Underwater Acoustic Source Localization by Vector Sensor Array Using Compressive Sampling

P. V. Nagesha*, G. V. Anand*, Sanjeev Gurugopinath*, and Akarsh Prabhakar†
*Department of ECE, PES University, Bengaluru 560085, India
†Department of ECE, National Institute of Technology Surathkal, Mangaluru 575025, India
Emails: {nagesh.uwa, gargeshwari.anand, sanju83, akarsh.9996}@gmail.com

***Abstract*—A MUSIC-ML (MML) processing algorithm for 3-dimensional localization of underwater acoustic sources using a 2-dimensional acoustic vector (AVS) array was proposed recently. The MML processor performs source localization with high accuracy and resolution. But the hardware and computational complexity of the MML processor is very high. In this paper we propose the use of compressive sampling of the AVS array output in the spatial domain for reducing the processor complexity. It is shown that a significant reduction in complexity with negligible performance degradation can be achieved by using compressive sampling. But complexity reduction by reducing the number of sensors without compressive sampling leads to a significant performance degradation.**

***Index Terms*—Acoustic vector sensor, complexity reduction, compressive sampling, MUSIC-ML processor, underwater acoustic source localization.**

I. Introduction

Methods of localization of underwater acoustic sources include the maximum likelihood (ML) technique [1], matched field processing techniques such as the Bartlett and Capon processors [2], or subspace techniques such as MUSIC [3]. Three-dimensional (3-D) localization requires the deployment of a two-dimensional array (2DA) of hydrophones, composed of a vertical linear array (VLA) for range-depth estimation and a horizontal linear array (HLA) for bearing estimation [4]. The ML technique is asymptotically efficient; but it is highly computation-intensive, particularly in a non-Gaussian noise environment that is frequently encountered in the ocean. The MUSIC algorithm has lower complexity, but its performance degrades rapidly as the signal-to-noise ratio (SNR) is reduced. A new method called MUSIC-ML (MML), involving a combination of MUSIC and ML, was proposed recently [5]. It was shown that the MML processor has a better low-SNR performance than MUSIC and lower complexity than ML. It was also shown that a significant improvement in the performance of the MML algorithm can be achieved if an acoustic vector sensor (AVS) array is used to sample the acoustic field. An AVS has four collocated sensors which measure the acoustic pressure p and three orthogonal components of particle velocity (ν_x, ν_y, ν_z) at the same point in space. Hence, an AVS is more informative than the conventional hydrophone which measures the acoustic pressure only. Consequently, an AVS array provides better localization performance than a hydrophone array with the same number of elements. But replacement of a hydrophone array by an AVS array entails a major increase in complexity.

An N-element AVS array has $4N$ outputs. Each output is connected to a separate receiver chain of frontend circuits that perform the operations of amplification, down-conversion, filtering, and analog-to-digital conversion on the received analog signal and provides a digital output for further signal processing. An HLA with a large number of elements N is often used for high resolution bearing estimation which is a critical requirement in many ocean acoustic applications. The hardware complexity of the source localization processor arises due to the presence of a large number of front end circuit chains, and the high software complexity arises due to the estimation algorithms involving manipulations of large-size data matrices. It is known that compressive sampling (CS) techniques can be used in array processing applications to reduce both types of complexity while maintaining high direction estimation accuracy and resolution [6]. Direction estimation may be done either through CS recovery or by applying an existing estimation technique such as MUSIC to the compressed measurement vector. In this paper we present a CS based MML algorithm for 3-D source localization in the ocean using an AVS array. It is shown that the performance of the proposed compressive MML (CMML) processor is almost on par with the existing MML processor if the signal-to-noise ratio (SNR) exceeds a certain threshold; but reduction in the number of sensors without compressive sampling leads to a significant performance degradation even at high SNR.

II. Data Models

A. AVS-VLA

We shall model the ocean as a range-independent channel composed of a water layer of constant depth h, density ρ, and sound speed c, overlying a fluid bottom of density ρ_b, sound speed c_b, and absorption coefficient ϵ. This model is chosen for the sake of computational simplicity; the analysis presented in this paper can be readily extended to a horizontally stratified ocean.

Consider an AVS-VLA of N sensors with uniform inter-sensor spacing d. The nth element of the VLA is located at $z_{vn} = z_{v1} + (n-1)d$, $n = 1, \cdots, N$. Let J uncorrelated sources be located at $\boldsymbol{\xi} = \{\boldsymbol{x}_j = (\boldsymbol{u}_j, \theta_j) = (r_j, z_j, \theta_j)\}$, $j = 1, \cdots, J$, in the far-field region of the 2DA, with ranges r_j, depths z_j, and bearings θ_j, radiate narrowband signals $\eta_j(t)$ of center frequency f_0 with means zero and variances σ_j^2,

978-1-5090-1537-5/16/$31.00 ©2016 IEEE

$j = 1, \cdots, J$. We shall not consider the vertical component of particle velocity v_z measured by an AVS since the inclusion of this measurement is found to increase the complexity of the array processor without yielding any significant additional improvement in performance. The output of the VLA at discrete time t is therefore denoted by the $3N$ dimensional data vector $\boldsymbol{y}_v(t) = \boldsymbol{s}_v(t) + \boldsymbol{w}_v(t)$, where $\boldsymbol{s}_v(t)$ is the signal vector, $\boldsymbol{w}_v(t)$ is the noise vector, and the subscript v denotes VLA. The VLA data vector can be written as

$$\boldsymbol{y}_v(t) = [\boldsymbol{a}_v(\boldsymbol{x}_1) \cdots \boldsymbol{a}_v(\boldsymbol{x}_J)]\boldsymbol{\eta}(t) + \boldsymbol{w}_v(t),\ t = 1, \cdots, L, \quad (1)$$

where

$$\boldsymbol{a}_v(\boldsymbol{x}_j) = [\boldsymbol{d}_{v1}(\boldsymbol{x}_j)^T \ \cdots \ \boldsymbol{d}_{vN}(\boldsymbol{x}_j)^T]^T,\ j = 1, \cdots, J, \quad (2)$$

