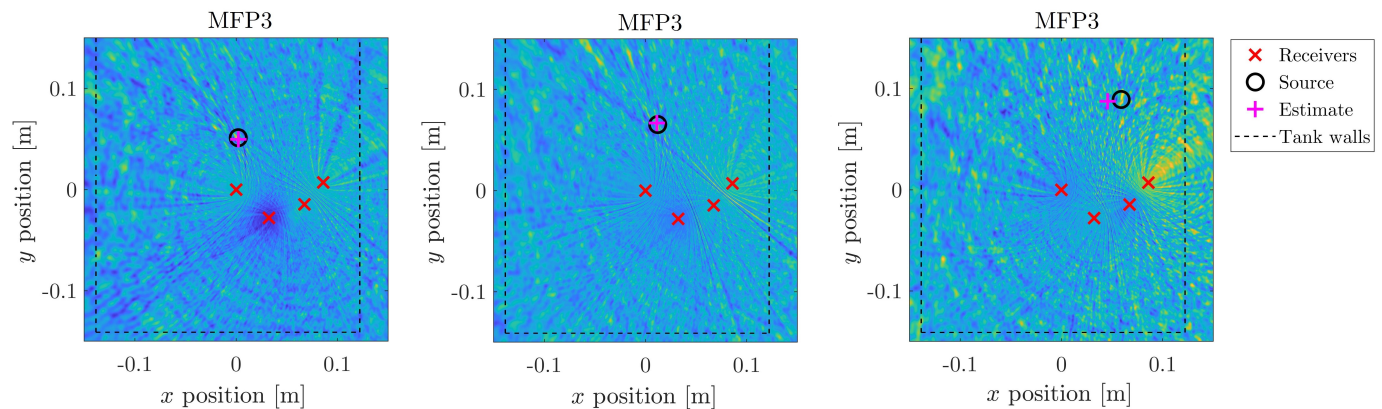

Fig. 11: MFP3 experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9.
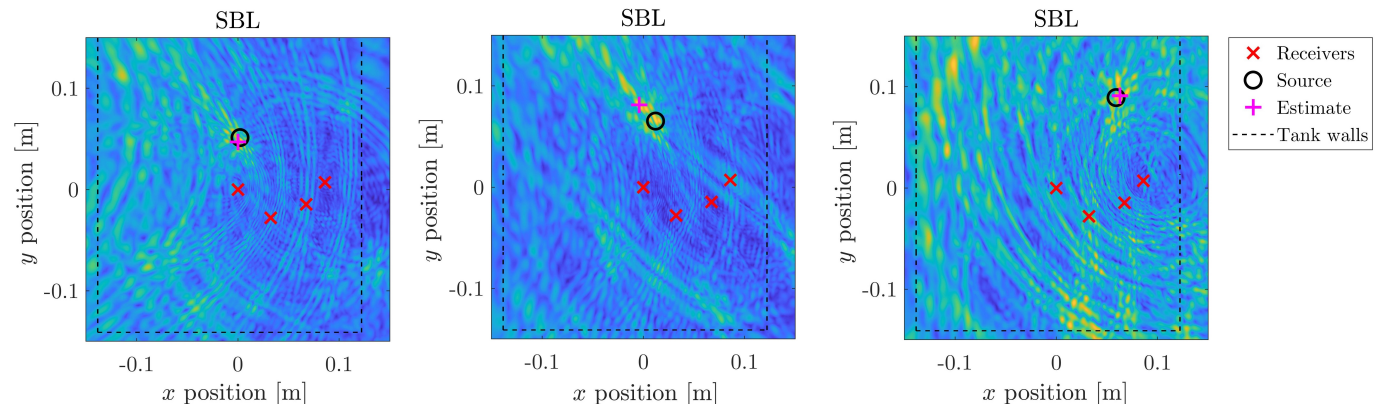

Fig. 12: SBL experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9.

an unknown object—a cylinder, stretching from the bottom to the surface of the tank, placed in the area between the source and the receivers, as depicted in Figs. 13–15. This unknown feature causes severe model mismatch; critically, if it blocks the LOS between the source and a receiver, then all the three modeled rays—LOS, and surface and bottom reflections—are essentially blocked. A flexible algorithm can in principle select (possibly implicitly) which receivers to use, and would be able to reject uninformative measurements, such as the ones acquired by a receiver "viewing" the occluded scenery.

