

On estimating parameters of a multi-component Chirp Model with equal chirp rates

Abhineek Shukla, Debasis Kundu, Amit Mitra, and Rhythm Grover

Abstract—Multi-component chirp signal models with equal chirp rates appear in various radar applications, e.g., synthetic aperture radar, echo signal of a rapid mobile target, etc. Many sub-optimal estimators have been developed for such models, however, these suffer from the problem of either identifiability or error propagation effect. In this paper, we have developed theoretical properties of the least squares estimators (LSEs) of the parameters of multi-component chirp model with equal chirp rates, where the model is contaminated with linear stationary errors. We also propose two computationally efficient estimators as alternative to LSEs, namely sequential combined estimators and sequential plugin estimators. Strong consistency and asymptotic normality of these estimators have been derived. Interestingly, it is observed that sequential combined estimator of the chirp rate parameter is asymptotically efficient. Extensive numerical simulations have been performed, which validate satisfactory computational and theoretical performance of all three estimators. We have also analysed a simulated radar data with the help of our proposed estimators of multi-component chirp model with equal chirp rates, which performs efficiently in recovery of inverse synthetic aperture radar (ISAR) image of a target from a noisy data.

Keywords: Consistency, asymptotic normality, equal chirp rates, multi-component chirps, stationary linear process, least squares estimators, sequential procedures.

I. INTRODUCTION

In this paper, we have considered the following real valued multi-component chirp model with equal chirp rates parameters:

$$y(n) = \sum_{k=1}^p \left(A_k^0 \cos(\alpha_k^0 n + \beta^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta^0 n^2) \right) + X(n), \quad n = 1, 2, 3, \dots, N, \quad (1)$$

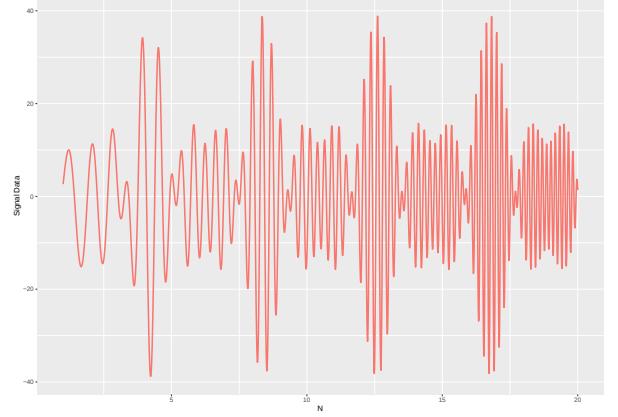
$$\text{where, } \theta^0 = (\mathbf{A}^{0\top}, \mathbf{\xi}^{0\top})^\top = \underbrace{(A_1^0, B_1^0, A_2^0, B_2^0, \dots, A_p^0, B_p^0)}_{\mathbf{A}^\top}, \underbrace{(\alpha_1^0, \alpha_2^0, \dots, \alpha_p^0, \beta^0)}_{\mathbf{\xi}^\top}$$

is the parameter vector, A_k^0, B_k^0 are the amplitude parameters, α_k^0 is the frequency parameter, β^0 is the chirp rate parameter of the k^{th} component, $k = 1, 2, 3, \dots, p$ and $X(n)$ is a noise term. We present Fig.1 for illustration of chirp waveform of the data generated from model (1) for two different values of $p = 5, 10$.

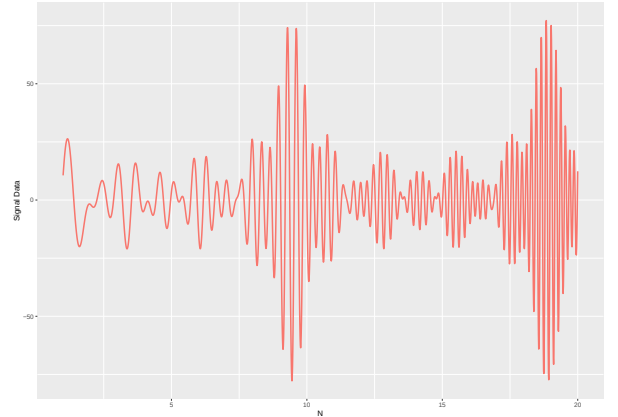
A. Shukla, D. Kundu and A. Mitra are with Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208016, India.

R. Grover is with Mehta Family School of Data Science and Artificial Intelligence, Indian Institute of Technology Guwahati, Assam-781039, India.

(Corresponding author: Abhineek Shukla, email: abhushukla@gmail.com)



(a) Five-component chirp model with equal chirp rates.



(b) Ten-component chirp model with equal chirp rates.

Fig. 1: Chirp waveform of a multi-component chirp model with equal chirp rates.

Consider the problem of radar tracking of moving targets, where translational motion of the target has been compensated, it has been observed in [18] that the Doppler phase of the signal reflected from the scatterer to be a quadratic function of time. Then, the cross-range signal in each range bin after slant range migration correction (see [18]) can be regarded as multi-component chirp model with a common unknown chirp rate. Another application where the model (1) emerges is, when a chirp signal is transmitted from a radar or a communication system (e.g., linear frequency modulated signal (LFM)) and passes through

a linear finite impulse response (FIR) channel (e.g., with multipath propagation), a multi-component chirp model emerges at the output with each component having identical highest order coefficient, which means that the filter output is composed of multi-component chirp model sharing the same highest order coefficient, i.e., chirp rate for quadratic phase, see [2] for more details. Multi-component chirp model with equal chirp rates also appear in studying high resolution radar imaging of the Earth from a satellite or an aircraft ([2]). It has been shown through analysis in [1], that the echo signal from each scatterer of fast-moving target is approximately a multi-component chirp signal with the equal chirp rates. It has also been concluded in [9] that the ISAR received signal from the moving target, which has many scatterers, is approximately a multi-component chirp signal with the equal chirp rates. Similar applications can be observed for an uncooperative target with multi-scatterers, the echo signals after heterodyne detection can be considered as multi-component LFM signals with the equal chirp rates,[20]. Some other applications mentioned in the literature are, the echo data in fast-time dimension has the form of multi-component chirp signals echo of high speed targets, [22], [26] etc. Therefore, model (1) has vast applications in radar and engineering domains and hence studying efficient estimation problem would be of considerable interest to the aerospace and engineering audience. References of such models may be found in [3], [12] and [11].

Substantial literature is available on parameter estimation of multi-component chirp models with distinct chirp rates (e.g., see [23], [24], [29], [10], [14], [27] and the references cited therein). On the other hand, a few works in the literature are dedicated for multi-component chirp models with equal chirp rates (e.g., see [2] and [22]). LSEs would be the most natural choice to obtain optimal estimators of parameters of model (1). However, computing LSEs would require solving a $(p + 1)$ dimensional non-linear optimization problem, which will be computationally prohibitive in practice. Hence, there is a need to develop computationally efficient methods with performance matching with the optimal LSEs. The model (1) is a sub-model of multi-component chirp signal with distinct chirp rates. Therefore, one may think of implementing efficient procedures proposed for multi-component chirp signal with distinct chirp rates, for the model (1). One such example is sequential least squares (see [10]), which is shown to be computationally simpler than LSEs, and also have same asymptotic distribution as that of LSEs. However, since chirp rates (β^0) of different components are same for (1), it loses the orthogonality structure, and hence it becomes a challenging and non-trivial problem to find a computationally efficient estimation procedure. Therefore, it is not immediate how the sequential estimators will behave both theoretically as well as empirically.

Sub-optimal estimators like product high order ambiguity function (PHAF) and integrated cubic phase function (ICPF) of [24] are implemented in [2] and [22], respectively, for the model (1). PHAF provides sub-optimal estimator while ICPF based estimators are computationally burdensome. However, theoretical analysis of PHAF and ICPF based estimators is quite challenging for a multi-component model, because of the interferences and/or interactions among different components and noise. These interferences occur due to non-linear operations of these methods on observed data (see [2] and [24]). Also, de-chirping is performed to obtain estimators of lower order coefficients parameters. This severely reduces the performance of estimators of lower order coefficients, caused by the error propagation from highest order coefficient estimates. The aim of our paper is two-fold:

- derive theoretical properties of least squares estimators for the model parameters of (1);
- propose computationally and asymptotically efficient estimators of the chirp rate parameters, as an alternative to LSEs.

We have considered a model which is of significant importance in radar applications of tracking mobile target. Our proposed estimation methods are motivated from the idea of reducing computational complexity of least squares estimators while, maintaining the optimal rates of convergence of the estimators. Combined estimator is based on the principle of improving estimators of chirp rate based on sequential process, while, plugin estimator is based on minimizing the computational load of LSEs to the least. Technical significance of our proposed estimators lies in obtaining same rates of convergence as that of LSEs along with the benefit of reducing computational complexity. We have also shown that combined estimator of chirp rate parameter to be asymptotically optimal. We have further analysed a radar data available in [16] and shown that our proposed estimators can be efficiently used to reconstruct the ISAR image of a noisy radar data.

