DIRECTION FINDING USING SPARSE LINEAR ARRAYS WITH MISSING DATA

Mianzhi Wang, Zhen Zhang, and Arye Nehorai

Preston M. Department of Electrical and Systems Engineering
Washington University in St. Louis, USA

ABSTRACT
We investigate the problem of direction of arrival (DOA) estimation using sparse linear arrays, such as co-prime and nested arrays, in the case of missing data resulting from sensor failures. We introduce a signal model where sensor failures occur after taking certain number of snapshots. We formulate a structured covariance estimation problem by exploiting the special geometry of sparse linear arrays, which also provides enhanced degrees of freedom. Numerical examples show that, by utilizing the information in both complete measurements and incomplete measurements, our method achieves better estimation accuracy than the traditional method using only complete measurements.

Index Terms—DOA estimation, missing data, maximum likelihood, coprime array, nested array

1. INTRODUCTION
Sparse linear arrays, such as minimum redundancy arrays (MRA) [1–3], nested arrays [4–6], and co-prime arrays [7–10], have the attractive property of providing $O(M^2)$ degrees of freedom with only $M$ sensors. The extra degrees of freedom are exploited by constructing an augmented covariance matrix from the difference coarray model [11, 12]. Due to their dependence on the coarray geometry, sparse linear arrays are more susceptible to sensor failures. If the measurements from one or more sensors are missing, the coarray structure will be partially destroyed, leading to performance degradation and loss of degrees of freedom.

Tackling missing data is important in robust DOA estimation, and many previous work has addressed the problem. Notably, in [13], Larsson et al. proposed a Cholesky parameterization based maximum likelihood estimator, and analyzed its asymptotic performance. However, their model is based on uniform linear arrays (ULAs), and requires a sequential failure pattern. In practice, any sensor may fail, so the sequential assumption may not be true. Recent advances in matrix completion [14, 15] and atomic norm minimization [16, 17] also bring new methods to tackle the missing data problem. By exploiting the low-rank property of the signal subspace, it is possible to extrapolate the missing data via semidefinite programming (SDP). However, when the number of measurements is large, the resulting SDP will be computationally expensive to solve. In this paper, we consider the direction finding problem using general sparse linear arrays with incomplete measurements. We do not assume a sequential failure pattern. We focus on deriving an algorithm that utilizes the information in both complete measurements and incomplete measurements based on the maximum-likelihood approach. We first estimate the augmented covariance matrix by exploiting its Toeplitz structure, and then apply the MUSIC algorithm [18] to obtain the DOA estimates. We derive the Cramèr Rao bound (CRB) and confirm the efficacy of our algorithms via numerical examples.

2. PROBLEM FORMULATION
We consider a sparse linear array whose sensors are located on a uniform grid. We represent the sensor locations by the integer set $D = \{d_1,d_2,\ldots,d_M\}$, where $M$ is the number of sensors. The actual sensor locations $d_i$ are given by $d_id_0$ for $i = 1,2,\ldots,M$, where $d_0$ denotes the grid size. Such an array can also be viewed as a thinned uniform linear array (ULA) that consists of $M_0 = d_M + 1$ sensors. For example, a co-prime array whose sensors are located at $[0, 2, 3, 4, 6, 9]d_0$ can be viewed as a 10-sensor ULA with the 2nd, 6th, 8th, and 9th sensors removed.

We consider $K$ far-field uncorrelated narrowband signals impinging on the array from directions $\theta = [\theta_1,\theta_2,\ldots,\theta_K]^T$. The received signal vectors are given by

$$y(t) = SA_{U}(\theta)x(t) + n(t), \quad t = 1,2,\ldots,N,$$

where $A_U(\theta) = [a_U(\theta_1), a_U(\theta_2),\ldots,a_U(\theta_K)]$ is the steering matrix of a $M_0$-sensor ULA [19]. $S$ is a $M \times M_0$ selection matrix, where $S_{mn}$ is one if and only if the $m$-th sensor in the sparse linear array corresponds to the $n$-th sensor in the ULA, and otherwise zero. $x(t)$ is the source signal, and $n(t)$ is the additive complex white Gaussian noise. We assume that the source signals follow the unconditional model [19], and there is no temporal correlation between each snapshot.