Figures 13, 14 and 15, presenting the objective functions of GCC-PHAT, MFP3 and SBL, respectively, for the same scenarios but with an occluder, corroborate the robustness of the SBL method. It is seen that GCC-PHAT and MFP3 are fragile when such unknown environmental features are present. The SBL method, which exhibits robustness in the presence of the occluder, is still able to localize the source. Since the attenuation coefficients are considered to be unknown, and are implicitly estimated (36), a perfectly valid estimated value (for some of them) is a value close or equal to zero. This essentially means that the SBL assigns different weights to measurements from different receivers, thus implicitly choosing to effectively ignore the less informative data acquired by receivers with occluded scenery.
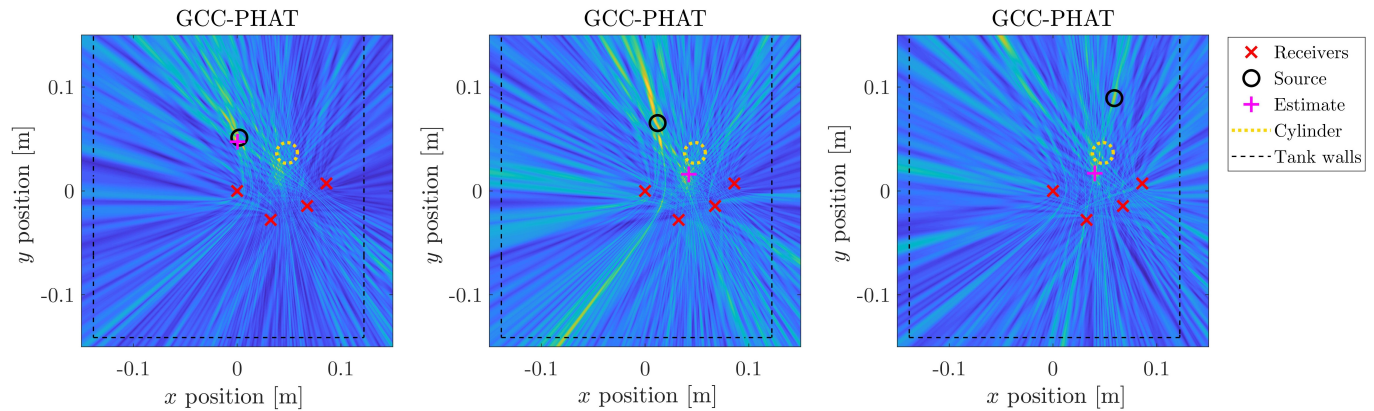
Fig. 13: GCC-PHAT experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9, with the presence of an unknown occluder (cylinder), modeling an effect of inaccurate environmental prior knowledge.
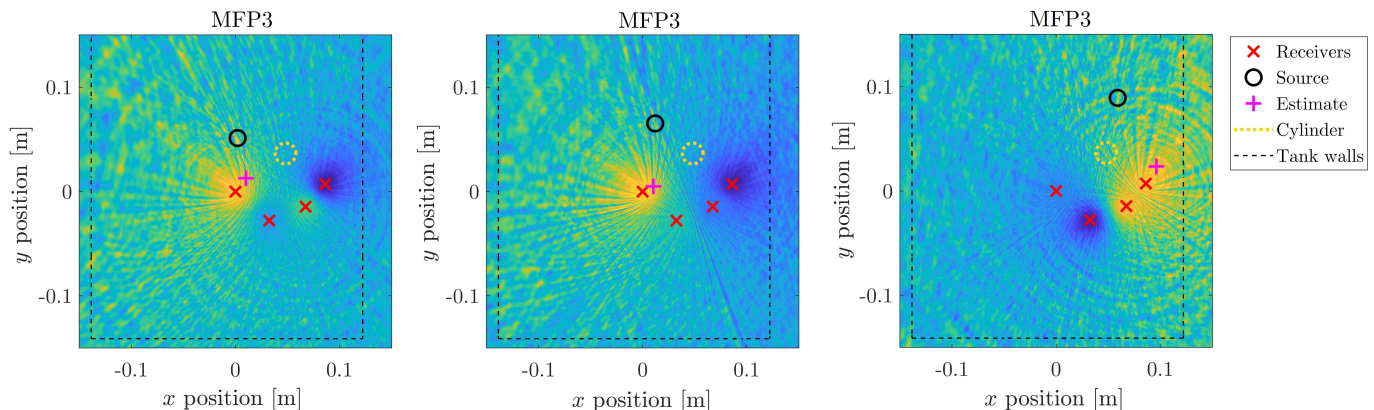
Fig. 14: MFP3 experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9, with the presence of an unknown occluder (cylinder), modeling an effect of inaccurate environmental prior knowledge.
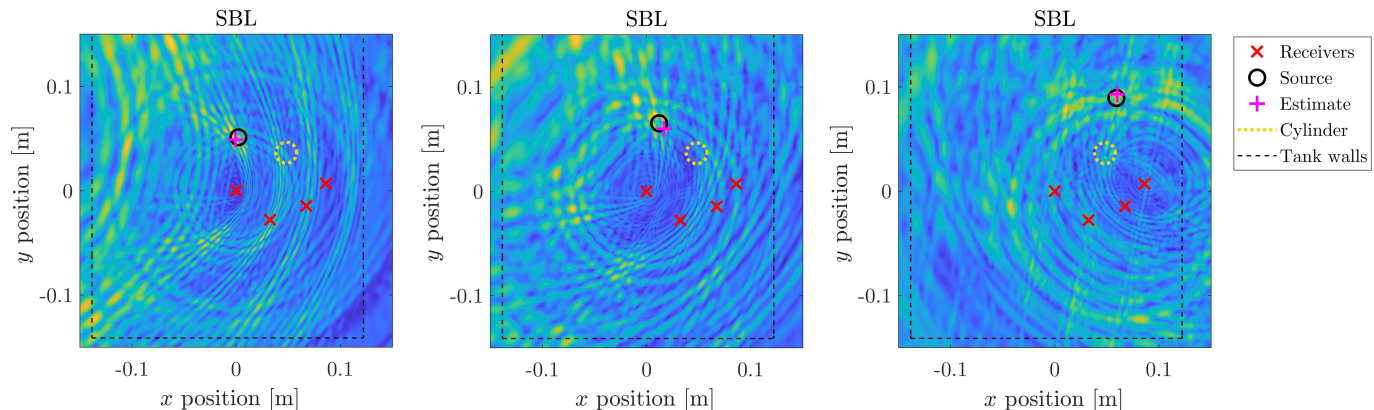
Fig. 15: SBL experimental results based on acoustic measurements acquired in the water tank using the system presented in Fig. 9, with the presence of an unknown occluder (cylinder), modeling an effect of inaccurate environmental prior knowledge.

VII. DISCUSSION AND EXTENSIONS

A natural extension of the direct localization problem that we considered, is to the case of multiple sources and/or a more complex channel model, which is beyond the scope of this paper and is left for future work. However, in this section we outline some key challenges of this setting to motivate this non-trivial extended problem. We then discuss the potential use of the SBL estimator in such scenarios, and point out important aspects of identifiability. Before addressing these topics, we first discuss some system design considerations for the particular setting described in Section II.

A. System Design Considerations

For given, limited resources, one may be interested in enhancing the performance of a system as much as possible with respect to the available degrees of freedom. In the specific case of our localization problem, performance can be understood as accuracy (e.g., in terms of (26)), and resources, perhaps, as the number of receivers, L.[14] While we defer the formulation of

[14]which are also the number of sensors in our formulation.

this notion into a well-defined problem (possibly using some function of the bound (33)) to future work, we comment on the important related aspects of the topology and the number of receivers.