The rest of the paper is organised as follows: we first describe the assumptions for motivating the estimation methodology of all the three proposed estimators in Section II. We also provide detailed estimation methodology and theoretical results of the estimators in the same Section II. Detailed numerical simulations are executed in Section III to study the performance of proposed estimators and also comparing the performance with ICPF, in terms of mean squared errors (MSEs) and computational time. We have also proposed some implementation procedures using PHAF based estimators, to overcome the computational challenge with the two-dimensional (2D) optimization problem associated with proposed estimators, along with its careful assessment on one radar data. The paper is concluded in Section IV followed by the appendices containing proofs of the theorems.

II. ESTIMATION METHODOLOGY

We first mention the assumptions on noise term and parameters of the model (1), required to motivate the layout of the proposed methodology, as well as to derive the asymptotic theoretical properties of all the three estimators.

A. Assumptions

Assumption 1. $X(n)$ is stationary linear process with the following form:

$$X(n) = \sum_{j=-\infty}^{\infty} a(j)\epsilon(n-j), \quad (2)$$

such that the coefficients $a(j)$ are absolutely summable, i.e.,

$$\sum_{j=-\infty}^{\infty} |a(j)| < \infty, \quad (3)$$

and $\{\epsilon(n)\}$ is a sequence of independently and identically distributed (i.i.d.) random variables with mean 0, variance σ^2 and finite fourth moment.

Assumption 2. True value of parameters $A_k^0, B_k^0 \in [-M, M]$ for some $M > 0$, $\alpha_k^0 \in (0, 2\pi)$, $\beta^0 \in (0, \pi/2)$ and $A_k^{0^2} + B_k^{0^2} > 0$ for $k = 1, 2, 3, \dots, p$.

Assumption 3. $A_1^{0^2} + B_1^{0^2} > A_2^{0^2} + B_2^{0^2} > \dots > A_p^{0^2} + B_p^{0^2}$.

Assumption 1 defines a large class of errors in the sense that it contains the usual assumption of i.i.d. Gaussian errors as well as moving average processes, autoregressive processes etc., and hence, including the dependent structure too. Assumption 2 is a standard assumption, which mainly indicates that the frequencies and frequency rate parameters are not on the boundary, and at the same time, it avoids any identifiability problem in possible values of the parameter space. Assumption 3 helps in the motivation of developing a natural sequential procedure and also guarantees the consistency of the estimators obtained through this sequential process. Assumption 3 also guides sequential procedure in identifying the different components in correct order.

Next, we describe the estimation methodologies for LSEs, sequential combined estimators and sequential plugin estimators of the parameters of model (1). These methods are based on the assumption that the number of components p is known apriori, for the model (1).

B. LSEs

Let us consider $\mathbf{Y} = (y(1), y(2), \dots, y(N))^T$ to be the data vector and define the following sum of squares:

$$Q(\boldsymbol{\theta}) = (\mathbf{Y} - \mathbf{W}(\boldsymbol{\xi})\mathbf{A})^T (\mathbf{Y} - \mathbf{W}(\boldsymbol{\xi})\mathbf{A}),$$

where, $\boldsymbol{\theta} = (\mathbf{A}^T, \boldsymbol{\xi}^T)^T$, $(n, 2k-1)^{th}$ and $(n, 2k)^{th}$ elements of the $N \times 2p$ order matrix $\mathbf{W}(\boldsymbol{\xi})$ are $\cos(\alpha_k n + \beta n^2)$

and $\sin(\alpha_k n + \beta n^2)$ respectively. Then, LSEs of $\boldsymbol{\theta}^0$ will be obtained as follows:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}). \quad (4)$$

Note that using separable linear regression technique, above optimization problem can be reduced to a $(p+1)$ dimensional from a $(3p+1)$ dimensional non-linear optimization problem (see e.g., [10] and [19]). So, after simplification, above estimators can be obtained equivalently as

$$\hat{\boldsymbol{\xi}} = \arg \min_{\boldsymbol{\xi}} (\mathbf{Y} - \mathbf{W}(\boldsymbol{\xi})\mathbf{A}(\boldsymbol{\xi}))^T (\mathbf{Y} - \mathbf{W}(\boldsymbol{\xi})\mathbf{A}(\boldsymbol{\xi})),$$

where, $\mathbf{A}(\boldsymbol{\xi}) = (\mathbf{W}(\boldsymbol{\xi})^T \mathbf{W}(\boldsymbol{\xi}))^{-1} \mathbf{W}(\boldsymbol{\xi})^T \mathbf{Y}$, for more details, one can refer [7]. Here, estimate of $\mathbf{A}$ is given by:

$$\hat{\mathbf{A}} = (\mathbf{W}(\hat{\boldsymbol{\xi}})^T \mathbf{W}(\hat{\boldsymbol{\xi}}))^{-1} \mathbf{W}(\hat{\boldsymbol{\xi}})^T \mathbf{Y}.$$

C. Sequential Combined Estimators

We use the idea of sequential estimation procedure [10] to find computationally simpler methods. This sequential estimator is based on estimating only one component at a time, hence requires only a 2D optimization problem to solve p number of times. We now define the sequential combined estimators of the parameters of model (1) in following steps:

- **Step-1** Consider the following sum of squares:

$$Q_1(\boldsymbol{\theta}_1) = \sum_{n=1}^N (y(n) - A_1 \cos(\alpha_1 n + \beta_1 n^2) - B_1 \sin(\alpha_1 n + \beta_1 n^2))^2.$$

First component parameter vector is $\boldsymbol{\theta}_1^0 = (A_1^0, B_1^0, \alpha_1^0, \beta_1^0)^T$ and let $\boldsymbol{\theta}_1 = (A_1, B_1, \alpha_1, \beta_1)^T$, then we define sequential combined estimator of the first component parameters as follows:

$$\check{\boldsymbol{\theta}}_1 = \arg \min_{\boldsymbol{\theta}_1} Q_1(\boldsymbol{\theta}_1). \quad (5)$$

Note that above (5) can be reduced to a 2D optimization problem by using separable linear regression technique (see [19]). We can simplify (5) similar to that of (4). It will also be a 2D optimization problem for the remaining component estimators.

- **Step-2** We eliminate the effect of first component from the data and then define sequential combined estimator $\check{\boldsymbol{\theta}}_2 = (\check{A}_2, \check{B}_2, \check{\alpha}_2, \check{\beta}_2)^T$ of second component parameter vector $\boldsymbol{\theta}_2^0 = (A_2^0, B_2^0, \alpha_2^0, \beta_2^0)^T$ by updating data as:

$$\check{y}_1(n) = y(n) - \check{A}_1 \cos(\check{\alpha}_1 n + \check{\beta}_1 n^2) - \check{B}_1 \sin(\check{\alpha}_1 n + \check{\beta}_1 n^2).$$

Consider,

$$Q_2(\boldsymbol{\theta}_2) = \sum_{n=1}^N (\check{y}_1(n) - A_2 \cos(\alpha_2 n + \beta_2 n^2) - B_2 \sin(\alpha_2 n + \beta_2 n^2))^2,$$

where $\theta_2 = (A_2, B_2, \alpha_2, \beta_2)^\top$, then the second component estimator is given by the following equation:

$$\check{\theta}_2 = \arg \min_{\theta_2} Q_2(\theta_2). \quad (6)$$

We can simplify (6) similar to that of (4).

- **Step-3** We repeat step-2 for the remaining components' sequential combined estimators, i.e. $\check{\theta}_k = (\check{A}_k, \check{B}_k, \check{\alpha}_k, \check{\beta}_k)^\top$, $k \geq 3$.
- **Step-4** Since all of the different components in model (1) have equal chirp rates parameters, so we propose combine estimators of β^0 from each component in an optimal manner. It is known that, when the chirp rates are different, then estimators of distinct components are asymptotically independent, (see [10]). Therefore, using this result as motivation, we propose sequential combined estimator $\check{\beta}$ of β^0 by minimizing the variance of linear combination of all component estimators $\check{\beta}_k$ among the following class:

$$\check{\beta} = \sum_{k=1}^p l_k \check{\beta}_k, \quad \sum_{k=1}^p l_k = 1.$$

Sequential combined estimator of chirp rate parameter β^0 is therefore, given by:

$$\check{\beta} = \frac{1}{\sum_{k=1}^p (\check{A}_k^{0^2} + \check{B}_k^{0^2})} \sum_{k=1}^p (\check{A}_k^{0^2} + \check{B}_k^{0^2}) \check{\beta}_k.$$

We have shown that $\check{\beta}$ is an asymptotically optimal estimator of chirp rate parameter β^0 . Finally, $\check{\theta} = (\check{A}_1, \check{B}_1, \check{A}_2, \check{B}_2, \dots, \check{A}_p, \check{B}_p, \check{\alpha}_1, \check{\alpha}_2, \dots, \check{\alpha}_p, \check{\beta})^\top$ is the proposed sequential combined estimator of parameter $\theta^0 = (A_1^0, B_1^0, A_2^0, B_2^0, \dots, A_p^0, B_p^0, \alpha_1^0, \alpha_2^0, \dots, \alpha_p^0, \beta^0)^\top$.

