ANALYTICAL PERFORMANCE ASSESSMENT OF 1-D STRUCTURED LEAST SQUARES

Florian Roemer and Martin Haardt
Ilmenau University of Technology, Communications Research Laboratory
P.O. Box 100565, D-98684 Ilmenau, Germany, http://tu-ilmenau.de/crl
{florian.roemer,martin.haardt}@tu-ilmenau.de

Abstract — In this paper, we derive the analytical performance of 1-D standard ESPRIT and 1-D Unitary ESPRIT using one iteration of Structured Least Squares to solve the shift invariance equations. First, we provide the estimation error of the k-th spatial frequency as an explicit expression of the noise realization, which requires no assumptions about the statistics of the noise. Then, we compute the statistical expectation over zero mean circularly symmetric white noise and provide explicit formulas for the resulting mean square errors. All expressions are asymptotic in the effective SNR, i.e., they become exact as either the number of snapshots or the SNR tends to infinity.

1. INTRODUCTION

Subspace-based high-resolution parameter estimation schemes represent an important tool in signal processing and can be applied to a broad range of applications, such as, communications, localization, biomedical imaging, channel modeling, and many more. A prominent family among these schemes are the ESPRIT-type algorithms, which are attractive since they are fully algebraic and hence simple to implement, e.g., standard ESPRIT [7], R-D Unitary ESPRIT [2], R-D standard Tensor-ESPRIT, R-D Unitary Tensor-ESPRIT [3], and many more.

In every ESPRIT-type algorithm the overdetermined shift-invariance equations are solved for an unknown matrix that contains the parameters of interest. This step requires an appropriate Least Squares (LS) scheme. If only the matrix of interest is considered unknown this gives rise to the LS estimator. However, LS is not optimal, since the shift-invariance equations have estimation errors on both sides of the equation. The Total Least Squares (TLS) [4] approach takes these estimation errors into account and hence improves the estimation accuracy. However, for an array with overlapping shift invariant subarrays, the estimation errors on both sides of the shift invariance equation are not independent anymore. Therefore, taking this structure into account, the estimation accuracy can be improved. This is achieved by the Structured Least Squares (SLS) [1] algorithm, which outperforms LS and TLS, especially for correlated sources and a large number of antennas.

In [6], we have introduced an analytical performance assessment for 2-D standard ESPRIT, 2-D Unitary ESPRIT, 2-D standard Tensor-ESPRIT, and 2-D Unitary Tensor-ESPRIT. We have shown that a first-order approximation of the estimation error in the spatial frequencies can be expressed as an explicit function of the noise realization, requiring no assumptions about the noise statistics. We have also computed the statistical expectation over circularly symmetric white noise, leading to compact MSE expressions. However, in [6], only the LS solution of the invariance equations was shown.

In this paper, we extend the work of [6] to the study of Structured Least Squares. We consider the 1-D case only and derive the estimation error for 1-D standard ESPRIT and 1-D Unitary ESPRIT using SLS. As in [6], the explicit expansion in terms of the noise realization as well as the MSE over zero mean circularly symmetric white noise is provided. Simulation results demonstrate the validity of our derivations.

The following notation is used throughout the paper: matrices are denoted by upper-case bold-faced letters, vectors by lower-case bold-faced letters, and scalars by italic letters. The superscripts $^*$, $^T$, $^H$, and $^\dagger$ represent complex conjugation, matrix transposition, Hermitian transposition, matrix inversion, and the Moore-Penrose pseudo inverse, respectively. The vec-operator vec $\{A\}$ aligns the elements of $A$ in a vector. Statistical expectation is represented by $E\{\cdot\}$. The Kronecker product between two matrices $A$ and $B$ is symbolized by $A \otimes B$. Estimated quantities are denoted by a hat, i.e., $\hat{A}$ is an estimate of $A$, such that $\hat{A} = A + \Delta A$. When approximate signs are written this implies an equation that becomes asymptotically true as the effective SNR goes to infinity. Consequently, all expressions we derive are asymptotic either in the SNR or in the number of snapshots.