Recall that a single receiver, equipped with a single sensor, receives (in general) three signal components, two of which are reflected from the surface and the bottom, which are respectively above and under the receiver. Thus, environmental knowledge regarding the position of the surface and the bottom is equivalent, in some sense, to having additional virtual receivers (e.g., [45]) above and under the surface and the bottom, respectively. In this respect, a single receiver already provides some vertical spatial diversity. Still, due to the blind nature of the problem, in which the emitted waveform and the channel coefficients are unknown, the information from a single sensor is insufficient for localization.

However, two sensors already contain six signal components, and can in principle contain sufficient information for localization. Intuitively, breaking the symmetry "as much as possible" relative to the environment in which the system is deployed would lead to better performance. In the two receivers case, placing the second receiver in a different horizontal location than the first would lead to an increased horizontal spatial diversity, which would in turn lead to enhanced performance. The principle of increasing spatial diversity with a fixed resources allocation can be formulated in some settings (e.g., [46]), and provides guidance and intuition for the design of a sensors network spatial distribution. For example, one may consider deploying a linear network of sensors *obliquely* relative to the surface.

B. Key Challenges in Extended Models

Incorporating multiple sources into the 3-ray signal model (5) changes the interplay between the (consequently increased) number of unknown parameters. As a result, it is no longer clear whether a simplified, efficiently computable expression for the objective function—as (22) in Proposition 1—can be obtained. Recall that this has a significant effect on the overall computational complexity of the method.

Moreover, assume that there are M sources to be localized (where M is known), and further assume that we have obtained such a simplified, computationally efficient expression for the objective function, which is a function of the sources' positions *only*, denoted by, say, $\boldsymbol{p}^{(1)}, \dots, \boldsymbol{p}^{(M)} \in \mathbb{R}^{3 \times 1}$. At this point, in order to obtain the optimal direct localization solution (in the sense of the extended criterion of (15)) for all M sources, one is required to solve a $3M$-dimensional nonlinear optimization problem, which may well be non-convex. Consequently, even for $M = 2$ sources, this is already difficult with a naive extension of our current proposed method, as it would require a 6-dimensional grid search (referring to the first step of the proposed solution), which is infeasible for reasonable resolutions. Hence, a different approach is perhaps required in order to solve the multiple sources direct localization problem.

Focusing again on the single source case, one may consider an extended K-ray model (with $K > 3$), assuming it would *accurately* describe the signal propagation, such that the K rays include primary *and* second- and higher-order reflections. In

that case, the performance (i.e., accuracy) improvement would be due to an increased effective/post-processing SNR. This can be understood from the interpretation given in Subsection IV-A, where the SBL method is seen as an implicit way to coherently add all the K reflections from all L sensors.

However, and since the ray-based propagation model is an approximation, while the deviations from the 3 primary rays can be small, the aggregated approximation errors in the time-delays of the higher-order reflections are likely to no longer be negligible. In that case, on top of additional computational burden, a naive extension of the current approach might yield a more fragile estimator, which is sensitive to model mismatch. The challenging task of exploiting more complex propagation-related phenomena for enhanced, computationally attractive direct localization remains to be explored in future work.

C. SBL as a Solution for Multiple Sources

Notwithstanding the above, our proposed algorithm can still be used for multiple sources localization as a sub-optimal, yet computationally feasible solution. Indeed, (22) can be viewed as a spatial quasi-likelihood map (as a function of $\boldsymbol{p}$), whose M highest maxima correspond to the M points in space, where sources are most likely to be present (under the mismatched model (5), treating, for each source, all the other $M-1$ sources as additive noise). While providing analytical guarantees for this case is beyond the scope of the current work, using the Matlab package provided in the supplementary material, one could easily verify that the SBL method still serves as a viable localization solution for this extended setting.

Given any set of parameters that describe a particular localization problem (i.e., bottom depth, locations of the receivers, etc.), the model (5) is guaranteed to be identifiable when the FIM is nonsingular, namely $\det(\boldsymbol{J}(\boldsymbol{\theta})) \neq 0$, and (33) is finite. However, when using the SBL for localization of multiple source (i.e., under mismatched model) as described above, this is obviously no longer true. Indeed, as least theoretically, there are certain "special" (however somewhat extreme) scenarios in which the sources not only could not be localized (reliably, or at all), but may also "disappear" from the resulting heatmap.

To illustrate this, consider the following case, which is depicted in Fig. 16. Assume, for example, that L receivers are all deployed at the same depth $h/2$ in a linear structure, namely $z_\ell = z_0 = h/2, y_\ell = y_0$ and $x_\ell = x_0 + \Delta(\ell - 1)$ for all $\ell \in \{1, \dots, L\}$, where Δ is the spacing between the receivers. Now, further assume that *two* sources are present, such that the second source is located at the same horizontal location as the first, but is located symmetrically about the half depth $h/2$ relative to the first. That is, if the first source is at $\boldsymbol{p} = (x_p, y_p, z_p)$, the second is at $\widetilde{\boldsymbol{p}} = (x_p, y_p, h - z_p)$. Finally, assume that an occluding object is present, such that (only) all LOS components at all L receivers are blocked from both of the sources. Denoting by $\widetilde{R}_{2\ell}, \widetilde{R}_{3\ell}$ the distances traveled by the NLOS surface and bottom associated rays, respectively, from the second source, it readily follows from (2)–(4) that $\widetilde{R}_{2\ell} = R_{3\ell}, \widetilde{R}_{3\ell} = R_{2\ell}$, hence

$$\tau_{2\ell}(\boldsymbol{p}) = \tau_{3\ell}(\widetilde{\boldsymbol{p}}), \quad \tau_{3\ell}(\boldsymbol{p}) = \tau_{2\ell}(\widetilde{\boldsymbol{p}}).$$
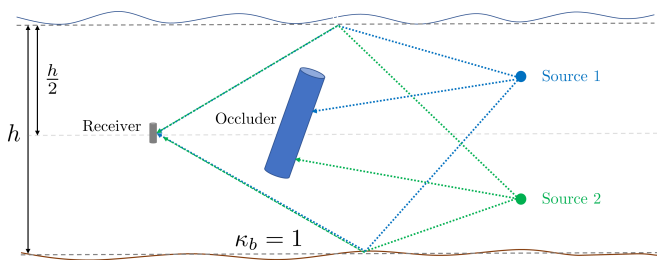
Fig. 16: A 2-dimensional illustration of the special case described in Subsection VII-C, wherein the sources are "acoustically invisible" in terms of the three-ray model, and cannot be localized by the SBL estimator.