D. Sequential Plugin Estimators

It is possible that the number of components in model (1) may be very large for many real-life applications. Sequential combined estimators will require p number of 2D optimization problems to solve. In this situation, implementing sequential combined estimators will involve a tedious computation. So, to resolve this issue, we perform 2D optimization only for the first component estimators and then, we plugin this estimate of β^0 in all remaining components. Hence the sequential plugin estimator will require one 2D and $(p-1)$ number of 1D optimization problems to solve, reducing the computational cost significantly. We now describe the detailed methodology of sequential plugin estimators of parameters of model (1) in following steps:

- **Step-I** Define following sum of squares to estimate first

component parameters of model (1):

$$Q_1(\bar{\theta}_1) = \sum_{n=1}^N (y(n) - A_1 \cos(\alpha_1 n + \beta n^2) - B_1 \sin(\alpha_1 n + \beta n^2))^2.$$

Further define $\bar{\theta}_1 = (A_1, B_1, \alpha_1, \beta)^\top$, then we estimate the first component parameters $\bar{\theta}_1^0 = (A_1^0, B_1^0, \alpha_1^0, \beta^0)^\top$ as follows:

$$\tilde{\theta}_1 = \arg \min_{\theta_1} Q_1(\bar{\theta}_1). \quad (7)$$

We can simplify (7) similar to that of (4). Equation (7) can be reduced to a 2D optimization problem. But for other components, it will be a 1D optimization problem (see (8)). It can be observed that sequential combined estimators and sequential plugin estimators of the first component parameters are same.

- **Step-II** To compute second component estimators $\tilde{\theta}_2 = (\tilde{A}_2, \tilde{B}_2, \tilde{\alpha}_2)^\top$ of $\bar{\theta}_2^0 = (A_2^0, B_2^0, \alpha_2^0)^\top$, we update the data as:

$$\tilde{y}_1(n) = y(n) - \tilde{A}_1 \cos(\tilde{\alpha}_1 n + \tilde{\beta} n^2) - \tilde{B}_1 \sin(\tilde{\alpha}_1 n + \tilde{\beta} n^2),$$

and define following sum of squares:

$$Q_2(\bar{\theta}_2) = \sum_{n=1}^N (\tilde{y}_1(n) - A_2 \cos(\alpha_2 n + \tilde{\beta} n^2) - B_2 \sin(\alpha_2 n + \tilde{\beta} n^2))^2,$$

where $\bar{\theta}_2 = (A_2, B_2, \alpha_2)^\top$. Second component sequential combined estimators are given by:

$$\bar{\theta}_2 = \arg \min_{\theta_2} Q_2(\bar{\theta}_2). \quad (8)$$

- **Step-III** Repeat step-II for remaining component parameters i.e., $\bar{\theta}_k^0 = (A_k^0, B_k^0, \alpha_k^0)$ for $k \geq 3$.

In this manner, we obtain sequential plugin estimator $\tilde{\theta} = (\tilde{A}_1, \tilde{B}_1, \tilde{A}_2, \tilde{B}_2, \dots, \tilde{A}_p, \tilde{B}_p, \tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_p, \tilde{\beta})^\top$ of the parameter $\theta^0 = (A_1^0, B_1^0, A_2^0, B_2^0, \dots, A_p^0, B_p^0, \alpha_1^0, \alpha_2^0, \dots, \alpha_p^0, \beta^0)^\top$. We emphasize that β^0 is estimated only once in sequential plugin method as compared to the case of sequential combined, where β^0 is estimated repeatedly for all different component parameters.

E. Main Theorems

Theorem 1. Estimators $\hat{\theta}$, $\check{\theta}$ and $\tilde{\theta}$ are strongly consistent for the parameter θ^0 of model (1) under the Assumption 1, 2 and 3, i.e., $\hat{\theta} \xrightarrow{a.s.} \theta^0$, $\check{\theta} \xrightarrow{a.s.} \theta^0$, $\tilde{\theta} \xrightarrow{a.s.} \theta^0$ as $N \rightarrow \infty$. Here, a.s. represents almost sure convergence.

Proof. Please refer Appendix A for the proof. $\square$

Theorem 2. Estimators $\hat{\theta}$, $\check{\theta}$ and $\tilde{\theta}$ follow asymptotically normal distributions under the Assumption 1, 2 and 3, with the scaling matrix

$$D^{-1} = \text{diag}(\underbrace{N^{1/2}, N^{1/2}, \dots, N^{1/2}, N^{1/2}}_{2p \text{ times}}, \underbrace{N^{3/2}, \dots, N^{3/2}}_p, N^{5/2}),$$

i.e.,

$$D^{-1}(\hat{\theta} - \theta^0) \xrightarrow{d} \mathcal{N}_{3p+1}(\mathbf{0}, 2c\sigma^2 \Sigma_1),$$

$$D^{-1}(\tilde{\theta} - \theta^0) \xrightarrow{d} \mathcal{N}_{3p+1}(\mathbf{0}, 2c\sigma^2 \Sigma_2) \text{ and}$$

$$D^{-1}(\tilde{\theta} - \theta^0) \xrightarrow{d} \mathcal{N}_{3p+1}(\mathbf{0}, 2c\sigma^2 \Sigma_3),$$

where $\text{diag}(a_1, a_2, \dots, a_k)$ represents $k \times k$ diagonal matrix with diagonal entries $a_1, a_2, \dots, a_k$. Diagonal elements of positive definite matrices Σ_1, Σ_2 and Σ_3 are described in equations (11)-(17) of Appendix B.

Proof. Please refer Appendix B for the proof. $\square$

Results similar to Theorem 1 and 2 exist for the estimators already exist for the estimators of parameters of multi-component chirp model with distinct chirp rates. However, the estimators of multi-component chirp model with equal chirp rates needed separate attention because in this case, the usual asymptotic orthogonality of different chirp components is lost. The asymptotic variance-covariance matrix of the estimators for equal chirp rates is different from the distinct chirp rates case. So, the novelty of our work lies in deriving separately all the properties of the estimators, which do not follow trivially from the properties of estimators developed for distinct chirp rates case.

Remark: If $X(n)$ is an i.i.d. sequence following normal distribution with mean 0 and variance σ^2 in Theorem 2, then diagonal entries of Σ_1 will provide the corresponding Cramer-Rao lower bound. We have reported asymptotic variance of LSEs of α_k^0 and β^0 in (11) and (12). It has been shown that sequential combined estimators of chirp rate parameter β^0 is asymptotically optimal, however, the asymptotic variance of sequential plugin estimator of frequency parameter α_k^0 is less than that of sequential combined estimator (see comparison in page 3 of the supplementary material). Therefore, $\Sigma_3 - \Sigma_2$ is neither positive definite, nor negative definite. These theoretical results are also validated by numerical simulations in the next section.

III. NUMERICAL SIMULATIONS

We have performed extensive numerical simulations to study the behaviour of the different proposed methods based on their MSEs, and also about the time needed to obtain these estimators. Numerical simulations have been categorized broadly into five subsections. We compare the finite sample performance of LSEs, sequential plugin estimators and sequential combined estimators of the non-linear parameters, in terms of MSEs, for a five-component chirp model with equal chirp rates in the subsection III-A. We have compared our proposed estimators with ICPF based estimator of chirp rate parameter in subsection III-B. Then, we have presented computational cost in terms of time to obtain these estimators in subsection III-C. We also provide some implementations and performance of the estimators in subsection III-D to reduce computational problem for the proposed estimators. Sub-section III-E includes assessment of the proposed implementation procedure using one radar

signal data. Throughout the numerical simulations, we have used analytical form of the model (1) for comparison, refer equation (1) of [2]. We have also considered the following two different kinds of error sequence $X(n)$ as:

- **E1:** $X(n)$ is a sequence of independently and identically distributed (i.i.d.) Gaussian errors with mean 0 and variance σ^2 ;
- **E2:** $X(n)$ is autoregressive moving average (ARMA) errors with following representation:

$$X(n) = 0.6X(n-1) + \epsilon(n) + 0.1\epsilon(n-1). \quad (9)$$

where $\epsilon(n)$ is a sequence of i.i.d. Gaussian random variables with mean 0 and variance σ^2 .