$$\boldsymbol{d}_{vn}(\boldsymbol{x}_j) = [p_{vn}(\boldsymbol{x}_j)\ \sqrt{2}\rho c v_{vxn}(\boldsymbol{x}_j)\ \sqrt{2}\rho c v_{vyn}(\boldsymbol{x}_j)]^T, \quad (3)$$
$$n = 1, \cdots, N,$$

$$\boldsymbol{\eta}(t) = [\eta_1(t) \ \cdots \ \eta_J(t)], \quad (4)$$

with $\boldsymbol{a}_v(\boldsymbol{x}_j)$ being the steering vector in the direction of $\boldsymbol{x}_j$, and $p_{vn}(\boldsymbol{x}_j)$, $v_{vxn}(\boldsymbol{x}_j)$, $v_{vyn}(\boldsymbol{x}_j)$ denoting the complex amplitudes of the acoustic pressure and horizontal (x, y) components of particle velocity at the nth sensor of the VLA due to a unit-strength source at $\boldsymbol{x}_j$. Using the normal mode theory of sound propagation in the ocean, we can write [7]

$$\boldsymbol{a}_v(\boldsymbol{x}_j) = \sum_{m=1}^{M} b_m(\boldsymbol{u}_j)\boldsymbol{c}_{v,m} \otimes \boldsymbol{v}_m(\theta_j), \quad (5)$$

$$b_m(\boldsymbol{u}_j) = \frac{\psi_m(r_j)}{\sqrt{\kappa_m r_j}} \exp\left(i\kappa_m r_j - \alpha_m r_j\right), \quad (6)$$

$$\boldsymbol{c}_{v,m} = [\psi_m(z_{v1}) \ \cdots \ \psi_m(z_{vN})]^T, \quad (7)$$

$$\boldsymbol{v}_m(\theta_j) = [1 \ \ \sqrt{2}\frac{\kappa_m}{\kappa_0}\cos(\theta_j) \ \ \sqrt{2}\frac{\kappa_m}{\kappa_0}\sin(\theta_j)]^T, \quad (8)$$

where $\psi_m(z)$, κ_m and α_m respectively are the mode functions, wavenumbers and attenuation coefficients of the mth normal mode, $\kappa_0 = 2\pi f_0/c$,

$$M = \left\lfloor \frac{\kappa_0 h}{\pi}\sqrt{1 - \left(\frac{c}{c_b}\right)^2} + \frac{1}{2} \right\rfloor, \quad (9)$$

and $\otimes$ denotes the Kronecker product. It is assumed that the noise vectors $\{\boldsymbol{w}_v(t), t = 1, \cdots, L\}$ are mutually independent circularly symmetric complex Gaussian with mean zero and covariance matrix $\sigma^2_{\boldsymbol{w}}\boldsymbol{I}_{3N}$, and also independent of $\{\boldsymbol{\eta}(t), t = 1, \cdots, L\}$.

B. AVS-HLA

Consider an AVS-HLA of N sensors with uniform inter-sensor spacing d, lying parallel to the x-axis at depth z_{h0}. The first element of the HLA is the same as one of the elements of the VLA, and the common element is the reference for range measurement. Using the subscript h to denote all HLA related quantities, we can write

$$\boldsymbol{y}_h(t) = [\boldsymbol{a}_h(\boldsymbol{x}_1) \ \cdots \ \boldsymbol{a}_h(\boldsymbol{x}_J)]\boldsymbol{\eta}(t) + \boldsymbol{w}_h(t), \quad (10)$$

$$\boldsymbol{a}_h(\boldsymbol{x}_j) = \sum_{m=1}^{M} b_m(\boldsymbol{u}_j)\boldsymbol{c}_{h,m}(\theta_j) \otimes \boldsymbol{v}_m(\theta_j), \quad (11)$$

$$\boldsymbol{c}_{h,m}(\theta_j) = \psi_m(z_{h0})[1 \ e^{ig_{mj}} \ \cdots \ e^{i(N-1)g_{mj}}]^T, \quad (12)$$

$$g_{mj} = \kappa_m \cos(\theta_j) \quad (13)$$

It is assumed that the noise vectors $\{\boldsymbol{w}_h(t), t = 1, \cdots, L\}$ are mutually independent circularly symmetric complex Gaussian with mean zero and covariance matrix $\sigma^2_{\boldsymbol{w}}\boldsymbol{I}_{3N}$, and also independent of $\{\boldsymbol{\eta}(t), t = 1, \cdots, L\}$.

III. Review of MML Method

Three dimensional localization of multiple sources by the MML method involves three stages of processing as described below.

A. Generation of Range-Depth Candidate Pool

Consider the problem of estimating the source ranges and depths $\{\boldsymbol{u}_j = (r_j, z_j), j = 1, \cdots, J\}$ by MUSIC. Computation of the 2-D MUSIC spectrum requires the knowledge of the array steering vector at multiple locations on the range-depth grid. For a source at location $\boldsymbol{x} = (\boldsymbol{u}, \theta)$, the steering vector $\boldsymbol{a}_v(\boldsymbol{x}) = \boldsymbol{a}_v(\boldsymbol{u}, \theta)$, depends not only on range-depth $\boldsymbol{u}$, but also on bearing θ. For range-depth estimation without prior knowledge of the source bearings, we employ the following strategy. Split the $3N$-dimensional data vector $\boldsymbol{y}_v(t)$ defined in (1)-(8) into three N-dimensional sub-vectors $\boldsymbol{y}_{v1}$, $\boldsymbol{y}_{v2}$, and $\boldsymbol{y}_{v3}$ corresponding to the $p-$, v_x- and v_y- channel measurements. The sub-vectors are given by

$$\boldsymbol{y}_{vi}(t) = [\boldsymbol{a}_v(\boldsymbol{x}_1) \cdots \ \boldsymbol{a}_v(\boldsymbol{x}_J)]\boldsymbol{\eta}(t) + \boldsymbol{w}_{vi}(t),\ i = 1, 2, 3, \quad (14)$$