In this case, if $\kappa_b = 1$ (of (11)), and if the two sources are collaborating and coordinated, then by transmitting the same waveform, they are essentially "acoustically invisible" (under the three ray model). Indeed, if we denote the waveform of the second source by $\widetilde{s}(t) = s(t)$, the baseband-converted signal from the ℓ-th receiver (as in (5)) would then be

$$x_\ell[n] = \sum_{r=2}^{3} b_{r\ell}\, s\left(t - \tau_{r\ell}(\boldsymbol{p})\right)\big|_{t=nT_s} +$$

$$\underbrace{\sum_{r=2}^{3} \widetilde{b}_{r\ell}\, \widetilde{s}\left(t - \tau_{r\ell}(\widetilde{\boldsymbol{p}})\right)\big|_{t=nT_s}}_{=-\sum_{r=2}^{3} b_{r\ell}\, s(t-\tau_{r\ell}(\boldsymbol{p}))|_{t=nT_s}} + v_\ell[n] = v_\ell[n],$$

since $\widetilde{b}_{2\ell} = \frac{-1}{R_{3\ell}} = -b_{3\ell}$ and $\widetilde{b}_{3\ell} = \frac{1}{R_{2\ell}} = -b_{2\ell}$, and we recall that due to the occluder, $b_{1\ell} = \widetilde{b}_{1\ell} = 0$ for all $\ell \in \{1, \ldots, L\}$. Thus, only noise is observed, and all the information is lost.

While the scenario above can certainly inspire underwater acoustic warfare techniques devised against single-source methods like the SBL, it nonetheless describes an extreme case of a perfectly tailored setting, where several conditions, which are exceptionally difficult to ensure, are fulfilled simultaneously. Therefore, and while this is only one example of a potential failure mechanism of the proposed method when used for localization of multiple sources, one may still gain a general impression of what should happen in order for SBL to completely fail in this setting. Generally, when the sources are not collaborating/coordinated and/or the symmetry is broken (e.g., by a non-regular deployment of the receivers), it is reasonable that the proposed method would still provide reliable localization for multiple sources.

VIII. CONCLUDING REMARKS

In the general context of underwater acoustics, based on the three-ray propagation model, we presented a semi-blind localization method, which incorporates environmental knowledge. A closed-form expression for the objective function was derived, along with an equivalent expression, which is more computationally appealing. Thanks to additional degrees of freedom in our model, the proposed method is more robust than its MFP counterpart, and can successfully localize a source in the absence of LOS components. Further, it exhibits stable performance enhancement with respect to methods modeling only LOS components, such as GCC-PHAT.

Since the proposed method is able to coherently "collect" three signal components from each sensor, the post-processing

SNR is consequently higher than any LOS-based method, which collect only one. This way, a given level of localization accuracy can generally be attained with shorter observation intervals. In turn, it is easier to incorporate the SBL method within an appropriate tracking algorithm (such as Kalman filtering), allowing for a more general framework that localizes the source and tracks its movement, assuming the source's velocity is sufficiently low. As a topic for future research, in such cases it may be possible to develop a computationally efficient update scheme for the objective function (23), based on eigenvalue perturbation theory. Another direction for future research, that is of great practical interest, is to apply coarse quantization to the collected data [47], thus reducing the required bandwidth for communication between the different receivers. The above is also true for the potential extensions for multiple sources, and for extended propagation models.

IX. ACKNOWLEDGEMENT

APPENDIX A
PROOF OF PROPOSITION 1

Proof. Using the identity $\mathrm{Diag}\left(\boldsymbol{D}_\ell(\boldsymbol{p})\boldsymbol{b}_\ell\right)\bar{\boldsymbol{s}} = \bar{\boldsymbol{S}}\boldsymbol{D}_\ell\boldsymbol{b}_\ell$, it is easily seen that, for every $\ell \in \{1, \ldots, L\}$, $\widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p}, \bar{\boldsymbol{s}}, \boldsymbol{B})$ of (16) is minimized with respect to $\boldsymbol{b}_\ell$ by

$$\widehat{\boldsymbol{b}}_\ell = \left(\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{-1}\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{x}}_\ell, \qquad (36)$$

assuming[15] $\mathrm{rank}\,(\boldsymbol{D}_\ell) = 3$ for all $\ell \in \{1, \ldots, L\}$ hereafter. Substituting $\widehat{\boldsymbol{B}} \triangleq [\widehat{\boldsymbol{b}}_1 \cdots \widehat{\boldsymbol{b}}_L]$ into $\widetilde{C}(\boldsymbol{p}, \bar{\boldsymbol{s}}, \boldsymbol{B})$ yields

$$\breve{C}(\boldsymbol{p}, \bar{\boldsymbol{s}})_{\mathrm{SBL}} \triangleq \widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p}, \bar{\boldsymbol{s}}, \widehat{\boldsymbol{B}}) = \sum_{\ell=1}^{L}\left\|\bar{\boldsymbol{x}}_\ell - \bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell\right\|_2^2$$

$$= \sum_{\ell=1}^{L}\left[\bar{\boldsymbol{x}}_\ell^{\mathrm{H}}\bar{\boldsymbol{x}}_\ell - \bar{\boldsymbol{x}}_\ell^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell - \widehat{\boldsymbol{b}}_\ell^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{x}}_\ell\right.$$

$$\left. + \widehat{\boldsymbol{b}}_\ell^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell\right]. \quad (37)$$