The real and imaginary parts are contaminated independently with the above two kind of errors E1 and E2. Parameters of the model for the subsections III-A, III-B and III-C are as follows:

$$\begin{aligned} p = 5, \quad A_1^0 = 3.35, \quad A_2^0 = 2.8, \quad A_3^0 = 2.1, \quad A_4^0 = 1.59, \\ A_5^0 = 0.9, \quad \alpha_1^0 = 0.89, \quad \alpha_2^0 = 0.96, \quad \alpha_3^0 = 0.76, \\ \alpha_4^0 = 0.56, \quad \alpha_5^0 = 0.37, \quad \beta^0 = 0.87. \end{aligned} \quad (10)$$

A. Comparing MSEs of LSEs with sequential combined and sequential plugin estimators

For the analytical model (1) of [2], we have studied finite sample performance of LSEs, sequential combined estimators and sequential plugin estimators in the sense of MSEs. We have taken different sets of sample sizes $N = 100, 200, \dots, 1000$. Both E1 and E2 are considered for error $X(n)$ with $\sigma = 2, 3$. For a generated data-set, we have obtained estimates of LSEs, sequential combined, and sequential plugin estimators. All of these estimators do not have any explicit form, so we have to resort to some standard numerical procedures. Here, we have implemented the Nelder-Mead simplex algorithm [15] (using “optim” function in R software) for optimization of the associated objective function to obtain all these estimators. We also need good initial guess for convergence of the Nelder-Mead algorithm to the true parameters. So, we have used grid-search on the intervals $[\alpha_k^0 - 1/N, \alpha_k^0 + 1/N]$ for $k = 1, 2, 3, 4, 5$ and $[\beta^0 - 1/N^2, \beta^0 + 1/N^2]$ to obtain initial guess.

Now, we calculate squared errors between the estimates and true values of the parameters. We then repeat this upto 10,000 number of times and find the average value of squared errors corresponding to each estimators, which yields MSEs. Plots of MSEs with respect to different sample sizes are provided in Fig. 2 for the case $\sigma = 2$. Similar results obtained for $\sigma = 3$ are provided in page 1, Figure 1 of the supplementary material.

Sub-optimal estimators usually suffer from a relatively higher signal to noise ratio (SNR) threshold. Therefore, we have also evaluated MSEs of estimators over 10,000 replications, with respect to 31 different values of SNR

ranging from -5 to 10. Errors are sampled randomly from normal distribution with mean 0 and variance according to the value of SNR. Obtained plot is presented in Fig. 3. A list of observations from these figures can be summarised as follows:

- MSEs of all the three estimators decrease as sample size increases, hence validating their consistency as stated in Theorem 1 (Fig. 2);
- MSEs of all the three estimators decrease as SNR increases (Fig. 3);
- If we consider chirp rate parameter β^0 , the sequential plugin estimator has more stable MSEs as compared with that of sequential combined estimator (Fig. 2);
- SNR threshold for all the three estimators seems to be shifting towards higher value of SNR as the number of components increase, e.g., SNR threshold for all the three estimators of α_3^0 is around -5 dB, whereas it is around 2.5 dB for α_5^0 . It also suggests that sequential procedures for models of type (1), perform as good as if we use LSEs in terms of SNR threshold, e.g., see performance of estimators of α_5^0 in Fig. 3;
- MSEs of sequential plugin estimator of the frequency parameters $\alpha_k^0, k = 2, 3, 4, 5$ are comparable with that of LSEs (Fig. 2, Fig. 3);
- Sequential combined estimator of chirp rate parameter β^0 performs at par with the LSEs for high value of SNRs (Fig. 3);
- Fig. 2 also validates the rate of convergence of obtained estimators, which is provided in Theorem 2. .

B. Comparison with ICPF

Implementation of ICPF to estimate chirp rate parameters is shown in [22]. We have compared MSEs of LSE, sequential combined estimator, sequential plugin estimator and ICPF based estimator of chirp rate parameter β^0 for the same model considered in (10). We have chosen 30 different values of SNR to compare their performances. MSEs are obtained over 10,000 replications. Simulated results are presented in Fig. 4, from which we can conclude the following points:

- Sequential combined estimator performs at par with LSEs, and better than the other computationally efficient estimators, ICPF based estimator and sequential plugin estimator for SNR higher than 2.5 dB;
- Sequential plugin estimator has stable MSEs than that of ICPF and sequential combined estimator.

Further, we have added Table 1 in supplementary material, for highlighting quantitative technical advantages of the proposed methods.

C. Computation Time Comparison of all estimators

We have further presented time cost associated in estimating all non-linear parameters using LSEs, sequential combined estimators and sequential plugin estimators. We

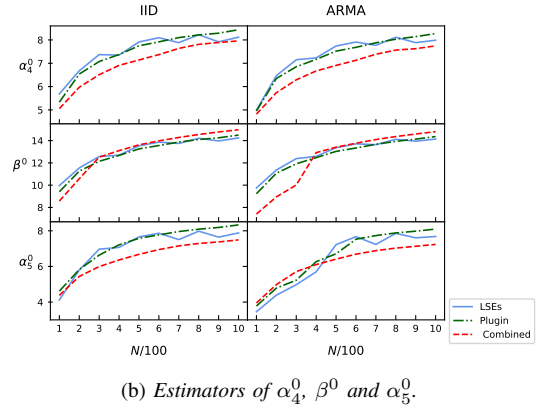
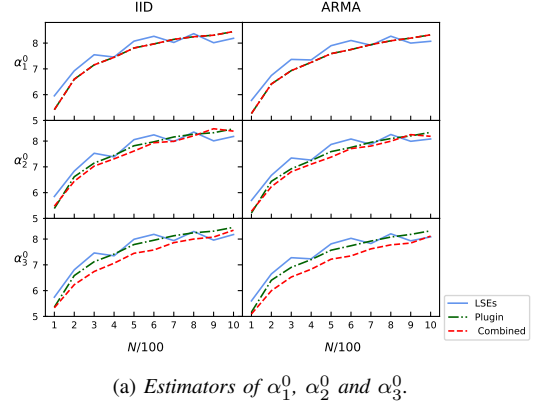
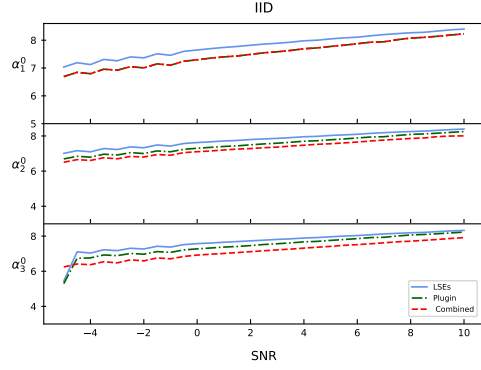


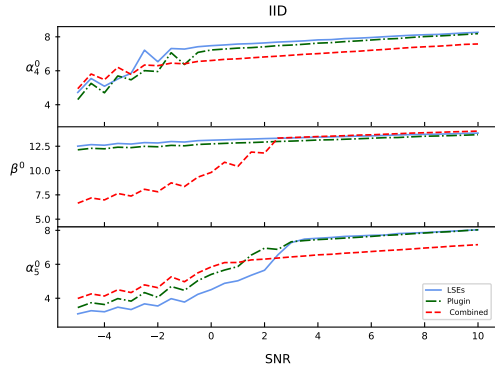
Fig. 2: MSEs (on $-\log_{10}$ scale) of LSEs, sequential plugin estimators (Plugin) and sequential combined estimators (Combined) versus the increasing sample size, for the case $\sigma = 2$.

have applied same approach of choosing initial guess and then applying Nelder-Mead to obtain all the final estimates, as suggested in subsection III-A. Figure 2, reported in the page 3 of the supplementary material, shows the ratio of average time (over 1000 replications) taken to compute the LSEs and sequential combined estimator, with respect to the average time to compute sequential plugin estimator. We conclude (Figure 2 of the supplementary material) that the time taken by LSEs is more than 6 times the time taken by the sequential plugin estimators. Also, the time taken by sequential combined estimator is more than 5 times the time taken by sequential plugin estimator for large sample sizes. We had not chosen ICPF in this comparison, because LSEs and sequential combined estimators would normally take much more time for finding good initial guess than ICPF as ICPF requires only 1D optimization problem to solve, whereas LSEs would require 6D optimization problem and sequential combined estimator would require five times 2D optimization problem to solve, for the set-up (10).