$$\boldsymbol{a}_{v1}(\boldsymbol{x}_j) = \widehat{\boldsymbol{a}}_{v1}(\boldsymbol{u}_j) = \sum_{m=1}^{M} b_m(\boldsymbol{u}_j)\boldsymbol{c}_{v,m} \quad (15)$$

$$\boldsymbol{a}_{v2}(\boldsymbol{x}_j) = \sum_{m=1}^{M} \left(\frac{\kappa_m}{\kappa_0}\right) b_m(\boldsymbol{u}_j)\boldsymbol{c}_{v,m}\cos(\theta_j) \quad (16)$$

$$\boldsymbol{a}_{v3}(\boldsymbol{x}_j) = \sum_{m=1}^{M} \left(\frac{\kappa_m}{\kappa_0}\right) b_m(\boldsymbol{u}_j)\boldsymbol{c}_{v,m}\sin(\theta_j) \quad (17)$$

and $\{\boldsymbol{w}_{vi}(t), i = 1, 2, 3\}$ are N-dimensional sub-vectors of $\boldsymbol{w}_v(t)$. Let

$$\widehat{\boldsymbol{R}}_{vi} = \frac{1}{L}\sum_{t=1}^{L} \boldsymbol{y}_{vi}(t)\boldsymbol{y}^H_{vi}(t) \quad (18)$$

denote the sample correlation matrices of the data sub-vectors $\boldsymbol{y}_{vi}(t)$. Since the steering vectors $\{\boldsymbol{a}_{vi}(\boldsymbol{x}), i = 1, 2, 3\}$ are highly correlated, we can use the approximation

$$\boldsymbol{a}_{v3}(\boldsymbol{x}) \approx \boldsymbol{a}_{v2}(\boldsymbol{x}) \approx \boldsymbol{a}_{v1}(\boldsymbol{x}) = \widehat{\boldsymbol{a}}_{v1}(\boldsymbol{u}) \equiv \widehat{\boldsymbol{a}}_{v2}(\boldsymbol{u}) \equiv \widehat{\boldsymbol{a}}_{v3}(\boldsymbol{u}) \quad (19)$$

Thus, we can construct the following 2D-MUSIC spectrum from the data sub-vectors:

$$B_v(\boldsymbol{u}) = \sum_{i=1}^{3} \frac{1}{\widehat{\boldsymbol{a}}^H_{v1}(\boldsymbol{u})\boldsymbol{E}_{Nvi}\boldsymbol{E}^H_{vi}\widehat{\boldsymbol{a}}_{v1}(\boldsymbol{u})} \quad (20)$$

where $\boldsymbol{E}_{Nvi}$ is the noise subspace matrix obtained from the eigendecomposition of $\widehat{\boldsymbol{R}}_{vi}$. The J tallest peaks of $B_v(\boldsymbol{u})$ are expected to provide estimates of $\{\boldsymbol{u}_j, j=1,\cdots,J\}$. In practice,large estimation errors may occur sometimes because one or more peaks corresponding to the true source positions are dwarfed by a large sidelobe. To avoid such errors, we form a candidate pool of range-depth estimate by selecting the arguments of the $J' > J$ tallest peaks of $B_v(\boldsymbol{u})$. The size of the pool should be large enough to ensure that all true range-depth estimates are included in the pool. We have performed extensive simulations to conclude that this condition is satisfied if $J' \approx 3J$. A method of obtaining source position estimates by weeding out the false peaks from the range-depth candidate pool is presented in the following sub-section.

B. Range-Depth Estimation

Let the locations of the $J'(\approx 3J)$ tallest peaks of $B_v(\boldsymbol{u})$ be denoted by $\{\widetilde{\boldsymbol{u}}_j, j=1,\cdots,J'\}$. For each of these J' range-depth pairs, we estimate the bearing by 1-D MUSIC. Thus we have

$$\widetilde{\theta}_{vj} = \arg\max_{\theta} \left\{ B_v^{(j)}(\theta|\widetilde{\boldsymbol{u}}_j) \right\},\ j=1,\cdots,J' \tag{21}$$

$$B_v^{(j)}(\theta|\widetilde{\boldsymbol{u}}_j) = \frac{1}{\boldsymbol{a}_v^H(\widetilde{\boldsymbol{u}}_j,\theta)\boldsymbol{E}_{Nv}\boldsymbol{E}_{Nv}^H\boldsymbol{a}_v(\widetilde{\boldsymbol{u}}_j,\theta)}. \tag{22}$$

Let $\widetilde{\boldsymbol{x}}_j \triangleq (\widetilde{\boldsymbol{u}}_j, \widetilde{\theta}_{vj})$. Each vector composed of J elements of the set $\{\widetilde{\boldsymbol{x}}_1,\cdots,\widetilde{\boldsymbol{x}}_{J'}\}$ is a candidate for the joint estimate of the location of J sources. The total number of candidates is $K=\binom{J'}{J}$. Let these candidates be denoted by

$$\boldsymbol{\xi}_i = [\widetilde{\boldsymbol{x}}_{i_1}\ \cdots\ \widetilde{\boldsymbol{x}}_{i_J}],\quad i=1,\cdots,K. \tag{23}$$

Let

$$\widetilde{\boldsymbol{A}}_{vi} = [\mathbf{a}_v(\widetilde{\boldsymbol{x}}_{i_1})\ \cdots\ \mathbf{a}_v(\widetilde{\boldsymbol{x}}_{i_J})] \tag{24}$$

be the steering-vector matrix for the ith candidate ξ_i. The matrix $\widetilde{\boldsymbol{A}}_{vi}(\widetilde{\boldsymbol{A}}_{vi}^H\widetilde{\boldsymbol{A}}_{vi})^{-1}\widetilde{\boldsymbol{A}}_{vi}^H$ is the orthogonal projection matrix onto the ith candidate signal subspace S_i = span $\{\mathbf{a}_v(\widetilde{\boldsymbol{x}}_{i_1})\ \cdots\ \mathbf{a}_v(\widetilde{\boldsymbol{x}}_{i_J})\}$. Consider the projection spectrum

$$P_v(\boldsymbol{\xi_i}) = \text{trace}(\widetilde{\boldsymbol{A}}_{vi}(\widetilde{\boldsymbol{A}}_{vi}^H\widetilde{\boldsymbol{A}}_{vi})^{-1}\widetilde{\boldsymbol{A}}_{vi}^H\widetilde{\boldsymbol{R}}_v),\ i{=}1,\cdots,K. \tag{25}$$