2. DATA MODEL AND 1-D STRUCTURED LEAST SQUARES

We consider $N$ subsequent observations of $d$ narrowband sources in the far field of an antenna array with $M$ sensors. The observations can be modeled as

$$X = X_0 + N,$$

$$X_0 = A \cdot S$$

where $X \in \mathbb{C}^{M \times N}$ is the measurement matrix, $A \in \mathbb{C}^{M \times d}$ represents the array steering matrix containing the $d$ array steering vectors, $S \in \mathbb{C}^{d \times N}$ represents the source symbol matrix, and $N$ consists of samples of the i.i.d. circularly symmetric complex Gaussian distributed noise samples with variance $\sigma_n^2$. Since we are applying ESPRIT to estimate the parameters, we require the antenna array to possess a shift invariance structure, i.e.,

$$J_1 \cdot A \cdot \Phi = J_2 \cdot A,$$

where $J_1$ and $J_2$ are the $m \times M$ selection matrices which select the first and the second subarray containing $m$ out of the $M$ sensors ($m < M$). Moreover, $\Phi$ is given by

$$\Phi = \text{diag}\{[e^{j \mu_1}, \ldots, e^{j \mu_d}]\} \in \mathbb{C}^{d \times d}.$$
If the array obeys the shift invariance according to (3) we can estimate the spatial frequencies via 1-D ESPRIT. First, we estimate the signal subspace by computing the \( d \) dominant left singular vectors of \( X \) via an SVD and align them in a matrix \( \hat{U}_s \in \mathbb{C}^{d \times d} \). Since \( \hat{U}_s \) and \( \Lambda \) span approximately the same column space, we can write
\[
\hat{U}_s \cdot T \approx A,
\]
for a non-singular matrix \( T \in \mathbb{C}^{d \times d} \). Inserting (5) into (3) we obtain the shift invariance equation in terms of the estimated subspace leading to
\[
J_1 \cdot \hat{U}_s \cdot \Psi \approx J_2 \cdot \hat{U}_s,
\]
where \( \Psi = T \cdot \Phi \cdot T^{-1} \). The matrix \( \Psi \) is estimated from (6) via an appropriate least squares method, which we will discuss below. Once an estimate of \( \Psi \) is available, the desired spatial frequencies are given by
\[
\{ \hat{\mu}_1, \ldots, \hat{\mu}_d \} = \text{angle}\{ \text{eig}(\hat{\Psi}) \}.
\]
The most straightforward way to estimate \( \Psi \) from (6) is using Least Squares (LS), i.e.,
\[
\hat{\Psi}_{LS} = \left( J_1 \cdot \hat{U}_s \right)^+ \cdot J_2 \cdot \hat{U}_s.
\]
However, this solution does not take into account that the matrix \( \hat{U}_s \) is not exactly known but prone to an estimation error. Consequently, the estimation accuracy can be improved by taking the estimation error explicitly into account. A famous solution is the Total Least Squares (TLS) algorithm [4], which allows for estimation errors on both sides of the shift invariance equation (6). However, these estimation errors are considered as being independent from each other. For an array where the subarrays overlap, it should be taken into account that the estimation errors are not independent. This is achieved by the Structured Least Squares (SLS) algorithm [1], which solves the following optimization problem
\[
\min_{\Delta \Psi, \Delta U_s} \left\| J_1 \cdot \hat{U}_s + \Delta U_s \right\|_F^2 + \kappa^2 \cdot \left\| \Delta U_s \right\|^2_F,
\]
where \( \kappa \) is a regularization constant. In other words, SLS improves the LS fit by allowing for an error in both \( \Psi \) and \( U_s \). The non-linear optimization problem in (9) can be solved iteratively by local linearization. However, as shown in [1], computing one iteration is sufficient, so that the resulting algorithm is not iterative in its nature. Consequently, the asymptotic performance analysis presented in this paper is also limited to the first iteration.