However, we have compared the sequential plugin estimator and ICPF estimator on the basis of time required



(a) Estimators of α_1^0 , α_2^0 and α_3^0 .



(b) Estimators of α_4^0 , β^0 and α_5^0 .

Fig. 3: MSEs (on $-\log_{10}$ scale) of LSEs, sequential plugin estimators (Plugin) and sequential combined estimators (Combined) versus the SNR, for fixed sample size $N = 500$.

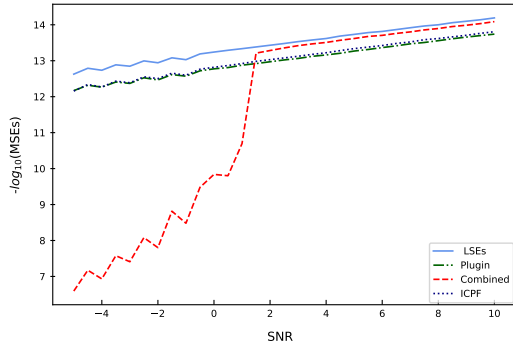


Fig. 4: MSEs (on $-\log_{10}$ scale) of LSEs, sequential plugin estimators (Plugin), sequential combined estimators (Combined) and ICPF based estimator of chirp rate parameter β^0 versus the SNR, for fixed sample size $N = 501$.

to compute the estimates of β^0 . We have used grid-search on the intervals $[\alpha_1^0 - 1/\sqrt{N}, \alpha_1^0 + 1/\sqrt{N}]$ and $[\beta^0 - 1/\sqrt{N}, \beta^0 + 1/\sqrt{N}]$ to obtain initial guess of both estimators for implementing Nelder-Mead algorithm. The motivation to use this interval is that we can always use high order ambiguity function (HAF) based estimators to find initial guess of β^0, α_1^0 and then perform grid-search on the interval of length $1/\sqrt{N}$ around the HAF estimates. We have presented time comparison between our most computationally efficient estimator (sequential plugin estimator) and ICPF based estimator. Here, average time (in seconds) is reported over 50 replications for both estimators. We can see clearly from the Table I that sequential plugin outperforms the ICPF estimator with a significant margin and hence, faster than it. Note that we have performed all simulations for time comparison in a system with following specifications: Intel(R) Xeon(R) CPU E7-8870 v3 @ 2.10GHz.

D. Implementation using PHAF

Our most computationally efficient estimator is the sequential plugin estimator, which requires 2D optimization problem to be solved for first component estimators and 1D problem for remaining components. 2D optimization will be burdensome when sample size is large. So, we now propose an implementation procedure that requires solving only a 1D optimization problem, but ICPF needs a large number of algebraic operations to perform on the data and hence making it computationally more burdensome as compared to the PHAF. Therefore, PHAF based estimators can help in reducing computational difficulty associated with obtaining good set of initial guesses for the sequential plugin and sequential combined estimators. We first obtain PHAF estimates of the first component estimators, then we provide following three different options to get initial guess for the first component sequential plugin estimators as:

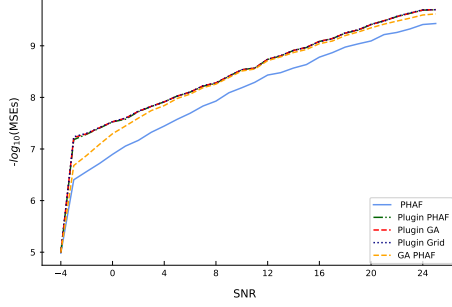
- 1) PHAF;
- 2) Use genetic algorithm (GA) in the $1/(2\pi N)$ neighbourhood of the the PHAF estimates;
- 3) Grid search around the $1/N^{0.9}$ and $1/N^{1.9}$ neighbourhood of PHAF estimators of frequency and frequency rate parameters.

We have used “GA” package [21] from R software to implement GA. Initial population size for GA is fixed at 500, elite count at 20, fitness function is negative of first component error sum of squares. The roulette wheel selection is used to determine parents for next generation of individuals. Probability of crossover is maintained at 0.9 and probability of mutation is set at 0.1. GA is then run for 100 generations.

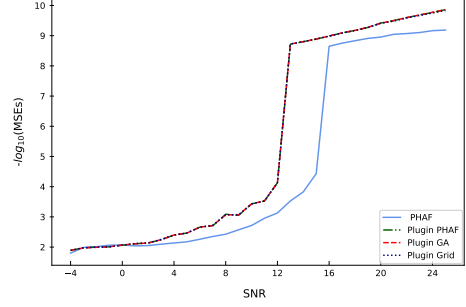
After this step, we are left with only 1D optimization problem to find sequential plugin estimators of frequency parameters of the remaining components, which can be done easily. We have compared MSEs of the sequential

Sample Size	101	121	141	161	181	201	221	241
ICPF	16.07	31.72	53.46	88.25	118.47	181.58	228.66	299.67
Plugin	7.36	13.43	20.7	30.18	36.92	47.91	60.62	74.46

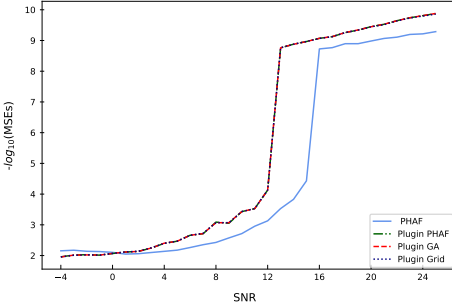
TABLE I: Computational complexity comparison of sequential plugin estimator and ICPF with respect to time (seconds).



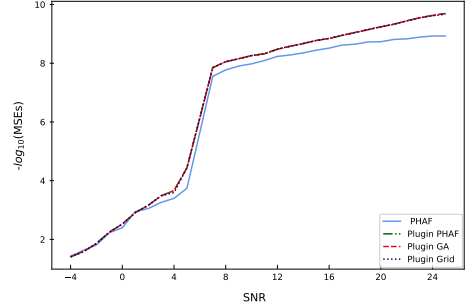
(a) Estimators of α_1^0 .



(a) Estimators of α_3^0 .



(b) Estimators of α_2^0 .



(b) Estimators of α_4^0 .

Fig. 5: MSEs (on $-\log_{10}$ scale) of PHAF estimators, sequential plugin estimators obtained by taking PHAF as initial guess (Plugin PHAF), sequential plugin estimators obtained by using GA as initial guess (Plugin GA), sequential plugin estimators obtained by grid search to get initial guess (Plugin Grid) and GA with PHAF implementation (GA PHAF) versus the SNR for fixed sample size $N = 500$.

Fig. 6: MSEs (on $-\log_{10}$ scale) of PHAF estimators, sequential plugin estimators obtained by taking PHAF as initial guess (Plugin PHAF), sequential plugin estimators obtained by using GA as initial guess (Plugin GA) and sequential plugin estimators obtained by grid search to get initial guess (Plugin Grid) versus the SNR for fixed sample size $N = 500$.

plugin estimators using above implementation, with that of the PHAF estimators across 30 different SNR levels. MSEs are observed over 1000 replications. The sample size is $N = 500$ with the non-linear parameters same as that mentioned in (10). GA algorithm in our case, took 56.6 seconds and grid search took 17.8 seconds to compute initial guess, on an average, over 30,000 replications. From Fig. 5, 6 and 7, it can be easily observed that:

- sequential plugin estimators of the first component parameters perform very well for even low SNR -3 dB;
- sequential plugin estimators for remaining component parameters attain better performance around 12 dB;
- different implementation procedures to find initial guess of the plugin estimators resulted in similar performance.

We suggest to use PHAF directly as initial guess for the data analysis purpose, as it will be the least time consuming out of the proposed three methods and at the same time, it has comparable numerical performance as the remaining two options.

E. Radar Data Analysis

We consider a radar data simulation obtained from [16] (Chapter 4) where, 110 point scatterers were placed to reproduce the outline of a fictitious airplane. Backscattered electric field for 64 different frequencies and 64 different aspects were collected. We further add i.i.d. normal errors with standard deviation 5 to the real and imaginary parts of the radar data, to make problem of estimation more realistic and difficult. The conventional small-bandwidth small-angle

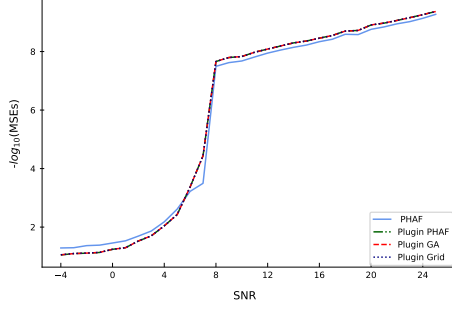
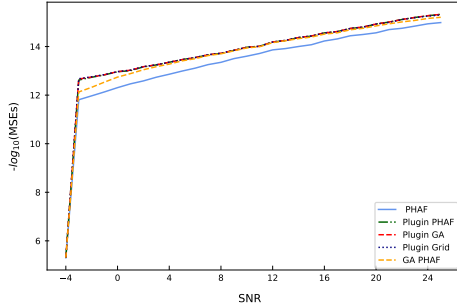
(a) Estimators of α_5^0 .(b) Estimators of β^0 .