We note that each element of the set $\{\boldsymbol{\xi}_i : i=1,\cdots,K\}$ is a point in a $3J$-dimensional space. The joint ML estimate of the location of all sources is given by

$$\widehat{\boldsymbol{\xi}}_v = (\widehat{\boldsymbol{u}}_1, \widehat{\theta}_{v1},\cdots,\widehat{\boldsymbol{u}}_J,\widehat{\theta}_{vJ}) = \arg\max_i\ \ P_v(\boldsymbol{\xi}_i) \tag{26}$$

The procedure outlined above is rather computation-intensive since it involves multiple maximizations of the MUSIC spectrum in (22). In order to reduce the computational load, a coarse search grid of size 3 degrees is used for obtaining the bearing estimates $\widetilde{\theta}_{vj}$ in (21), since high bearing-estimation accuracy is not required for the joint range-depth estimates obtained from the maximization of (26). After obtaining the range-depth estimates $\{\widehat{\boldsymbol{u}}_j, j=1,\cdots,J\}$ using the method described above, more accurate bearing estimates are obtained using the HLA data as explained below.

C. Bearing Estimation

For obtaining improved bearing estimates, we use a modified version of the MML method used for range-depth estimation in Sections III-A and III-B. We use the HLA data $\{\mathbf{y}_h(t), t=1,\cdots,L\}$ to compute the 1-D MUSIC spectrum corresponding to the range-depth estimate $\widehat{\boldsymbol{u}_j}$ of each source. Thus we have

$$B_h^{(j)}(\theta){=}\frac{1}{\boldsymbol{a}_h^H(\widehat{\boldsymbol{u}}_j,\theta)\boldsymbol{E}_{Nh}\boldsymbol{E}_{Nh}^H\boldsymbol{a}_h(\widehat{\boldsymbol{u}}_j,\theta)},\ j{=}1,\cdots,J. \tag{27}$$

where $\boldsymbol{a}_h(\widehat{\boldsymbol{u}}_j,\theta)$ is the steering vector of the AVS HLA and $\boldsymbol{E}_{Nh}$ is the noise subspace matrix obtained from the eigen decomposition of the HLA data sample correlation matrix $\widehat{\mathbf{R}}_h$ defined as

$$\widehat{\mathbf{R}}_h = \frac{1}{L}\sum_{t=1}^{L}\mathbf{y}_h(t)\mathbf{y}_h^H(t). \tag{28}$$

It is found that, for each $j \in \{1,\cdots,J\}$, $B_h^{(j)}(\theta)$ has large peaks of comparable magnitudes close to all $\{\theta_j, j=1,\cdots,J\}$, and also some randomly located large false peaks. To overcome the resultant ambiguity, we define the spectrum

$$B_h(\theta) = \sum_{j=1}^{J}\frac{1}{\boldsymbol{a}_h^H(\widehat{\boldsymbol{u}}_j,\theta)\boldsymbol{E}_{Nh}\boldsymbol{E}_{Nh}^H\boldsymbol{a}_h(\widehat{\boldsymbol{u}}_j,\theta)}, \tag{29}$$

and consider the set of the tallest $J''(\approx 3J)$ peaks of $B_h(\theta)$ denoted by $\bar{\theta}_j, j=1,\cdots,J''$. The range-depth estimate $\widehat{\boldsymbol{u}}_j$ of each source can be paired with any of the $\bar{\theta}_j, j=1,\cdots,J''$. Therefore, the total number of candidates for the joint estimate of the location of sources is $\bar{K}=J''(J''-1)\cdots(J''-J+1)$, assuming that the sources have distinct bearings. Let these candidates be denoted by

$$\bar{\boldsymbol{\xi}}_i = [\widehat{\boldsymbol{u}}_1,\bar{\theta}_{i_1},\cdots,\widehat{\boldsymbol{u}}_J,\bar{\theta}_{i_J}],\quad i=1,\cdots,\bar{K} \tag{30}$$

Let

$$\bar{\boldsymbol{A}}_{hi} = [\boldsymbol{a}_h(\widehat{\boldsymbol{u}}_1,\bar{\theta}_{i_1})\ \cdots\ \boldsymbol{a}_h(\widehat{\boldsymbol{u}}_J,\bar{\theta}_{i_J})] \tag{31}$$

be the steering vector for $\bar{\boldsymbol{\xi}}_i$. The matrix $\bar{\boldsymbol{A}}_{hi}(\bar{\boldsymbol{A}}_{hi}^H\bar{\boldsymbol{A}}_{hi})^{-1}\bar{\boldsymbol{A}}_{hi}^H$ is the orthogonal projection matrix onto the ith candidate signal subspace $\bar{S}i$ = span$\{\boldsymbol{a}_h(\widehat{\boldsymbol{u}}_1,\bar{\theta}_{i_1})\ \cdots\ \boldsymbol{a}_h(\widehat{\boldsymbol{u}}_J,\bar{\theta}_{i_J})\}$. Consider the spectrum

$$P_h(\bar{\boldsymbol{\xi}}_i){=}\text{trace}(\bar{\boldsymbol{A}}_{hi}(\bar{\boldsymbol{A}}_{hi}^H\bar{\boldsymbol{A}}_{hi})^{-1}\bar{\boldsymbol{A}}_{hi}^H\widehat{\boldsymbol{R}}_h),\ i{=}1,\cdots,\bar{K}. \tag{32}$$

The joint ML estimate of the location of all sources is given by

$$\widehat{\boldsymbol{\xi}}_{2DA} = (\widehat{\boldsymbol{u}}_1,\widehat{\theta}_1,\cdots,\widehat{\boldsymbol{u}}_J,\widehat{\theta}_J) = \arg\max_i P_h(\bar{\boldsymbol{\xi}}_i). \tag{33}$$

The final bearing estimates are $\{\widehat{\theta}_j, j=1,\cdots,J\}$.