From (36), we observe that

$$\widehat{\boldsymbol{b}}_\ell^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell = \widehat{\boldsymbol{b}}_\ell^{\mathrm{H}}\left(\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\right)^{\mathrm{H}}\bar{\boldsymbol{x}}_\ell,$$

with which (37) simplifies to

$$\breve{C}_{\mathrm{SBL}}(\boldsymbol{p}, \bar{\boldsymbol{s}}) = \underbrace{\sum_{\ell=1}^{L}\bar{\boldsymbol{x}}_\ell^{\mathrm{H}}\bar{\boldsymbol{x}}_\ell}_{\substack{\text{constant with respect to} \\ \boldsymbol{p} \text{ and } \bar{\boldsymbol{s}}}} - \sum_{\ell=1}^{L}\bar{\boldsymbol{x}}_\ell^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell. \quad (38)$$

Therefore, using (38), (15) can now be written as

$$\min_{\substack{\bar{\boldsymbol{s}}\in\mathcal{S}_N \\ \boldsymbol{B}\in\mathbb{C}^{3\times L}}} \widetilde{C}_{\mathrm{SBL}}(\boldsymbol{p}, \bar{\boldsymbol{s}}, \boldsymbol{B}) = \max_{\bar{\boldsymbol{s}}\in\mathcal{S}_N}\sum_{\ell=1}^{L}\bar{\boldsymbol{x}}_\ell^{\mathrm{H}}\bar{\boldsymbol{S}}\boldsymbol{D}_\ell\widehat{\boldsymbol{b}}_\ell. \quad (39)$$

[15]We ignore the extreme, unrealistic cases in which $\boldsymbol{D}_\ell$ are not full rank, which occur only for very specific settings of the receivers' and source's positions. Nonetheless, the initial optimization is performed via a grid search, hence we can discard points giving rise to these rare, singular settings.

At this point, notice that using

$$\bar{\boldsymbol{x}}_\ell^{\mathrm{H}} \bar{\boldsymbol{S}} = \bar{\boldsymbol{s}}^{\mathrm{T}} \bar{\boldsymbol{X}}_\ell^* \Longrightarrow \left(\bar{\boldsymbol{S}} \boldsymbol{D}_\ell\right)^{\mathrm{H}} \bar{\boldsymbol{x}}_\ell = \left(\bar{\boldsymbol{X}}_\ell^{\mathrm{H}} \boldsymbol{D}_\ell\right)^{\mathrm{H}} \bar{\boldsymbol{s}}^*,$$

we may write

$$\widehat{\boldsymbol{b}}_\ell = \left(\boldsymbol{D}_\ell^{\mathrm{H}} \bar{\boldsymbol{S}}^{\mathrm{H}} \bar{\boldsymbol{S}} \boldsymbol{D}_\ell\right)^{-1} \boldsymbol{D}_\ell^{\mathrm{H}} \bar{\boldsymbol{X}}_\ell \bar{\boldsymbol{s}}^*. \tag{40}$$

Substituting $\widehat{\boldsymbol{b}}_\ell$ from (40) into (39), using (40), $\bar{\boldsymbol{S}}^{\mathrm{H}} \bar{\boldsymbol{S}} = \mathrm{Diag}(|\bar{\boldsymbol{s}}|^2) \triangleq \boldsymbol{P}_{\bar{\boldsymbol{s}}}$ and simplifying further yields

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \sum_{\ell=1}^{L} \bar{\boldsymbol{x}}_\ell^{\mathrm{H}} \bar{\boldsymbol{S}} \boldsymbol{D}_\ell \widehat{\boldsymbol{b}}_\ell =$$

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \sum_{\ell=1}^{L} \bar{\boldsymbol{s}}^{\mathrm{T}} \bar{\boldsymbol{X}}_\ell^* \boldsymbol{D}_\ell \left(\boldsymbol{D}_\ell^{\mathrm{H}} \boldsymbol{P}_{\bar{\boldsymbol{s}}} \boldsymbol{D}_\ell\right)^{-1} \boldsymbol{D}_\ell^{\mathrm{H}} \bar{\boldsymbol{X}}_\ell \bar{\boldsymbol{s}}^* = \tag{41}$$

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \left(\sum_{\ell=1}^{L} \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{P}_{\bar{\boldsymbol{s}}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}}\right) \bar{\boldsymbol{s}}, \tag{42}$$

where from (41) to (42) we have used that $\boldsymbol{P}_{\bar{\boldsymbol{s}}} \in \mathbb{R}_+^{N \times N}$, and that (41) is real-valued (and nonnegative).

By assumption, $\boldsymbol{\mathcal{E}} = \boldsymbol{O}$, hence $\boldsymbol{P}_{\bar{\boldsymbol{s}}} = P_s \cdot \boldsymbol{I}_N$ from (20). Thus, in this case (42) simplifies further to

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \left(\sum_{\ell=1}^{L} \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}}\right) \bar{\boldsymbol{s}} =$$

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \boldsymbol{Q}(\boldsymbol{p}) \bar{\boldsymbol{s}} = \lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{p})\right),$$

where $\boldsymbol{Q}(\boldsymbol{p})$ is defined in (21). Therefore, we conclude that when $\boldsymbol{\mathcal{E}} = \boldsymbol{O}$, the SBL position estimate is given by

$$\widehat{\boldsymbol{p}}_{\mathrm{SBL}} = \underset{\boldsymbol{p} \in \mathbb{R}^{3 \times 1}}{\arg\max} \lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{p})\right).$$

∎

APPENDIX B
PROOF OF PROPOSITION 2

Proof. Observe that in the proof of Proposition A, (42) holds for the general case, where $\boldsymbol{\mathcal{E}}$ is not necessarily equal to $\boldsymbol{O}$. Therefore, starting from (42), and focusing on the inverse matrix of a single matrix element in the sum, we now have

$$\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{P}_{\bar{\boldsymbol{s}}} \boldsymbol{D}_\ell^*\right)^{-1} = \frac{1}{P_s} \left(\boldsymbol{D}_\ell^{\mathrm{T}} \left(\boldsymbol{I}_N + \boldsymbol{\mathcal{E}}\right) \boldsymbol{D}_\ell^*\right)^{-1} =$$

$$\frac{1}{P_s} \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^* + \boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{\mathcal{E}} \boldsymbol{D}_\ell^*\right)^{-1} =$$

$$\frac{1}{P_s} \left[\left(\boldsymbol{I}_3 + \boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{\mathcal{E}} \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1}\right) \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)\right]^{-1} =$$

$$\frac{1}{P_s} \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\boldsymbol{I}_3 + \boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{\mathcal{E}} \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1}\right)^{-1}.$$