Fig. 7: MSEs (on $-\log_{10}$ scale) of PHAF estimators, sequential plugin estimators obtained by taking PHAF as initial guess (Plugin PHAF), sequential plugin estimators obtained by using GA as initial guess (Plugin GA), sequential plugin estimators obtained by grid search to get initial guess (Plugin Grid) and GA with PHAF implementation (GA PHAF) versus the SNR for fixed sample size $N = 500$.

ISAR imaging algorithm is applied to obtain ISAR image of the airplane as shown in Fig. 8 with a certain level of contrast adjusted. Using motivation from [18], we have fitted twenty-component chirp with equal chirp rates model to the backscattered data observed in each range bin cell. After obtaining fitted data from these twenty-component chirp models in each range bins, we apply the conventional small-bandwidth small-angle ISAR imaging algorithm to get the ISAR image at the same contrast level as that of contaminated data Fig. 8 (a).

One can observe from Fig. 8 that using estimators with optimal rates of convergence can satisfactorily help in fitting the data obtained from backscattered electric field after hitting a target, whereas the ISAR image obtained using PHAF in Fig. 8 (d) is distorted and noisy. Combined estimator performs better than the plugin and PHAF based estimators. Note that these plots are presented at the same level of contrast. Our proposed estimators can be combined further with a suitable contrasting algorithm, and then it can be used to improve ISAR imaging even in

more difficult situations. We have proposed our estimators after assuming known number of chirp components, so selecting twenty-component chirp model is subjective and existing methods like Bayesian information criteria (BIC) or maximum a posteriori probability (MAP) based approach fail to estimate appropriate number of components for such kind of radar datasets. We believe that the methods based on bayesian posterior inference methods will be effective in solving such problems. Hence, a comprehensive separate study is required to study estimation of the number of components for such a dataset.

IV. CONCLUSION

We have established theoretical asymptotic properties of three estimators namely the LSEs, sequential combined estimators and sequential plugin estimators of the parameters of a chirp signal model with equal chirp rates. We observe that among the three estimators, the sequential plugin estimator is computationally least burdensome and provides a computationally efficient alternative to the LSEs. Empirical results indicate stable performance of the sequential plugin estimator for the chirp rate parameter. We further observe that sequential combined estimator of chirp rate parameter β^0 is asymptotically optimal, with its asymptotic variance matching with that of the LSEs. Analysis of radar signal data indicate that model (1) can be used satisfactorily to analyse these type of signals and therefore, our proposed estimators have potential to be used further in radar applications. It is observed that fitting radar signal with model (1) performs satisfactorily using our proposed estimators, specially combined estimator performs better than the plugin and PHAF based estimators. Extensive numerical simulations support the derived theoretical asymptotic properties of the proposed estimators. Future work may be directed towards estimating number of chirp signal components present in such radar signal data..

ACKNOWLEDGMENTS

We are thankful to the Associate Editor Dr. Hasan Mir for his encouragement, and the three unknown reviewers for their constructive suggestions. Their suggestions have helped improving the manuscript significantly.

V. APPENDIX A

We provide a very brief sketch of the proof of Theorem 1 and 2.

A. Strong Consistency of LSEs

Proof: Proof of $\hat{\theta} \xrightarrow{a.s.} \theta^0$ follows along the similar lines of proof of Theorem 1 of [8] with appropriate changes. ■

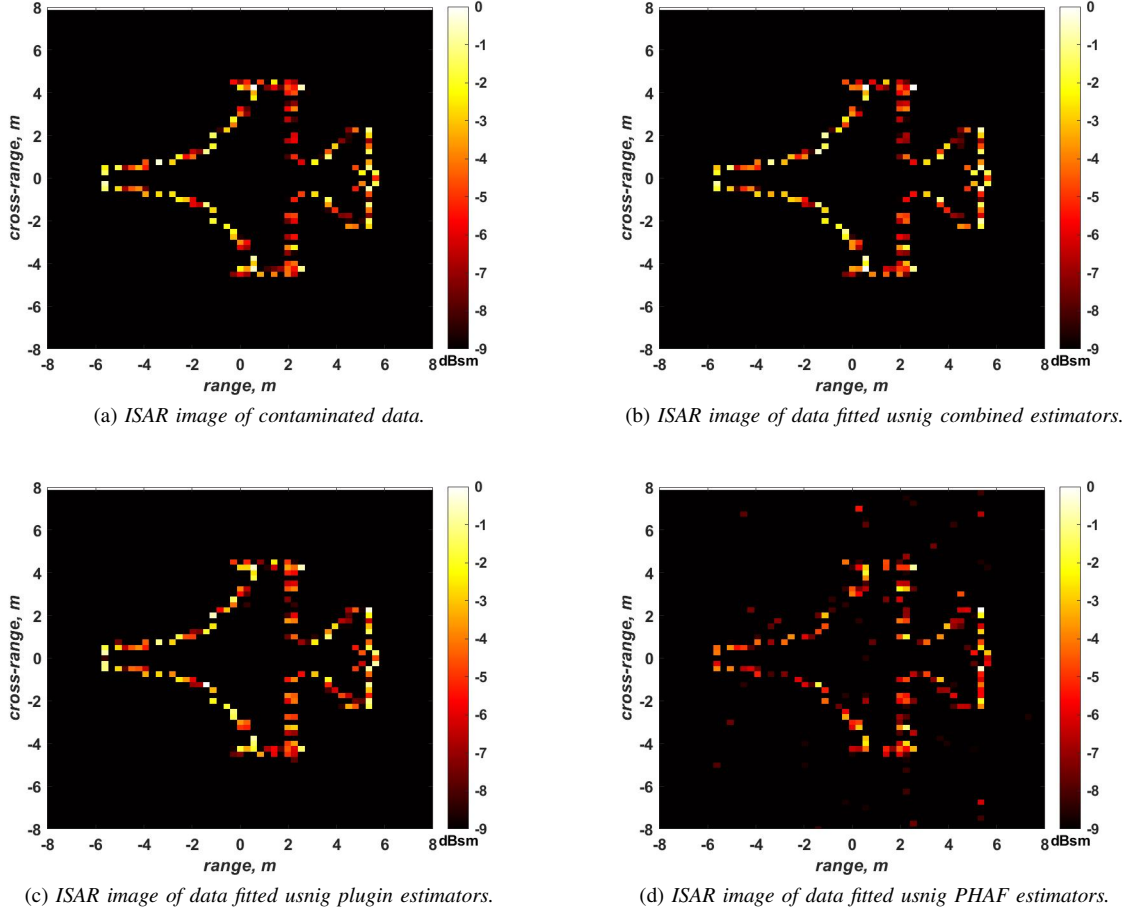


Fig. 8: ISAR image obtained from fitting twenty-component chirp model with equal chirp rates in each range bin signals.

B. Strong Consistency of Sequential Combined Estimators

Proof: Proof of $\check{\theta}_k \xrightarrow{a.s.} \theta_k^0$ follows along the similar lines of proof of Theorem 3, 4 and 5 of [10]. Note that $\check{\beta} = \frac{1}{\sum_{k=1}^p (\check{A}_k^{0^2} + \check{B}_k^{0^2})} \sum_{k=1}^p (\check{A}_k^{0^2} + \check{B}_k^{0^2}) \check{\beta}_k$ is a continuous function of strongly consistent estimators and hence, by continuous mapping theorem, $\check{\beta}$ is strongly consistent for β^0 . ■

C. Strong Consistency of Sequential Plugin Estimators

Proof: We know that first component sequential plugin estimators are same as first component sequential combined estimators, so

$$\tilde{\theta}_1 \xrightarrow{a.s.} (A_1^0, B_1^0, \alpha_1^0, \beta^0)^\top$$

follows from previous argument (see Section V-B). We provide a sketch of proof of consistency for $\bar{\theta}_2$, and strong consistency of $\hat{\theta}_k$ will follow similar to it for $k \geq 3$. First consider the following lemma:

Lemma 1. Consider the set $S_c = \{\bar{\theta}_2 : \bar{\theta}_2 \in \bar{\Theta}_2, |\bar{\theta}_2 - \theta_2^0| \geq c\}$, where $\bar{\theta}_2 = (A_2, B_2, \alpha_2)^\top$, $\theta_2^0 = (A_2^0, B_2^0, \alpha_2^0)^\top$

and $\bar{\Theta}_2 = [-M, M] \times [-M, M] \times (0, 2\pi)$. If

$\liminf_{\bar{\theta}_2 \in S_c} \inf_{\bar{\theta}_2 \in S_c} \frac{1}{N} (Q_2(\bar{\theta}_2) - Q_2(\theta_2^0)) > 0$ a.s. for any $c > 0$, then $\bar{\theta}_2$, the minimizer of $Q_2(\bar{\theta}_2)$, is a strongly consistent estimator of θ_2^0 .