IV. Localization by Compressive Sampling

A. Compressive Sampling

The purpose of compressive sampling (CS) is to mitigate the complexity of the processor used for source localization. In the CS framework, the analog signal received by the $3N$ channels of an N-sensor AVS array is compressed in the analog domain and then passed through a fewer number of front-end circuit chains to obtain the digital baseband signal for further processing [6]. Hence, compressive sampling is equivalent to applying a compression matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times 3N}$, with $M < 3N$, to the $3N$-dimensional data vector $\boldsymbol{y}(t)$ to generate the compressed M-dimensional measurement vector $\boldsymbol{z}(t) = \mathbf{\Phi}\boldsymbol{y}(t)$. We can write

$$\boldsymbol{y}(t)=[\boldsymbol{a}(\bar{\boldsymbol{x}}_1) \;\cdots\; \boldsymbol{a}(\bar{\boldsymbol{x}}_K)]\boldsymbol{\zeta}(t)+\boldsymbol{w}(t)=\bar{\boldsymbol{A}}\boldsymbol{\zeta}(t)+\boldsymbol{w}(t), \quad (34)$$

for $t = 1, \cdots, L$, where $\boldsymbol{\zeta}(t) \in \mathbb{C}^K$ is a sparse signal vector with J non-zero elements, $K > 3N$ and $K \gg J$, where J is the number of sources located at $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_J\} \subseteq \{\bar{\boldsymbol{x}}_1, \cdots, \bar{\boldsymbol{x}}_K\}$. The support of $\boldsymbol{\zeta}(t)$ is the same for $t = 1, \cdots, L$. Thus, we have

$$\boldsymbol{z}(t)=\mathbf{\Phi}\bar{\boldsymbol{A}}\boldsymbol{\zeta}(t)+\mathbf{\Phi}\boldsymbol{w}(t)=\mathbf{\Psi}\boldsymbol{\zeta}(t)+\boldsymbol{v}(t),\; t=1,\cdots,L. \quad (35)$$

We have chosen the elements of $\mathbf{\Phi}$ to be independent samples of a complex circular Gaussian random variable, but other choices are possible according to the CS theory. The compressive sampling AVS array architecture is as shown in Figure 1.

The problem of localization may be solved by recovering the support of the sparse signal vector $\boldsymbol{\zeta}(t)$ from the compressed measurement vectors $\{\boldsymbol{z}(t), t = 1, \cdots, L\}$. It is known that exact recovery of $\boldsymbol{\zeta}(t)$ is possible if noise $\boldsymbol{v}(t) = \mathbf{0}$ and the matrix $\mathbf{\Psi} \in \mathbb{C}^{M\times K}$ satisfies the restricted isometry property (RIP) or the computationally verifiable mutual incoherence property (MIP [8]. The mutual coherence of $\mathbf{\Psi}$ is defined as

$$\mu = \max_{i\neq j} \left|\boldsymbol{\psi}_i^H \boldsymbol{\psi}_j\right|, \quad (36)$$

where $\boldsymbol{\psi}_i$ is the ith column of $\mathbf{\Psi}$. A sufficient condition for perfect recovery is $\mu < \frac{1}{2J-1}$. In general, K should be sufficiently small and/or M should be sufficiently large for satisfying the MIP condition. On the other hand, we need a large value of K for achieving high bearing resolution, and a small value of M for complexity mitigation. In the case of noisy measurements, the probability of perfect recovery reduces as the noise intensity is increased. In this paper, localization by CS recovery is not considered. We seek to solve the localization problem by applying the MML algorithm to the compressed measurement vectors $\{\boldsymbol{z}_v(t) = \mathbf{\Phi}_v\boldsymbol{y}_v(t) = \mathbf{\Phi}_h\boldsymbol{y}_h(t)\}$, derived from an AVS VLA of N_v sensors and an AVS HLA of N sensors. But the conditions for perfect recovery mentioned above play an important role in the choice of the parameters K and M.

B. Range-Depth Estimation

Before considering compressive sampling, we shall illustrate the range-depth estimation performance of the MML processor using an AVS VLA of N_v sensors. Consider a Pekeris channel