3. PERFORMANCE ANALYSIS

To analyze the benefit of SLS, we use the analytical framework established in [5]. In this paper it is shown that the estimation error of the \( k \)-th spatial frequency estimated via LS can be computed as
\[
\Delta \hat{\mu}_k \approx \text{Im} \left\{ \frac{\text{vec} \left[ \Delta P^{(k)} \cdot \Delta \Psi_{LS} \cdot q_k \right]}{\lambda_k} \right\},
\]
\[
\Delta \Psi_{LS} \approx \left( J_1 \cdot U_s \right)^+ \cdot \left( J_2 \cdot \Delta U_s - J_1 \cdot \Delta U_s \cdot \Psi \right).
\]
Here, the vectors \( q_k \) and \( P^{(k)} \) represent the \( k \)-th column of the matrix \( Q \) and the \( k \)-th row of the matrix \( P = Q^{-1} \), respectively, where \( Q \) is given by the EVD of \( \Psi \) according to
\[
\Psi = Q \cdot \Lambda \cdot Q^{-1}.
\]
Moreover, \( \lambda_k = e^{j \nu_k} \) represents the \( k \)-th eigenvalue of \( \Psi \), i.e., the \( k \)-th diagonal entry of the matrix \( \Lambda \). The quantity \( \Delta U_s \) in (10) is the estimation error in the subspace for which [5] provides the explicit first-order expansion
\[
\Delta U_s \approx U_n \cdot U_n^H \cdot N \cdot V_n \cdot S_n^{-1}.
\]
Here, the columns of \( U_n \in \mathbb{C}^{M \times (M-d)} \) and \( V_n \in \mathbb{C}^{N \times d} \) span the null space and the row space, and \( S_n \in \mathbb{C}^{d \times d} \) represents the diagonal matrix of singular values of the unperturbed matrix \( X_0 \), respectively. These matrices are defined via the SVD of \( X_0 \)
\[
X_0 = [U_s, U_n] \cdot [S_n, 0] [0, 0]^H \cdot [V_s, V_n]^H.
\]
It should be emphasized that (13) is an explicit function of the noise realization \( N \) and therefore requires no assumptions about its particular statistics. In [6], we have shown that the statistical expectation over the squared estimation error from (10) for zero mean circularly symmetric white noise can be computed. This leads to an explicit expression for the MSE given by
\[
E \{ (\Delta \hat{\mu}_k)^2 \} = \frac{\sigma_n^2}{2} \left\| W_{U,N}^T \cdot r_{ls} \right\|_2^2, \quad \text{where}
\]
\[
r_{ls} = q_k \otimes \left( [(J_1 \cdot U_s)^+ \cdot (J_2 \cdot \lambda_k - J_1)]^T \cdot P_k \right).
\]
Moreover, \( W_{U,N} \) describes the linear mapping between \( \text{vec} \{ \Delta U_s \} = \Delta u_s \) and \( \text{vec} \{ N \} = n \) according to
\[
\Delta u_s \approx W_{U,N} \cdot n.
\]