With the above assumptions, the covariance matrix is given by

$$R = \mathbb{E}[y(t)y^H(t)] = SR_US^T,$$

where $R_U = A_UPA_U^T + \sigma_i^2I$, $P = \text{diag}(p_1,p_2,\ldots,p_K)$, and $p_k$ is the power of $k$-th source. Therefore the covariance matrix of a sparse linear array is a compressed version of the covariance matrix of a ULA. By vectorizing $R$, we obtain

$$r = (S \otimes S)(A_U^T \otimes A_U)p + \sigma_i^2I,$$

where $r = \text{vec}(R)$, $p = [p_1,p_2,\ldots,p_K]^T$, and $I = \text{vec}(I)$, $\otimes$ denotes the Kronecker product, $\odot$ denotes the Khatri-Rao product (i.e., column-wise Kronecker product), and vec($A$) converts $A$ into a column vector by stacking the columns of $A$ [20]. Model (2) resembles a measurement model with deterministic sources and noise, and $(S \otimes S)(A_U^T \otimes A_U)$ embeds a steering matrix of a virtual array with enhanced degrees of freedom, whose sensor locations are given by $D_{co} = \{(d_m - d_n)d_m, d_n \in D\}$. If $D_{co}$ consists of consecutive integers from $-M_0 + 1$ to $M_0 - 1$, we call the sparse linear array complete. If a sparse linear array is complete (e.g., minimum

This work was supported by the ONR Grant N000141310050.
we define during the first period, we can simply remove them and form a new sparse linear array whose sensors are all functional during the first period. During the l-th period (2 ≤ l ≤ L), sensors fail and the measurement data from these sensors is missing. Let M_l be the number of working sensors during the l-th period. Let T_l be a selection matrix of size M_l × M_l such that the (i, j)-th element of T_l is one if and only if the j-th sensor in the sparse linear array is the i-th sensor in the l-th period. Let T_1 be a selection matrix of size M_1 × M_1 such that the (i, j)-th element of T_1 is one if and only if the j-th sensor in the sparse linear array is the i-th sensor in the l-th period, and otherwise zero. For notational simplicity, we define T_l = I_{M_l}. After discarding the measurements from the malfunctioning sensors, the snapshots taken during the l-th period are given by

\[ y_l(t) = T_l[S A_u(\theta)] x(t) + n_l(t), \quad (3) \]

for \( t = N_1 + \cdots + N_{l-1} + 1, \ldots, N_1 + \cdots + N_{l-1} + N_l \), where \( N_l \) is the number of snapshots collected during the l-th period. The total number of snapshots is denoted by \( N = \sum_{l=1}^{L} N_l \). Correspondingly, we can collect L different sample covariance matrices \( \hat{R}_l = 1/N_l \sum_{t = N_{l-1} + 1}^{N_l} y_l(t) y_l^H(t), \) \( l = 1, 2, \ldots, L \). We also define their expectations as

\[ R_l = E[\hat{R}_l] = T_l S R U T_l^T + \sigma_n^2 I, \quad (4) \]

whose vectorized versions are given by

\[ r_l = \text{vec}(R_l) = (T_l S \otimes T_l S) (A_u^\perp \otimes A_u) p + \sigma_n^2 i. \quad (5) \]

Because of the missing data, \( (T_l S \otimes T_l S) (A_u^\perp \otimes A_u) \) no longer embeds a desired virtual array steering matrix and existing methods cannot be directly applied. If we use only \( R_1 \) for estimation, we lose much information provided in \( \hat{R}_l \) (2 ≤ l ≤ L). Therefore an estimator that utilizes all the information in \( \hat{R}_l \) (1 ≤ l ≤ L) is desired.