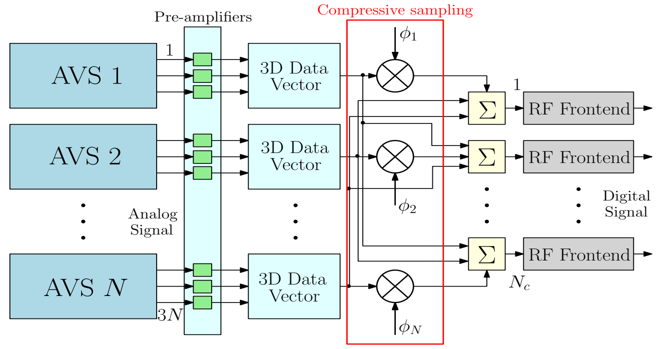

Fig. 1: Compressive sampling AVS array architecture

with the following parameters: depth $h = 100$ m, sound speed in water $c = 1500$ m/s, sound speed in bottom $c_b = 1700$ m/s, density ratio $\frac{\rho_b}{\rho} = 1.5$, and bottom attenuation coefficient $\epsilon = 0.2$ dB/λ_b (where λ_b is the acoustic wavelength in the ocean bottom). The signal frequency is $f_0 = 200$ Hz. The Kraken normal mode program [9] is used to compute the modal eigenfunctions, wavenumbers, and attenuation coefficients, $\psi_m(z)$, κ_m, and α_m. We consider two uniform VLAs, one with $N_v = 20$ sensors, and the other with $N_v = 9$ sensors. For both arrays, the depth of the topmost sensor is $z_{v1} = 5$ m and the array length is 76 m. We consider three sources whose coordinates are independent samples of uniformly distributed random variables, $r_j \sim \mathcal{U}(4\text{ km}, 7\text{ km})$, $z_j \sim \mathcal{U}(10\text{ m}, 100\text{ m})$, and $\theta_j \sim \mathcal{U}(0°, 360°)$, subject to the constraint that the separation between sources is not less than 40 m in range, 4 m in depth and 10 degrees in bearing. Signals received from different sources have the same signal-to-noise ratio (SNR) at the receiver array. Only the pressure components of the signal and noise at each AVS are considered for computing the SNR. The SNR (in dB) of the jth source is defined as

$$(\text{SNR})_j = 10\log_{10}\left(\frac{\sum_{n=1}^N |p_{nj}|^2}{N\sigma_{\mathbf{w}}^2}\right), \quad (37)$$

where p_{nj} is the acoustic pressure at the nth sensor due to the jth source. Simulation results comparing the performances of the two arrays, obtained from $L = 100$ snapshots, are shown in Figs. 2 and 3. These figures show plots of average root-mean-square errors (ARMSE) of range and depth estimates with respect to SNR. The ARMSE is defined as

$$\text{ARMSE(range)} = \frac{1}{J}\sum_{j=1}^{J}\sqrt{\sum_{s=1}^{S}\frac{1}{S}(\widehat{r}_j^{(s)} - r_j^{(s)})^2}, \quad (38)$$

where $r_j^{(s)}$ and $\widehat{r}_j^{(s)}$ are the true and estimated values of the range of the jth source in the sth Monte Carlo simulation. ARMSE(depth) and ARMSE(bearing) are defined in a similar fashion. The results in this figure and all subsequent figures are obtained from $S = 500$ Monte Carlo simulations. It is evident from Figs. 2 and 3 that pretty good range-depth estimates can be obtained even with a VLA with the relatively

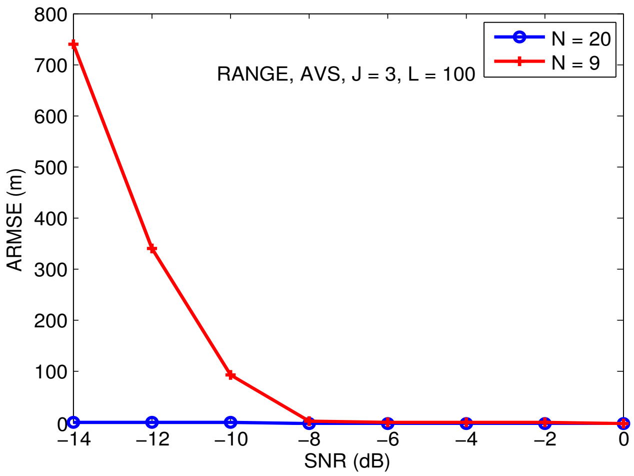

Fig. 2: Plots of ARMSE vs. SNR for range estimation by MML method, for three sources at random locations.

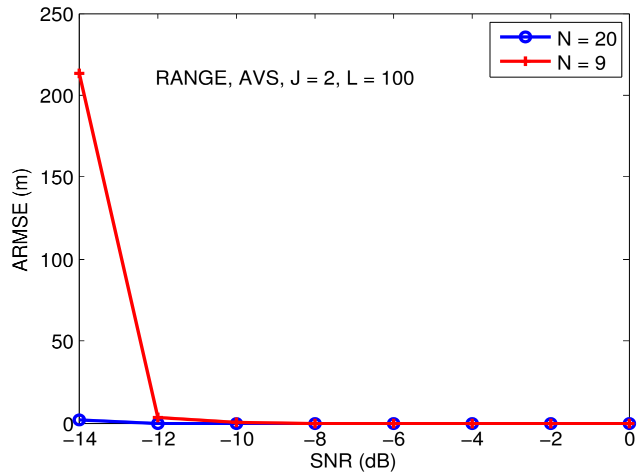

Fig. 4: Plots of ARMSE vs. SNR for range estimation by MML method, for two sources at fixed locations.

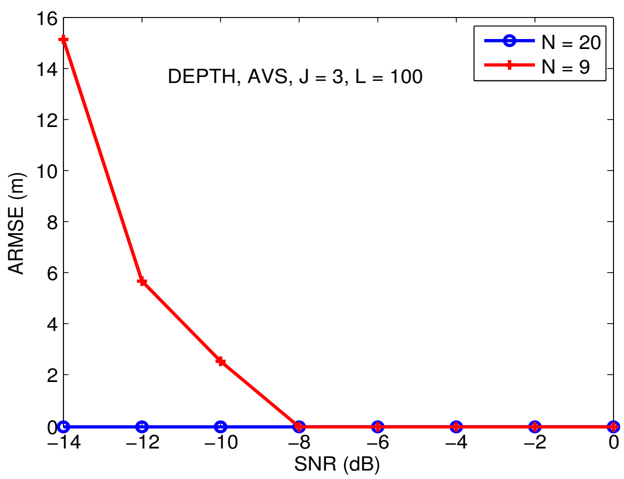

Fig. 3: Plots of ARMSE vs. SNR for depth estimation by MML method, for three sources at random locations.