Next, we derive similar relationships for the SLS-based estimate of the spatial frequencies. As mentioned in the previous section, we only consider one iteration. Moreover, since our analysis is asymptotic in the effective SNR we consider \( \kappa = 0 \), i.e., no regularization, since regularization is typically not needed for high SNRs. Under this assumption, the cost function (9) can be expressed as
\[
\Delta u_{s,SLS}, \Delta \psi_{SLS} = \arg \min_{\Delta \psi, \Delta U_s} \left\| r_{LS} + F_{SLS} \cdot \left[ \begin{array}{c} \Delta \psi \\ \Delta U_s \end{array} \right] \right\|_2^2,
\]
where \( \Delta \psi_{SLS} = \text{vec} \{ \Delta \Psi_{SLS} \} \), \( \Delta u_{s,SLS} = \text{vec} \{ \Delta U_{s,SLS} \} \), \( \Delta \psi = \text{vec} \{ \Delta \Psi \} \), and the residual vector \( r_{LS} \) as well as the update matrix \( F_{SLS} \) are given by
\[
r_{LS} = \text{vec} \left\{ J_1 \cdot \hat{U}_s \cdot \hat{\Psi}_{LS} - J_2 \cdot \hat{U}_s \right\},
\]
\[
F_{SLS} = \left( I_d \otimes (J_1 \cdot U_s) \right) \cdot \left( \hat{\Psi}_{LS}^T \otimes J_1 \right) - \left( I_d \otimes J_2 \right).
\]
Since (19) is linear, its solution is
\[
\Delta \psi_{SLS} \approx \Delta u_{s,SLS} = -F_{SLS}^+ \cdot r_{LS}.
\]
To arrive at a first order perturbation expansion for \( \Delta \psi_{SLS} \) from (21) we first rewrite \( r_{LS} \) from (20) by inserting \( \hat{U}_s = U_s + \Delta U_s \) and \( \hat{\Psi}_{LS} = \Psi + \Delta \Psi_{LS} \). With the help of (11) we can obtain an explicit first order expansion of \( r_{LS} \) with respect to \( \Delta u_s \)
\[
r_{LS} \approx \text{vec} \left\{ \frac{\hat{J}_1 \cdot (U_s + \Delta U_s)}{U_s + \Delta U_s} \cdot (\Psi + \Delta \Psi_{LS}) \right\}
\]
\[
- \frac{J_2 \cdot (U_s + \Delta U_s)}{U_s + \Delta U_s}) \right\}
\approx W_{R,U} \cdot \Delta u_s.
\]
where $W_{R,U}$ is given by
\[
W_{R,U} = \left( \mathbf{J}_1 \otimes \mathbf{J}_2 \right) + \mathbf{I}_d \otimes \left( \mathbf{J}_1 \cdot \mathbf{U}_k \cdot \mathbf{J}_1 \mathbf{U}_k + \mathbf{J}_2 \right) \\
- \mathbf{J}_1 \cdot \mathbf{U}_k \cdot \mathbf{J}_1 - \mathbf{I}_d \otimes \mathbf{J}_2.
\]

This provides a first order expansion for $r_{k}$ in $\Delta \mathbf{u}_k$, which is readily transformed into a first order expansion in $\mathbf{r}$ with the help of (17). To expand (21) we still need to consider $\hat{F}_{\text{SLS}}^+$. First of all, note that $\hat{F}_{\text{SLS}}^+$ is of size $(d \cdot m) \times (d \cdot M + d^2)$ and hence never tall. Moreover, it has full row rank with probability one, given that all spatial frequencies are distinct. Furthermore, we note that $\hat{F}_{\text{SLS}}^+$ can be seen as an estimate of the matrix $F_{\text{SLS}}$ given by
\[
F_{\text{SLS}} = \left[ \mathbf{I}_d \otimes (\mathbf{J}_1 \cdot \mathbf{U}_k), (\hat{\mathbf{J}}_1 \otimes \mathbf{J}_1) - (\mathbf{I}_d \otimes \mathbf{J}_2) \right]. \tag{24}
\]

Consequently, $\hat{F}_{\text{SLS}}^+ = F_{\text{SLS}}^+ + \Delta F_{\text{SLS}}$, where $\Delta F_{\text{SLS}}$ only depends on $\Delta \hat{\mathbf{J}}_1$ and $\Delta \mathbf{U}_k$, and hence is a small perturbation term itself. This shows that $F_{\text{SLS}}^+$ can be expressed as
\[
\hat{F}_{\text{SLS}}^+ = F_{\text{SLS}}^+ + O(\Delta), \tag{25}
\]
i.e., the pseudo-inverse of $\hat{F}_{\text{SLS}}^+$ is equal to the pseudo inverse of $F_{\text{SLS}}^+$ plus a term which scales linear in the perturbation, plus higher-order terms. Inserting (25) and (23) into (21) we observe that the linear term in the expansion of $\hat{F}_{\text{SLS}}^+$ gives to a quadratic term in (21), since $r_{k}$ is already linear in the perturbation. This quadratic term can be neglected and we arrive at
\[
\Delta \hat{\mathbf{J}}_{\text{SLS}} \approx - F_{\text{SLS}}^+ \cdot W_{R,U} \cdot \Delta \mathbf{u}_k. \tag{26}
\]