3. Estimation in the Presence of Missing Data

3.1. Ad-hoc Estimator

The ad-hoc estimator for our signal model is inspired by redundancy averaging [22, 23], and is an extension of the ad-hoc estimator in [13]. Let \( V_k = \{ (m, n) | d_{m} - d_{n} = k, d_{m}, d_{n} \in D \} \). Let \( A_{m,n} \) denote the set of snapshot indices when both the m-th and the n-th sensor are working. We define

\[ u_k = \frac{\sum_{(m,n)\in V_k} \sum_{\ell \in A_{m,n}} y_m(t) y_n(t)}{\sum_{(m,n)\in V_k} |A_{m,n}|}, \quad (6) \]

where \( y(t) = [y_1(t), \ldots, y_M(t)] \) is the full measurement vector before discarding invalid data, \( y_m(t) \) is the output of the m-th sensor, and \(|A|\) denotes the cardinality of \( A \). For complete arrays, we can obtain \( u_k \) for \( k = -M_0 + 1, -M_0 + 2, \ldots, M_0 - 1 \), and the ad-hoc estimate of \( R_U \) is given by

\[ \hat{R}_U^{(\text{ad-hoc})} = \begin{bmatrix} u_0 & u_{-1} & \cdots & u_{-M_0+1} \\ u_1 & u_0 & \cdots & u_{-M_0+2} \\ \vdots & \ddots & \ddots & \vdots \\ u_{M_0-1} & \cdots & u_0 & u_{M_0} \end{bmatrix}. \quad (7) \]

We can then apply MUSIC or other DOA estimation methods to \( \hat{R}_U^{(\text{ad-hoc})} \) to obtain the DOA estimates.

3.2. Maximum-Likelihood Based Estimators

Based on the results in [24], the negative log-likelihood function of our model is given by

\[ L(R_1, \ldots, R_L) = \sum_{l=1}^{L} N_l \log |R_l| + \text{tr}(R_l^{-1} \hat{R}_l), \quad (8) \]

where we have omitted the constants.

Observe that \( R_U \) is Hermitian Toeplitz. It is possible to reparameterize \( R_U \) by exploiting the Toeplitz structure, and the estimation of \( R_U \) becomes a structure covariance estimation problem. In the following discussion, we consider only complete arrays. Extension to non-restricted arrays will be discussed in the remarks.

Following the idea in [25], we construct the structured matrices as follows. Let \( I_M^{(i)} \) denotes the \( M \times M \) matrix whose elements are zero except for the i-th upper diagonal (i.e., \( I_M^{(i)}(m, n) = \delta(n - m - i) \)), where \( \delta(n) \) is the Kronecker delta. For a given positive integer M, we define the matrices \( Q_M^{(i)} \) as

\[ Q_M^{(i)} = \begin{cases} M, & i = 1, \\ (1^{(i-1)})^T + j(1^{(i-1)} - M)^T, & 2 \leq i \leq M, \\ -(1^{(i-M)})^T + j(1^{(i-M)})^T, & M + 1 \leq i \leq 2M - 1. \end{cases} \quad (9) \]

Then we are able to express \( R_U \) as

\[ R_U = \sum_{i=1}^{2M_0-1} c_i Q_M^{(i)} \quad (10) \]

where \( c = [c_1, c_2, \ldots, c_{2M_0-1}]^T \in \mathbb{R}^{2M_0-1} \) is the Hermitian Toeplitz parameterization of \( R_U \). After obtaining its estimate, we can reconstruct \( R_U \) from (10) and then perform DOA estimation. Substituting (10) into (8) and taking the derivative with respect to \( c_i \), we obtain

\[ \frac{\partial L(c)}{\partial c_i} = \sum_{l=1}^{L} N_l \text{tr} \left[ T_l S Q_M^{(i)} T_l^H R_l^{-1} (R_l - \hat{R}_l) R_l^{-1} \right] \]