Hence, using the Neumann series [48], we have[16]

$$\left(\boldsymbol{I}_3 + \boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{\mathcal{E}} \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1}\right)^{-1} = \boldsymbol{I}_3 + \mathcal{O}(\boldsymbol{\mathcal{E}}) \Longrightarrow$$

$$\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{P}_{\bar{\boldsymbol{s}}} \boldsymbol{D}_\ell^*\right)^{-1} = \frac{1}{P_s} \left[\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} + \mathcal{O}(\boldsymbol{\mathcal{E}})\right]. \tag{43}$$

---

[16]By denoting $\boldsymbol{\Phi} = \mathcal{O}(\boldsymbol{\mathcal{E}})$, we mean that $|\lambda_{\max}(\boldsymbol{\Phi})| = \mathcal{O}(\varepsilon_{\max})$, where $\varepsilon_{\max} = |\lambda_{\max}(\boldsymbol{\mathcal{E}})|$. Therefore, $\boldsymbol{\Phi} \to \boldsymbol{O}$ when $\varepsilon_{\max} \to 0$.

As expected, the last term in (43) indicates that this approximation holds when the deviations from a constant spectral level, quantified here by $\boldsymbol{\mathcal{E}}$, are sufficiently small with respect to the normalized average power (see (20)).

Proceeding, by substituting (43) into (42), and using well-known eigenvalue perturbation theory results [49], we obtain

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \left(\sum_{\ell=1}^{L} \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{P}_{\bar{\boldsymbol{s}}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}}\right) \bar{\boldsymbol{s}} =$$

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \left(\sum_{\ell=1}^{L} \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left[\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} + \mathcal{O}(\boldsymbol{\mathcal{E}})\right] \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}}\right) \bar{\boldsymbol{s}} =$$

$$\max_{\bar{\boldsymbol{s}} \in \mathcal{S}_N} \bar{\boldsymbol{s}}^{\mathrm{H}} \left[\boldsymbol{Q}(\boldsymbol{p}) + \mathcal{O}(\boldsymbol{\mathcal{E}})\right] \bar{\boldsymbol{s}} = \lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{p})\right) + \mathcal{O}(\varepsilon_{\max}),$$

where we recall that $\varepsilon_{\max} = |\lambda_{\max}(\boldsymbol{\mathcal{E}})|$ (see (20)). It follows that

$$\widehat{\boldsymbol{p}}_{\mathrm{SBL}} = \underset{\boldsymbol{p} \in \mathbb{R}^{3 \times 1}}{\arg\max} \lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{p})\right) + \mathcal{O}(\varepsilon_{\max}).$$

∎

APPENDIX C
PROOF OF PROPOSITION 3

Proof. A key observation is that $\boldsymbol{Q}(\boldsymbol{p})$ is low-rank. Indeed, by definition, $\boldsymbol{Q}(\boldsymbol{p})$ is a sum of the following L matrices,

$$\boldsymbol{Q}_\ell(\boldsymbol{p}) \triangleq \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}} \in \mathbb{C}^{N \times N}, \tag{44}$$

where each is low-rank. Specifically, recall that $\boldsymbol{D}^{(\ell)} \in \mathbb{C}^{N \times 3}$, hence

$$\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \in \mathbb{C}^{3 \times 3} \implies \mathrm{rank}\left(\left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1}\right) = 3, \tag{45}$$

where we recall that $\mathrm{rank}(\boldsymbol{D}_\ell) = 3$ by assumption. In turn, this implies that

$$\mathrm{rank}\left(\boldsymbol{Q}_\ell(\boldsymbol{p})\right) = \mathrm{rank}\left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \left(\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^*\right)^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^*\right)^{\mathrm{H}}\right) = 3$$

$$\implies \mathrm{rank}\left(\boldsymbol{Q}(\boldsymbol{p})\right) = 3L,$$

assuming $\{\boldsymbol{Q}_\ell(\boldsymbol{p})\}$ are linearly independent.[17] Thus, we conclude that $\boldsymbol{Q}(\boldsymbol{p})$ has only $3L$ nonzero eigenvalues. Since typically $L \ll N$, we have established that $\boldsymbol{Q}(\boldsymbol{p})$ is low-rank.

Next, observe that $\boldsymbol{Q}(\boldsymbol{p})$ is a sum of L positive semi-definite matrices, and is therefore a positive semi-definite matrix as well. Due to its special structure (21), it is possible to compute a different matrix, $\widetilde{\boldsymbol{Q}}(\boldsymbol{p})$, with *exactly* the same eigenvalues as those of $\boldsymbol{Q}(\boldsymbol{p})$. For this, define the Cholesky decompositions [50]

$$\boldsymbol{D}_\ell^{\mathrm{T}} \boldsymbol{D}_\ell^* \triangleq \boldsymbol{\Gamma}_\ell^{\mathrm{H}} \boldsymbol{\Gamma}_\ell \in \mathbb{C}^{3 \times 3}, \ \forall \ell \in \{1, \dots, L\}, \tag{46}$$

where $\boldsymbol{\Gamma}_\ell \in \mathbb{C}^{3 \times 3}$. With these L 3-dimensional square matrices, substituting (46) into (21), we may now write

$$\boldsymbol{Q}(\boldsymbol{p}) = \sum_{\ell=1}^{L} \bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \boldsymbol{\Gamma}_\ell^{-1} \left(\bar{\boldsymbol{X}}_\ell \boldsymbol{D}_\ell^* \boldsymbol{\Gamma}_\ell^{-1}\right)^{\mathrm{H}} \in \mathbb{C}^{N \times N},$$

where we emphasize that $\det(\boldsymbol{\Gamma}_\ell) \neq 0$ is guaranteed for all $\ell \in \{1, \dots, L\}$ due to (45). Now, define (as in (18))

$$\boldsymbol{U}(\boldsymbol{p}) \triangleq \left[\bar{\boldsymbol{X}}_1 \boldsymbol{D}_1^* \boldsymbol{\Gamma}_1^{-1} \cdots \ \bar{\boldsymbol{X}}_L \boldsymbol{D}_L^* \boldsymbol{\Gamma}_L^{-1}\right] \in \mathbb{C}^{N \times 3L},$$