Proof. It follows along the same lines of proof of Lemma 1 of [25]. □

Now, we need to show that $\liminf_{\bar{\theta}_2 \in S_c} \inf_{\bar{\theta}_2 \in S_c} \frac{1}{N} (Q_2(\bar{\theta}_2) - Q_2(\theta_2^0)) > 0$ a.s.. We will use the result from [10] for first component sequential plugin estimator, i.e., $\tilde{\beta} = \beta^0 + o(1/N^2)$. Expanding the difference $(Q_2(\bar{\theta}_2) - Q_2(\theta_2^0))$, and following the arguments of proof of Theorem 3 in [10], it follows that $\liminf_{\bar{\theta}_2 \in S_c} \inf_{\bar{\theta}_2 \in S_c} \frac{1}{N} (Q_2(\bar{\theta}_2) - Q_2(\theta_2^0)) > 0$ a.s. holds for any $c > 0$. ■

VI. APPENDIX B

A. Asymptotic Normality of LSEs

Proof: We will follow the notations from subsection II-B. Using multivariate Taylor Series expansion, we have:

$$\mathbf{Q}'(\hat{\boldsymbol{\theta}}) - \mathbf{Q}'(\boldsymbol{\theta}^0) = \mathbf{Q}''(\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),$$

where, $\hat{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$.

Above yields $\mathbf{D}^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) = -[\mathbf{D}\mathbf{Q}''(\hat{\boldsymbol{\theta}})\mathbf{D}]^{-1}\mathbf{D}\mathbf{Q}'(\boldsymbol{\theta}^0)$,

where $\mathbf{Q}'(\hat{\boldsymbol{\theta}}) = \mathbf{0}$ and $\mathbf{D}$ matrix is defined in Theorem 2. Let us now consider the derivative vector $\mathbf{D}\mathbf{Q}'(\boldsymbol{\theta}^0)$. By using the central limit theorem of stochastic processes (see [5]), we obtain asymptotic distribution of $\mathbf{D}\mathbf{Q}'(\boldsymbol{\theta}^0)$ as multivariate normal with mean vector $\mathbf{0}$ and variance co-variance matrix $2c\sigma^2\boldsymbol{\Sigma}_1^{-1}$ which can be represented in block matrix form as: $2c\sigma^2 \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$. Inversion of this block-matrix will provide elements of $\boldsymbol{\Sigma}_1$. Here, $\boldsymbol{\Sigma}_{11}$ is the identity matrix of order $2p \times 2p$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^\top$,

$$\boldsymbol{\Sigma}_{12} = \begin{bmatrix} B_1^0/2 & 0 & 0 & 0 & \cdots & 0 & B_1^0/3 \\ -A_1^0/2 & 0 & 0 & 0 & \cdots & 0 & -A_1^0/3 \\ 0 & B_2^0/2 & 0 & 0 & \cdots & 0 & B_2^0/3 \\ 0 & -A_2^0/2 & 0 & 0 & \cdots & 0 & -A_2^0/3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & B_p^0/2 & B_p^0/3 \\ 0 & 0 & 0 & 0 & \cdots & -A_p^0/2 & -A_p^0/3 \end{bmatrix}_{(2p \times (p+1))}$$

and for $k = 1, 2, \dots, p$, the k^{th} row of $\boldsymbol{\Sigma}_{22}$ contains $\frac{(A_k^{0^2} + B_k^{0^2})}{3}$ at k^{th} column and $\frac{(A_k^{0^2} + B_k^{0^2})}{4}$ at $(p+1)^{th}$ column. The $(p+1) \times (p+1)^{th}$ element of $\boldsymbol{\Sigma}_{22}$ is $\sum_{k=1}^p \frac{(A_k^{0^2} + B_k^{0^2})}{5}$. Remaining elements of $(p+1)^{th}$ row are obtained by symmetry. The rest of the elements of $\boldsymbol{\Sigma}_{22}$ are 0. Using expression of $\boldsymbol{\Sigma}_1^{-1}$, we can easily calculate each term of the matrix $\boldsymbol{\Sigma}_1$.

It can be seen that $[\mathbf{D}\mathbf{Q}''(\hat{\boldsymbol{\theta}})\mathbf{D}]^{-1} \xrightarrow{n \rightarrow \infty} \boldsymbol{\Sigma}_1$, where $2c\sigma^2\boldsymbol{\Sigma}_1^{-1}$ is the variance co-variance matrix of $\mathbf{D}\mathbf{Q}'(\boldsymbol{\theta}^0)$. By simple calculations, we get following results:

$$AsymVar\left(N^{-5/2}(\hat{\beta} - \beta^0)\right) = \frac{360c\sigma^2}{\sum_{k=1}^p (A_k^{0^2} + B_k^{0^2})}, \quad (11)$$

$$AsymVar\left(N^{-3/2}(\hat{\alpha}_k - \alpha_k^0)\right) = \frac{360c\sigma^2}{\sum_{k=1}^p (A_k^{0^2} + B_k^{0^2})} + \frac{24c\sigma^2}{(A_k^{0^2} + B_k^{0^2})}. \quad (12)$$

B. Asymptotic Normality of Sequential Combined Estimators:

Proof: We will follow the notations from subsection II-C. If we define $\boldsymbol{\theta}_1 = (A_1, B_1, \alpha_1, \beta_1)^\top$, then for the following sum of squares,

$$Q_1(\boldsymbol{\theta}_1) = \sum_{n=1}^N (y(n) - A_1 \cos(\alpha_1 n + \beta_1 n^2) - B_1 \sin(\alpha_1 n + \beta_1 n^2))^2.$$

Using multivariate Taylor Series expansion, we have:

$$\mathbf{Q}'_1(\check{\boldsymbol{\theta}}_1) - \mathbf{Q}'_1(\boldsymbol{\theta}_1^0) = \mathbf{Q}''_1(\check{\boldsymbol{\theta}}_1)(\check{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0),$$

where, $\check{\boldsymbol{\theta}}_1$ is a point between $\check{\boldsymbol{\theta}}_1$ and $\boldsymbol{\theta}_1^0$. For $\mathbf{D}_1^{-1} = \text{diag}(N^{1/2}, N^{1/2}, N^{3/2}, N^{5/2})$, this yields,

$$\mathbf{D}_1^{-1}(\check{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0) = -[\mathbf{D}_1\mathbf{Q}''_1(\check{\boldsymbol{\theta}}_1)\mathbf{D}_1]^{-1}\mathbf{D}_1\mathbf{Q}'_1(\boldsymbol{\theta}_1^0),$$

where, $\mathbf{Q}'_1(\check{\boldsymbol{\theta}}_1) = \mathbf{0}$. We first calculate the derivative vector $\mathbf{Q}'_1(\boldsymbol{\theta}_1^0)$ as:

$$\begin{aligned} \left. \frac{\partial Q_1(\boldsymbol{\theta}_1)}{\partial A_1} \right|_{\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^0} &= -2 \sum_{n=1}^N (y(n) - A_1^0 \cos(\alpha_1^0 n + \beta^0 n^2) - B_1^0 \sin(\alpha_1^0 n + \beta^0 n^2)) \cos(\alpha_1^0 n + \beta^0 n^2) \\ &= -2 \sum_{n=1}^N (X(n) + A_2^0 \cos(\alpha_2^0 n + \beta^0 n^2) + B_2^0 \sin(\alpha_2^0 n + \beta^0 n^2)) \cos(\alpha_1^0 n + \beta^0 n^2) \\ &= -2 \sum_{n=1}^N \left(X(n) \cos(\alpha_1^0 n + \beta^0 n^2) \right) + o(\sqrt{N}). \end{aligned}$$

Last equality above, is observed using conjectures mentioned in [10] and Lemma 2 of [6]. Similar calculations hold for other derivative terms. It can then be shown that the first component estimators follow asymptotic properties same as the LSEs of mono-component chirp model parameters. If we consider second component estimators $\boldsymbol{\theta}_2 = (\check{A}_2, \check{B}_2, \check{\alpha}_2, \check{\beta}_2)^\top$, and define

$$\check{y}_1(n) = y(n) - \check{A}_1 \cos(\check{\alpha}_1 n + \check{\beta}_1 n^2) - \check{B}_1 \sin(\check{\alpha}_1 n + \check{\beta}_1 n^2), \text{ and}$$

$$Q_2(\boldsymbol{\theta}_2) = \sum_{n=1}^N (\check{y}_1(n) - A_2 \cos(\alpha_2 n + \beta_2 n^2) - B_2 \sin(\alpha_2 n + \beta_2 n^2))^2.$$