for \( i = 1, 2, \ldots, 2M_0 - 1 \). Because \( \text{vec}(AXB) = (B^T \otimes A) \text{vec}(X) \), and because \( (A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \) for non-singular \( A, B \) [20], we have

\[ \text{vec}(T_l S Q_M^{(i)} T_l^H) = \Phi_i d_M^{(i)} \quad (11) \]
where $q_{M_0}^{(i)} = \text{vec}(Q_{M_0}^{(i)})$, and $\Phi_i = T_i S \otimes T_i S$. We also have
\[ \text{vec}(R_i^{-1}(R_i^- \tilde{R}_i)R_i^{-1}) = W_i^{-1}(\Phi_i Q_{M_0} c - \tilde{r}_i), \] (12)
where $W_i = R_i^T \otimes R_i$, $Q_{M_0} = [q_{M_0}^{(1)}, q_{M_0}^{(2)}, \cdots, q_{M_0}^{(2M_0-1)}]$, and $\tilde{r}_i = \text{vec}(R_i)$. Let all the partial derivatives with respect to $c_i$ be zero. Then, we utilize (11) and (12) to obtain
\[ \left( \sum_{i=1}^{L} N_i G_i \right) c = \sum_{i=1}^{L} N_i h_i, \] (13)
where $G_i = Q_{M_0}^{(i)} \Phi_i^T W_i^{-1} \Phi_i Q_{M_0}$, and $h_i = Q_{M_0}^{(i)} \Phi_i^T W_i^{-1} \tilde{r}_i$. Note that if we have sufficient snapshots in each period, $\tilde{R}_i$ will be very close to $R_i$, and we can replace $W_i$ with its estimate $\hat{W}_i = R_i^T \otimes \hat{R}_i$. In this case the only unknown in (13) will be $c$, whose estimate can be readily given by
\[ \hat{c}_{\text{WLS}} = \left[ \sum_{i=1}^{L} N_i G_i \right]^{-1} \left[ \sum_{i=1}^{L} N_i \hat{h}_i \right]. \] (14)
where $\hat{G}_i$ denotes $G_i$ with $W_i$ replaced by $\hat{W}_i$, and $\hat{h}_i$ denotes $h_i$ with $W_i$ replaced by $\hat{W}_i$. Lemma 2 ensures that (14) produces real results.

**Lemma 1.** Let $A, B, C$ be Hermitian symmetric. Then $\text{tr}(ABC)$ is real.

**Proof.** This can be shown by the fact that
\[ \text{tr}(ABC)^* = \text{tr}((ABC)^H) = \text{tr}(CAB)A = \text{tr}(ABC). \]
\[ \square \]

**Lemma 2.** Both $\hat{G}_i$ and $\hat{h}_i$ are real.

**Proof.** Through algebraic manipulations, the $(m, n)$-th element of $\hat{G}_i$ can be rewritten as
\[ \text{tr}[(\hat{R}_i^{-1}T_i S)^{\Theta} M_0^{(m)} S^* T_i^\dagger \hat{R}_i^{-1} T_i S Q_{M_0}^{(n)} S^* T_i^\dagger]. \]
By the definition of $Q_{M_0}^{(m)}$ in (9), we know that $T_i S Q_{M_0}^{(m)} S^* T_i^\dagger$ is Hermitian symmetric. Because $\hat{R}_i^{-1}$ is also Hermitian symmetric, we know that each element of $\hat{G}_i$ is real by Lemma 2. The proof for the second claim follows the same idea.
\[ \square \]

We call (14) the “weighted least squares” (WLS) estimate, because (14) is the solution to the weighted least squares problem: $\min_c \sum_{i=1}^{L} N_i \| \Phi_i Q_{M_0} c - \tilde{r}_i \|^2_{W_i^{-1}}$, where $\| x \|^2_W = \sqrt{x^H W x}$.