---

[17]This holds with probability one, due to the randomness in $\{\bar{\boldsymbol{X}}_\ell\}$.

with which

$$\boldsymbol{Q}(\boldsymbol{p}) = \boldsymbol{U}(\boldsymbol{p})\boldsymbol{U}(\boldsymbol{p})^{\mathrm{H}}.$$

However, we have that

$$
\begin{aligned}
\Lambda_+ \left(\boldsymbol{Q}(\boldsymbol{p})\right) &= \Lambda_+ \left(\boldsymbol{U}(\boldsymbol{p})\boldsymbol{U}(\boldsymbol{p})^{\mathrm{H}}\right) \\
&= \Lambda_+ \left(\boldsymbol{U}(\boldsymbol{p})^{\mathrm{H}}\boldsymbol{U}(\boldsymbol{p})\right) = \Lambda_+ \left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right),
\end{aligned}
$$

where $\Lambda_+ \left(\boldsymbol{C}\right)$ denotes the set of the nonzero eigenvalues of the semi-positive definite matrix $\boldsymbol{C}$, and $\widetilde{\boldsymbol{Q}}(\boldsymbol{p}) \in \mathbb{C}^{3L \times 3L}$. Put simply, $\widetilde{\boldsymbol{Q}}(\boldsymbol{p})$ has the same spectrum as $\boldsymbol{Q}(\boldsymbol{p})$. In particular,

$$\lambda_{\max}\left(\boldsymbol{Q}(\boldsymbol{p})\right) = \lambda_{\max}\left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right).$$

Since $\dim\left(\widetilde{\boldsymbol{Q}}(\boldsymbol{p})\right) = 3L < N = \dim\left(\boldsymbol{Q}(\boldsymbol{p})\right)$, we have reduced the computational burden, which is now governed by L, rather than N. Specifically, the complexity is $\mathcal{O}(NL^3)$, due to the required Cholesky decompositions (46) applied to L 3-dimensional matrices [50], and the subsequent application of the power method to the $3L$-dimensional matrix $\widetilde{\boldsymbol{Q}}(\boldsymbol{p})$. ∎

References

[1] A. Bahr, J. J. Leonard, and M. F. Fallon, "Cooperative localization for autonomous underwater vehicles," *Int. J. Robotics Res.*, vol. 28, no. 6, pp. 714–728, 2009.

[2] P. Corke, C. Detweiler, M. Dunbabin, M. Hamilton, D. Rus, and I. Vasilescu, "Experiments with underwater robot localization and tracking," in *Proc. IEEE Int. Conf. Robotics, Automation*, 2007, pp. 4556–4561.

[3] J. Waterston, J. Rhea, S. Peterson, L. Bolick, J. Ayers, and J. Ellen, "Ocean of things: Affordable maritime sensors with scalable analysis," in *Prof. OCEANS Conf.*, Marseille, France, 2019, pp. 1–6.

[4] H. P. Bucker, "Use of calculated sound fields and matched-field detection to locate sound sources in shallow water," *J. Acoust. Soc. Am.*, vol. 59, no. 2, pp. 368–373, 1976.

[5] H.-P. Tan, R. Diamant, W. K. G. Seah, and M. Waldmeyer, "A survey of techniques and challenges in underwater localization," *Ocean Engineering*, vol. 38, no. 14-15, pp. 1663–1676, 2011.

[6] V. Chandrasekhar, W. K. G. Seah, Y. S. Choo, and H. V. Ee, "Localization in underwater sensor networks: survey and challenges," in *Proc. ACM Int. Workshop Underwater Networks*, 2006, pp. 33–40.

[7] P. C. Etter, *Underwater acoustic modeling: principles, techniques and applications*. CRC Press, 1995.

[8] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*. Springer Science & Business Media, 2011.

[9] M. Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *J. Acoust. Soc. Am.*, vol. 122, no. 5, pp. 2580–2586, 2007.

[10] X. Cheng, H. Shu, Q. Liang, and D. H.-C. Du, "Silent positioning in underwater acoustic sensor networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1756–1766, 2008.

[11] B. Kouzoundjian, F. Beaubois, S. Reboul, J. B. Choquel, and J.-C. Noyer, "A TDOA underwater localization approach for shallow water environment," in *Proc. IEEE OCEANS Conf.*, Aberdeen, Scotland, 2017, pp. 1–4.

[12] G. Tuna and V. C. Gungor, "A survey on deployment techniques, localization algorithms, and research challenges for underwater acoustic sensor networks," *Int. J. Commun. Syst.*, vol. 30, no. 17, 2017.

[13] M. Stojanovic, "Underwater acoustic communication," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–12, 1999.

[14] R. Diamant, H.-P. Tan, and L. Lampe, "LOS and NLOS classification for underwater acoustic localization," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 311–323, 2012.

[15] L. Emokpae and M. Younis, "Surface based anchor-free localization algorithm for underwater sensor networks," in *Proc. IEEE Int. Conf. Commun. (ICC-2011)*, 2011, pp. 1–5.

[16] L. E. Emokpae, S. DiBenedetto, B. Potteiger, and M. Younis, "UREAL: Underwater reflection-enabled acoustic-based localization," *IEEE Sensors J.*, vol. 14, no. 11, pp. 3915–3925, 2014.

[17] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE J. Ocean. Eng.*, vol. 18, no. 4, pp. 401–424, 1993.