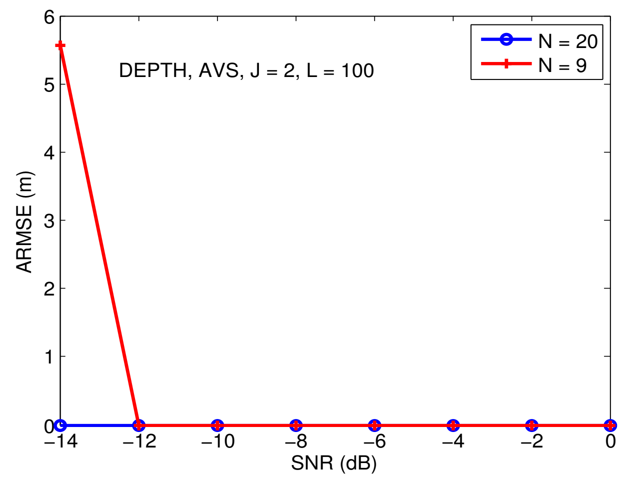

Fig. 5: Plots of ARMSE vs. SNR for depth estimation by MML method, for two sources at fixed locations.

small number of 9 sensors. Similar plots for the case of two sources at fixed positions, viz. $(3000 \text{ m}, 25 \text{ m}, 75°)$ and $(5200 \text{ m}, 75 \text{ m}, 125°)$, are shown in Figs. 4 and 5. In this case, the ARMSE values are lower at low SNR (< -8 dB) for two reasons: (1) the number of sources is less, and (2) the possibility of low estimation accuracy due to the placement of a source at an unfavorable position is avoided.

Table I presents a comparison of hardware and computational complexity for acoustic pressure sensor (APS), AVS, and compressively sampled AVS (CS-AVS) arrays with N sensors. It is seen that significant reduction in complexity is achieved if N is large. There is little scope for complexity mitigation by compressive sampling for a 9-sensor AVS VLA. We shall therefore restrict the application of compressive MML (CMML) to the problem of bearing estimation by AVS HLA.

C. Bearing Estimation

For bearing estimation we consider an AVS HLA of $N = 20$ sensors. Each snapshot of the HLA data vector $\boldsymbol{y}_h(t)$ of $3N = 60$ elements is compressed to a measurement vector $\boldsymbol{z}_h(t) = \boldsymbol{\Phi}_h \boldsymbol{y}_h(t)$ of $M = 3N_c$ elements, $N_c < N$. First, we consider 3 sources at random locations as described in Section IV-B. The signals from the three sources have equal SNR at the receiver. The compressive MML (CMML) processor is used for bearing estimation. Plots of ARMSE(bearing) vs. SNR for three different values of N_c, viz., $N_c = 3, 6, 9$ are shown in Fig. 6. The range-depth estimates required for the bearing estimation are obtained by the MML algorithm using an AVS VLA of (a) 20 sensors, and (b) 9 sensors. Results similar to those in Fig. 6, for the case of two sources at the fixed locations of $(3000 \text{ m}, 25 \text{ m}, 75°)$ and $(5200 \text{ m}, 75 \text{ m}, 125°)$, are shown in Fig. 7. Once again, it is seen that the ARMSE values are

Array type	**APS**	**AVS**	**CS-AVS**
Number of channels	N	$3N$	$J\log(3N)$
Signal acquisition hardware	N	$3N$	$J\log(3N)$
Correlation matrix size	$N\times N$	$3N\times 3N$	$J\log(3N)\times J\log(3N)$
Matrix inversion complexity	$\mathcal{O}(N^3)$	$\mathcal{O}((3N)^3)$	$\mathcal{O}((J\log(3N))^3)$
Eigendecomposition complexity	$\mathcal{O}(N^3)$	$\mathcal{O}((3N)^3)$	$\mathcal{O}((J\log(3N))^3)$
Maximum number of snapshots	N	$3N$	$J\log(3N)$

TABLE I. Comparison of complexity for different arrays with N sensors.

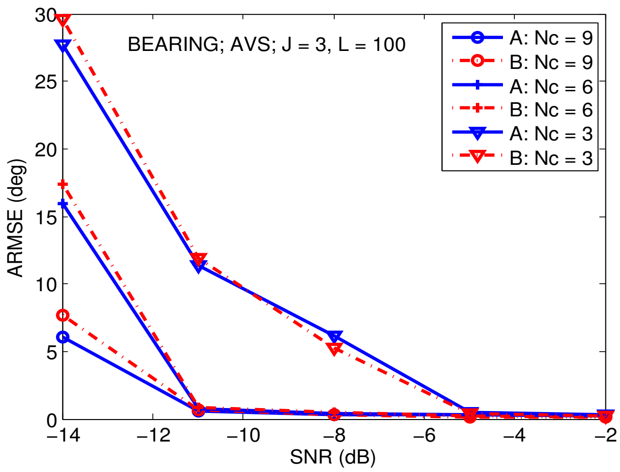

Fig. 6: Plots of ARMSE vs. SNR for bearing estimation by CMML ($N = 9$) method for different values of N_c, for three sources at random locations. Range-depth estimates were obtained by MML method using (A) 20-sensor VLA, (B) 9-sensor VLA.

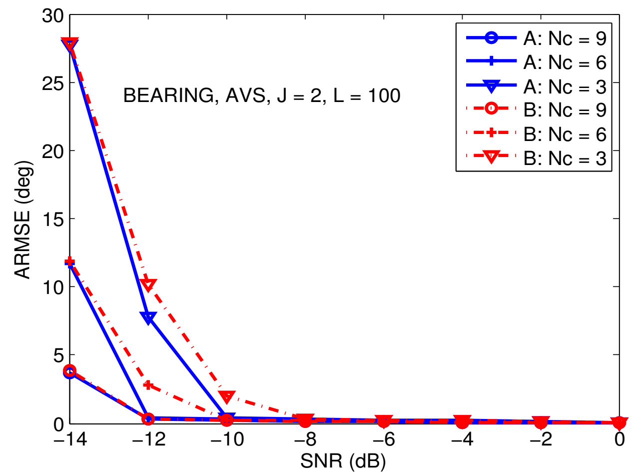

Fig. 7: Plots of ARMSE vs. SNR for bearing estimation by CMML method for different values of N_c, for two sources at fixed locations. Range-depth estimates were obtained by MML method using (A) 20-sensor VLA, (B) 9-sensor VLA.

lower at low SNR (< -8 dB) for reasons set out in Section IV-C.