Since $\Delta \mathbf{u}_k$ is not directly needed as long as only a single iteration is performed, we rewrite (26) slightly to extract only $\Delta \hat{\mathbf{J}}_{\text{SLS}}$. This can be achieved by exploiting the fact that $F_{\text{SLS}}$ is a flat matrix with full row rank so that its pseudo-inverse is given by $F_{\text{SLS}}^+ = F_{\text{SLS}}^H \cdot (F_{\text{SLS}}^H \cdot F_{\text{SLS}}^H)^{-1}$. Consequently, to extract $\Delta \hat{\mathbf{J}}_{\text{SLS}}$, we can simplify the second block of $F_{\text{SLS}}^+$ and then arrive at
\[
\Delta \hat{\mathbf{J}}_{\text{SLS}} = - W_{\mathbf{J},R} \cdot W_{R,U} \cdot \Delta \mathbf{u}_k, \tag{27}
\]
where
\[
W_{\mathbf{J},R} = (\mathbf{I}_d \otimes (\mathbf{J}_1 \cdot \mathbf{U}_k))^H \cdot (F_{\text{SLS}}^H \cdot F_{\text{SLS}}^H)^{-1}. \tag{28}
\]

The final estimate for $\mathbf{J}$ via SLS is formed as $\hat{\mathbf{J}}_{\text{SLS}} + \Delta \hat{\mathbf{J}}_{\text{SLS}}$. Therefore, the estimation error in $\hat{\mathbf{J}}$ after the first iteration of SLS is given by
\[
\hat{\mathbf{J}}_{\text{SLS}} - \mathbf{J} = \Delta \hat{\mathbf{J}}_{\text{SLS}} + \Delta \hat{\mathbf{J}}_{\text{SLS}}. \tag{29}
\]

Consequently, we can express the first order expansion in the $k$-th spatial frequency by inserting (29) into (10) which yields
\[
\Delta \mu_{k,\text{SLS}} \approx \text{Im} \left\{ p_k^T \cdot (\Delta \mathbf{J}_{\text{SLS}} + \Delta \hat{\mathbf{J}}_{\text{SLS}}) \cdot q_k \right\} / \lambda_k \tag{30}
\]
Inserting (27) and (11) this can be simplified into
\[
\Delta \mu_{k,\text{SLS}} \approx \text{Im} \left\{ r_{k,\text{SLS}}^T \cdot \Delta \mathbf{u}_k \right\} \tag{31}
\approx \text{Im} \left\{ r_{k,\text{SLS}}^T \cdot W_{U,N} \cdot \mathbf{n} \right\}. \tag{32}
\]
where the vector $r_{k,\text{SLS}}^T$ is given by
\[
r_{k,\text{SLS}}^T = q_k^T \otimes p_k^T \cdot (\mathbf{J}_1 \cdot \mathbf{U}_k)^+ \cdot (\mathbf{J}_1 / \lambda_k - \mathbf{J}_1) \\
- (q_k^T \otimes p_k^T \cdot (\mathbf{J}_1 \cdot \mathbf{U}_k)^H / \lambda_k) \cdot (F_{\text{SLS}} \cdot F_{\text{SLS}}^H)^{-1} \cdot W_{R,U}. \tag{33}
\]

Finally, from (32) we can compute the MSE under the assumption that the elements of the noise vector are zero mean, i.i.d., and circularly symmetric. Proceeding like in [6], we obtain the following closed-form MSE expression
\[
E \left\{ (\Delta \mu_{k,\text{SLS}})^2 \right\} = \frac{\sigma_n^2}{2} \cdot \left\| W_{U,N}^T \cdot r_{k,\text{SLS}} \right\|_2^2. \tag{34}
\]