[18] M. D. Collins and W. A. Kuperman, "Focalization: Environmental focusing and source localization," *J. Acoust. Soc. Am.*, vol. 90, no. 3, pp. 1410–1422, 1991.

[19] E. A. Iscar Ruland, A. Shree, N. Goumas, and M. Johnson-Roberson, "Low cost underwater acoustic localization," in *Proc. ASA Mtgs. Acoust.*, vol. 30, no. 1, 2017.

[20] B. Gerondeau, L. Galeota, A. Caudwell, R. Gouge, A. Martin, R. Séguin, and R. Zitouni, "Low-cost underwater localization system," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*. IEEE, 2020, pp. 1153–1158.

[21] W. Mantzel, J. Romberg, and K. Sabra, "Compressive matched-field processing," *J. Acoust. Soc. Am.*, vol. 132, no. 1, pp. 90–102, 2012.

[22] K. L. Gemba, W. S. Hodgkiss, and P. Gerstoft, "Adaptive and compressive matched field processing," *J. Acoust. Soc. Am.*, vol. 141, no. 1, pp. 92–103, 2017.

[23] E. Dubrovinskaya, P. Casari, and R. Diamant, "Bathymetry-aided underwater acoustic localization using a single passive receiver," *J. Acoust. Soc. Am.*, vol. 146, no. 6, pp. 4774–4789, 2019.

[24] A. A. Saucan and M. Z. Win, "Information-seeking sensor selection for ocean-of-things," *IEEE Internet of Things J.*, vol. 7, no. 10, pp. 10 072–10 088, 2020.

[25] A. J. Weiss, "Direct position determination of narrowband radio frequency transmitters," *IEEE Signal Process. Lett.*, vol. 11, no. 5, pp. 513–516, 2004.

[26] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, 1976.

[27] M. S. Brandstein and H. F. Silverman, "A robust method for speech signal time-delay estimation in reverberant rooms," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP-1997)*, vol. 1, 1997, pp. 375–378.

[28] F. Grondin and J. Glass, "A study of the complexity and accuracy of direction of arrival estimation methods based on GCC-PHAT for a pair of close microphones," *arXiv preprint arXiv:1811.11787*, 2018.

[29] G. Wang, W. Zhu, and N. Ansari, "Robust TDOA-based localization for IoT via joint source position and NLOS error estimation," *IEEE Internet of Things J.*, vol. 6, no. 5, pp. 8529–8541, 2019.

[30] Y. Zou and H. Liu, "TDOA localization with unknown signal propagation speed and sensor position errors," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1024–1027, 2020.

[31] W. Xiong, C. Schindelhauer, H. C. So, J. Bordoy, A. Gabbrielli, and J. Liang, "TDOA-based localization with NLOS mitigation via robust model transformation and neurodynamic optimization," *Signal Processing*, vol. 178, p. 107774, 2021.

[32] W. S. Hodgkiss and J. C. Preisig, "Kauai ACOMMS MURI 2011 (KAM11) experiment," in *Proc. Euro. Conf. Underwater Acoust. (ECUA-2012)*, 2012, pp. 993–1000.

[33] P. C. Etter, *Underwater acoustic modeling and simulation*. CRC press, 2018.

[34] J. M. Hovem, *Ray trace modeling of underwater sound propagation. Documentation and use of the PlaneRay model*, 2011.

[35] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 2009.

[36] R. Aubauer, M. O. Lammers, and W. W. L. Au, "One-hydrophone method of estimating distance and depth of phonating dolphins in shallow water," *J. Acoust. Soc. Am.*, vol. 107, no. 5, pp. 2744–2749, 2000.

[37] F. Schulz, R. Weber, A. Waldhorst, and J. Bohme, "Performance enhancement of blind adaptive equalizers using environmental knowledge," in *Proc. OCEANS Conf.*, vol. 4, 2003, pp. 1793–1799.

[38] L. Emokpae and M. Younis, "Surface-reflection-based communication and localization in underwater sensor networks," *ACM Trans. Sensor Networks*, vol. 10, no. 3, pp. 1–51, 2014.

[39] A. Weiss, "Blind direction-of-arrival estimation in acoustic vector-sensor arrays via tensor decomposition and Kullback-Leibler divergence covariance fitting," *IEEE Trans. Signal Process.*, vol. 69, pp. 531–545, 2021.

[40] A. R. Conn, N. I. Gould, and P. L. Toint, *Trust Region Methods*. SIAM, 2000.

[41] A. B. Baggeroer, "The stochastic Cramér-Rao bound for source localization and medium tomography using vector sensors," *J. Acoust. Soc. Am.*, vol. 141, no. 5, pp. 3430–3449, 2017.

[42] A. B. Baggeroer, W. Kuperman, and H. Schmidt, "Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem," *J. Acoust. Soc. Am.*, vol. 83, no. 2, pp. 571–587, 1988.

[43] S. L. Collier, "Fisher information for a complex Gaussian random variable: Beamforming applications for wave propagation in a random medium," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4236–4248, 2005.

[44] W. A. P. van Kleunen, K. C. H. Blom, N. Meratnia, A. B. J. Kokkeler, P. J. M. Havinga, and G. J. M. Smit, "Underwater localization by combining time-of-flight and direction-of-arrival," in *Proc. OCEANS Conf.*, Taipei, Taiwan, 2014, pp. 1–6.

[45] M. Siderius, D. R. Jackson, D. Rouseff, and R. Porter, "Multipath compensation in shallow water environments using a virtual receiver," *J. Acoust. Soc. Am.*, vol. 102, no. 6, pp. 3439–3449, 1997.

[46] G. Shulkind, S. Jegelka, and G. W. Wornell, "Sensor array design through submodular optimization," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 664–675, 2018.

[47] A. Weiss and G. W. Wornell, "One-bit direct position determination of narrowband Gaussian signals," in *Proc. IEEE Statistical Signal Processing Workshop (SSP-2021)*, 2021, pp. 466–470.

[48] G. W. Stewart, *Matrix Algorithms: Volume 1: Basic Decompositions*. SIAM, 1998.

[49] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*. SIAM, 1997, vol. 50.

[50] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 2013.