The bearing estimation performances of CMML with $N = 20$, $N_c = 9$, MML ($N = 20$), and MML ($N = 9$), for 3 sources at random locations, are compared in Figs. 8 and 9. Figure 8 shows the variation of ARMSE with SNR for 200 snapshots, and Fig. 9 shows the variation of ARMSE with the number of snapshots at -8 dB SNR. It is seen from these figures that the performance of CMML $(N = 20, N_c = 9)$ is very close to that of MML ($N = 20$) for SNR > -11 dB. But the ARMSE is much higher for MML ($N = 9$). Hence we conclude that (a) significant reduction in complexity with negligible performance degradation can be achieved by using compressive sampling, and (b) reduction in the number of sensors without compressive sampling leads to a significant performance degradation even at high SNR.

V. Conclusion

Three dimensional localization of underwater acoustic sources can be done with high resolution and accuracy by the MML processor using data measured by an AVS 2DA, composed of an AVS VLA for range-depth estimation and an AVS

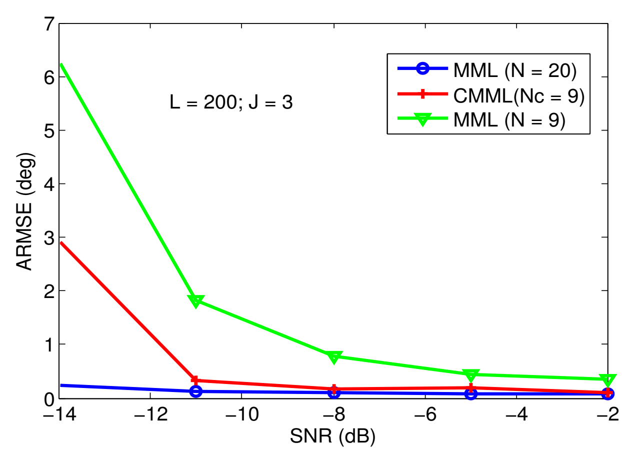

Fig. 8: Plots of ARMSE vs. SNR for bearing estimation of 3 sources at random locations by CMML $(N = 20, N_c = 9)$, MML $(N = 20)$ and MML $(N = 9)$ methods.

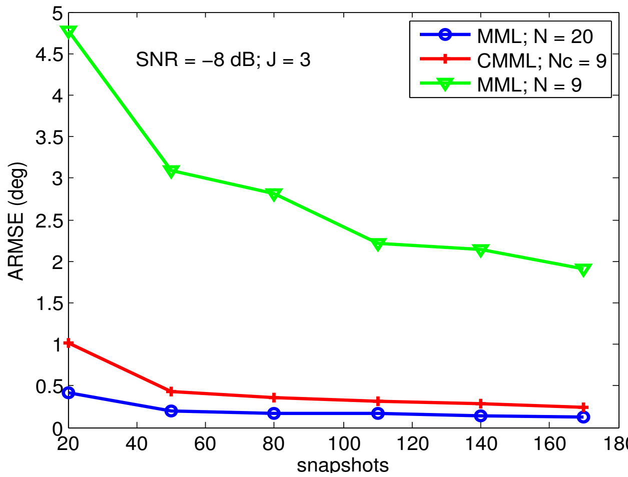

Fig. 9: Plots of ARMSE vs. number of snapshots for bearing estimation of 3 sources at random locations by CMML ($N = 20, N_c = 9$), MML ($N = 20$) and MML ($N = 9$) methods.

HLA for bearing estimation. This processor has a high level of hardware and software complexity because (1) each AVS acquires data on three channels, and (2) the HLA should have a large number of sensors to achieve high bearing resolution. In this paper we have proposed the use of compressive sampling of the AVS HLA output in the spatial domain for reducing the processor complexity. We have presented simulation results to illustrate the performance of the compressive MML processor. We have considered compression of the 60-dimensional data vector at the output of a 20-sensor AVS HLA to a 27-dimensional vector (equivalent to an AVS HLA of 9 sensors). It is shown that the performance of the MML processor for bearing estimation of 3 randomly placed sources using the compressively sampled data is very close to that of the MML processor using the original 20-sensor data, for SNR above -11 dB. At higher SNR, complexity can be reduced further by increasing the compression ratio, without compromising the performance. But the performance of the MML processor using uncompressed data at the output of a 9-sensor AVS HLA is significantly inferior even at high SNR.

References

[1] M. J. Hinich, "Maximum likelihood signal processing for a vertical array," *J. Acoust. Soc. Amer.*, vol. 54, no. 2, pp. 499–503, Aug 1973.

[2] H. P. Bucker, "Use of calculated sound fields and matched field detection to locate sound sources in shallow water," *J. Acoust. Soc. Amer.*, vol. 59, no. 2, pp. 368–372, Feb 1976.

[3] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, no. 3, pp. 276–280, Mar 1986.

[4] J. Tabrikian and H. Messer, "Three-dimensional source localization in a waveguide," *IEEE Trans. Sig. Process.*, vol. 44, no. 1, pp. 1–13, Jan 1996.

[5] P. V. Nagesha and G. V. Anand, "Three-dimensional localization of underwater acoustic sources by a linear vector sensor array," in *Proc. 12th Western Pacific Acoustics Conference*, Dec 2015, pp. 604–612.

[6] Y. Wang, G. Leus, and P. Pandharipande, "Direction estimation using compressive sampling array processing," in *Proc. IEEE/SP 15th Workshop on Statistical Signal Processing*, Sep 2009, pp. 626–629.

[7] K. G. Nagananda and G. V. Anand, "Subspace intersection method of high resolution bearing estimation in shallow ocean using acoustic vector sensors," *Sig. Process.*, vol. 90, no. 1, pp. 105–118, Jan 2010.

[8] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov 2001.

[9] M. B. Porter, "The kraken normal mode program," in *Tech. rep., SACLANTCEN Memorandum, U. S. department of commerce, National Technical Information Service*, 1991.