The expressions for the estimation error (10), (30) as well as the MSEs (15), (34) have been derived for 1-D standard ESPRIT. However, as shown in [6], the corresponding derivations for 1-D Unitary ESPRIT are very similar, requiring only minor adaptations. This is still true for 1-D Unitary ESPRIT with SLS. Equations (10), (30) remain the same except that all quantities in it are not computed from $X_0$ but from the forward-backward-averaged matrix $Z_0$ given by
\[
Z_0 = [X_0, \Pi \cdot X_0^* \cdot \Pi] \in \mathbb{C}^{M \times 2N}. \tag{35}
\]

On the other hand, the MSE expressions (15) and (34) have to be slightly extended. For 1-D Unitary ESPRIT with SLS we obtain
\[
E \left\{ (\Delta \mu_{k,\text{SLS}})^2 \right\} = \frac{\sigma_n^2}{2} \cdot \left\| W_{U,N}^T \cdot \mathbf{r}_{k,\text{SLS}} \right\|_2^2 - \Re \left\{ W_{U,N}^T \cdot \mathbf{r}_{k,\text{SLS}} \cdot \Pi_{2MN} \cdot W_{U,N}^T \cdot \mathbf{r}_{k,\text{SLS}} \right\} \tag{36}
\]
where $W_{U,N}$ and $\mathbf{r}_{k,\text{SLS}}$ are computed like $W_{U,N}$ and $\mathbf{r}_k$ replacing $X_0$ by $Z_0$. Similarly, the MSE for 1-D Unitary ESPRIT with SLS is obtained by modifying (34) in the same manner as in (36).

![Fig. 1. Analytical and empirical MSEs vs. the SNR for $M = 12$, $N = 10$, $d = 3$ correlated sources ($\rho = 0.99$) at $\mu_1 = 1$, $\mu_2 = 0$, $\mu_3 = -1$.](image-url)
4. SIMULATION RESULTS

To demonstrate the validity of our results, we have performed Monte Carlo simulations to compare the empirical estimation errors with the results found analytically. In Figure 1, we consider $d = 3$ sources with complex Gaussian distributed source symbols and a pair-wise correlation between sources of $\rho = 0.99$. Moreover, a $M = 12$ element ULA is used to collect $N = 10$ snapshots. The sources’ true positions are given by $\mu_1 = 1, \mu_2 = 0, \mu_3 = -1$. From standard ESPRIT (SE) and Unitary ESPRIT (UE) with LS and SLS we display the empirical estimation error found by averaging over the Monte Carlo trials (“emp”) with the two types of analytical MSE expressions averaged over zero mean circularly symmetric white noise. Simulation results have demonstrated the validity of our asymptotic results. The improvement due to SLS is particularly pronounced for correlated sources and a large number of sensors. However, as we have shown, even cases where the performance difference between LS and SLS is only minor can be reliably predicted by the analytical MSE expressions.

All simulation results show that our analytical expressions agree well with the empirical results.

In Figure 3 we consider $d = 4$ uncorrelated sources, $M = 8$ antenna elements, and $N = 3$ snapshots. The sources are positioned at $\mu_1 = 1.0, \mu_2 = 0.7, \mu_3 = -0.6, \mu_4 = -0.3$. Note that since $N < d$, standard ESPRIT cannot be applied and hence we show Unitary ESPRIT only. In this scenario, the improvement due to SLS is only small. This shows that even minor differences in the performance of different algorithms can be reliably predicted by the analytical MSE expressions.

All simulation results show that our analytical expressions agree well with the empirical results for medium to high SNRs, even in the case of a single source or a small number of snapshots $N$.

5. CONCLUSIONS

In this paper, we have derived the analytical performance of 1-D standard ESPRIT and 1-D Unitary ESPRIT using SLS to solve the shift invariance equations. We have provided the estimation error as an explicit function of the noise realization, which requires no assumption about the noise statistics, as well as the closed-form MSE expressions averaged over zero mean circularly symmetric white noise. Simulation results have demonstrated the validity of our asymptotic results. The improvement due to SLS is particularly pronounced for correlated sources and a large number of sensors. However, as we have shown, even cases where the performance difference between LS and SLS is only minor can be reliably predicted by the analytical MSE expressions.

REFERENCES