We can also observe that (13) leads to the following fixed-point type iteration:
\[ \hat{c}_{k+1}^{(l)} = \left[ \sum_{i=1}^{L} N_i G_i \left( \hat{c}_{k+1}^{(l)} \right) \right]^{-1} \left[ \sum_{i=1}^{L} N_i h_i \left( \hat{c}_{k+1}^{(l)} \right) \right], \] (15)
where $G_i (\hat{c}_{k+1}^{(l)})$ and $h_i (\hat{c}_{k+1}^{(l)})$ are constructed from $\hat{c}_{k+1}^{(l)}$.

**Remark 1.** In practice, the computation of $G_i$ and $h_i$ can be efficiently implemented by exploiting the properties of Kronecker product and the fact that $\Phi_i$ are Kronecker products of simple selection matrices. In our experiments, by setting the initial value as $\hat{c}_{\text{WLS}}, \{ \hat{c}_{k+1}^{(l)} \}$ showed good convergence in a few iterations.

Because the measurements are assumed independent, the $(m, n)$-th element of the Fisher information matrix (FIM) for our signal model is given by [24, 27]:
\[ \text{FIM}_{mn} = \sum_{i=1}^{L} N_i \text{tr} \left[ \frac{\partial R_i}{\partial \eta_m} \frac{\partial R_i}{\partial \eta_n} - \hat{R}_i^{-1} \right]. \]
Using the properties of the Kronecker product, we can express the FIM as
\[ \text{FIM}_{mn} = \sum_{i=1}^{L} N_i \left[ \text{tr} \left( \frac{\partial r_i}{\partial \eta_m} \right)^H \Phi_i^H (R_i^T \otimes R_i) \Phi_i \frac{\partial r_i}{\partial \eta_n} \right], \]
where $r_i = \text{vec}(R_i)$. Therefore, for complete arrays, the FIM for the Toeplitz parametrization is given by
\[ \text{FIM}_{c} = \sum_{i=1}^{L} N_i Q_{M_0}^{(i)} \Phi_i^H (R_i^T \otimes R_i) \Phi_i Q_{M_0}, \] (16)
For incomplete arrays, as stated in Remark 3, not all elements in $R_i$ are present in $\hat{R}_i$. Therefore $Q_{M_0}$ is no longer full rank, and we cannot perform the matrix inversion in (14) or (15). In this case, we first delete the elements we cannot estimate from $c$ and their corresponding basis matrices from $\{Q_{M_0}^{(i)}\}_{i=1}^{2M_0-1}$ to form $\hat{c}$ and $\hat{Q}_{M_0}$. We then estimate $\hat{c}$ using (14) or (15), with $Q_{M_0}$ replaced by $\hat{Q}_{M_0}$. Finally, we construct a submatrix of $\hat{R}_i$ from the estimated $\hat{c}$.

4. PERFORMANCE BOUNDS

5. NUMERICAL EXAMPLES

We consider the following two sparse linear array configurations in the numerical examples:
- Nested array: $[0, 1, 2, 3, 7, 11, 15, 19]d_0$.
- Coprime array: $[0, 3, 5, 6, 9, 10, 12, 15, 20, 25]d_0$.
Fig. 1 illustrates the performance of different algorithms for the nested array configuration. We observe that TML-FP achieves the best performance, and is very close to the CRB, while “First” results in the worst performance because it cannot utilize the information in \( \hat{\mathbf{R}}_l \) \((l \geq 2)\). We observe similar results for the co-prime configuration in Fig. 2. However, a gap exists between the RMSE of TML-FP and the CRB, which may be attributed to the fact that the co-prime array is incomplete.

### 6. CONCLUSION AND FUTURE WORK

In this paper, we discussed the problem of direction finding using sparse linear arrays with incomplete measurements. By exploiting the coarray structure, we proposed to reconstruct a covariance matrix with enhanced degrees of freedom using the Toeplitz parameterization. Specifically, by applying our method to co-prime and nested arrays, we can resolve more sources than the number of sensors in the missing data case. We used numerical examples to show that our method has better accuracy than the traditional method using only the complete measurements. Potential future work includes performance and identifiability analysis in the presence of missing data.

### 7. REFERENCES