Derivative vector of $Q_2(\boldsymbol{\theta}_2)$ at $\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2^0$ will have following elements:

$$\begin{aligned} \left. \frac{\partial Q_2(\boldsymbol{\theta}_2)}{\partial A_2} \right|_{\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2^0} &= -2 \sum_{n=1}^N \left(X(n) \cos(\alpha_2^0 n + \beta^0 n^2) \right) + o(\sqrt{N}), \\ \left. \frac{\partial Q_2(\boldsymbol{\theta}_2)}{\partial \alpha_2} \right|_{\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2^0} &= -2 \sum_{n=1}^N \left(X(n) (A_2^0 n \sin(\alpha_2^0 n + \beta^0 n^2) - B_2^0 n \cos(\alpha_2^0 n + \beta^0 n^2)) \right) + o(N\sqrt{N}), \end{aligned}$$

$$\left. \frac{\partial Q_2(\theta_2)}{\partial \beta} \right|_{\theta_2=\theta_2^0} = -2 \sum_{n=1}^N \left(X(n) (A_2^0 n^2 \sin(\alpha_2^0 n + \beta^0 n^2) - B_2^0 n^2 \cos(\alpha_2^0 n + \beta^0 n^2)) \right) + o(N\sqrt{N}).$$

Arguing along the same lines of proof of the first component, we get for $k \geq 1$, θ_k asymptotically follows Normal distribution with variance covariance matrix

$$\frac{2c\sigma^2}{A_k^{0^2} + B_k^{0^2}} \begin{bmatrix} A_k^{0^2} + 9B_k^{0^2} & -8A_k^0 B_k^0 & -36B_k^0 & 30B_k^0 \\ -8A_k^0 B_k^0 & 9A_k^{0^2} + B_k^{0^2} & 36A_k^0 & -30A_k^0 \\ -36B_k^0 & 36A_k^0 & 192 & -180 \\ 30B_k^0 & -30A_k^0 & -180 & 180 \end{bmatrix}.$$

Therefore, asymptotic variance of sequential combined estimators of α_k^0 is given by:

$$AsymVar\left(N^{-3/2}(\check{\alpha}_k - \alpha_k^0)\right) = \frac{384c\sigma^2}{(A_k^{0^2} + B_k^{0^2})}. \quad (13)$$

Note that different component sequential combined estimators are asymptotically independent. So, the final estimator of β^0 is given by minimizing the variance of linear combination:

$$\check{\beta} = \sum_{k=1}^p l_k \check{\beta}_k, \quad l_k \geq 0, \sum_{k=1}^p l_k = 1$$

$$\text{Solving above yields } l_k = \frac{(A_k^{0^2} + B_k^{0^2})}{\sum_{k=1}^p (A_k^{0^2} + B_k^{0^2})}.$$

We can use consistent estimators of l_k as they are functions of amplitude parameters. After simple calculations, it can be shown that asymptotic variance of $\check{\beta}$ matches with that of LSEs. ■

C. Asymptotic Normality of Sequential Plugin Estimators

Proof: Note that sequential plugin estimators of first component parameters are same as the sequential combined estimators. Thus, we have $D_1^{-1}(\tilde{\theta}_1 - \theta_1^0)$ converging in distribution to normal with asymptotic mean $\mathbf{0}$ and variance covariance matrix

$$\frac{2c\sigma^2}{A_1^{0^2} + B_1^{0^2}} \begin{bmatrix} A_1^{0^2} + 9B_1^{0^2} & -8A_1^0 B_1^0 & -36B_1^0 & 30B_1^0 \\ -8A_1^0 B_1^0 & 9A_1^{0^2} + B_1^{0^2} & 36A_1^0 & -30A_1^0 \\ -36B_1^0 & 36A_1^0 & 192 & -180 \\ 30B_1^0 & -30A_1^0 & -180 & 180 \end{bmatrix}.$$

where $D_1^{-1} = \text{diag}(N^{1/2}, N^{1/2}, N^{3/2}, N^{5/2})$. Now for the second component estimators, $\tilde{\theta}_2 = (\tilde{A}_2, \tilde{B}_2, \tilde{\alpha}_2)^T$ of $\theta_2^0 = (A_2^0, B_2^0, \alpha_2^0)^T$, we update the data as: $\tilde{y}_1(n) = y(n) - \tilde{A}_1 \cos(\tilde{\alpha}_1 n + \tilde{\beta} n^2) - \tilde{B}_1 \sin(\tilde{\alpha}_1 n + \tilde{\beta} n^2)$ and define the following sum of squares:

$$Q_2(\tilde{\theta}_2) = \sum_{n=1}^N \left(\tilde{y}_1(n) - A_2 \cos(\alpha_2 n + \tilde{\beta} n^2) - B_2 \sin(\alpha_2 n + \tilde{\beta} n^2) \right)^2.$$

Now, we calculate the elements of derivative vector $Q'_2(\tilde{\theta}_2^0)$ in following steps:

$$\left. \frac{\partial Q_2(\tilde{\theta}_2)}{\partial A_2} \right|_{\tilde{\theta}_2=\tilde{\theta}_2^0} = -2 \sum_{n=1}^N \left(X(n) \cos(\alpha_2^0 n + \beta^0 n^2) \right) + B_2^0 N^3 (\tilde{\beta} - \beta^0)/3 + o(\sqrt{N}).$$

Similarly, it can be shown that

$$\left. \frac{\partial Q_2(\tilde{\theta}_2)}{\partial B_2} \right|_{\tilde{\theta}_2=\tilde{\theta}_2^0} = -2 \sum_{n=1}^N \left(X(n) \sin(\alpha_2^0 n + \beta^0 n^2) \right) - A_2^0 N^3 (\tilde{\beta} - \beta^0)/3 + o(\sqrt{N}),$$

$$\left. \frac{\partial Q_2(\tilde{\theta}_2)}{\partial \alpha_2} \right|_{\tilde{\theta}_2=\tilde{\theta}_2^0} = 2 \sum_{n=1}^N \left(nX(n) \left(A_2^0 \sin(\alpha_2^0 n + \beta^0 n^2) - B_2^0 \cos(\alpha_2^0 n + \beta^0 n^2) \right) + (A_2^{0^2} + B_2^{0^2}) N^4 (\tilde{\beta} - \beta^0)/4 + o(N\sqrt{N}) \right).$$

$$D_2 Q_2''(\tilde{\theta}_2^0) D_2 \xrightarrow{n \rightarrow \infty} \begin{bmatrix} 1 & 0 & B_2^0/2 \\ 0 & 1 & -A_2^0/2 \\ B_2^0/2 & -A_2^0/2 & (A_2^{0^2} + B_2^{0^2})/3 \end{bmatrix} = \bar{\Sigma}_2, \text{ where } D_2^{-1} = \text{diag}(N^{1/2}, N^{1/2}, N^{3/2}), \text{ and } [D_2 Q_2''(\tilde{\theta}_2^0) D_2]^{-1} \xrightarrow{n \rightarrow \infty}$$

$$\frac{1}{A_2^{0^2} + B_2^{0^2}} \begin{bmatrix} A_2^{0^2} + 4B_2^{0^2} & -3A_2^0 B_2^0 & -6B_2^0 \\ -3A_2^0 B_2^0 & 4A_2^{0^2} + B_2^{0^2} & 6A_2^0 \\ -6B_2^0 & 6A_2^0 & 12 \end{bmatrix}.$$

Derivative vectors properties for any k^{th} component with $k = 3, 4, \dots, p$ can be obtained similarly. Some important observations on asymptotic variances and covariances are provided in the supplementary material (page 1, 2 and 3). Collecting these variances and covariances, we write $2c\sigma^2 \Sigma_3^k$ as the asymptotic variance covariance matrix of $D_2^{-1} Q'_k(\tilde{\theta}_k^0)$, then asymptotic variance covariance matrix of $\tilde{\theta}_k$ is given by:

$$2c\sigma^2 \bar{\Sigma}_k^{-1} \Sigma_3^k \bar{\Sigma}_k^{-1},$$

where $\lim_{n \rightarrow \infty} D_2 Q_2''(\tilde{\theta}_k^0) D_2 = \bar{\Sigma}_k$. After some straightforward calculations, we observe that

$$AsymVar(\tilde{\alpha}_k) = \frac{24c\sigma^2}{(A_k^{0^2} + B_k^{0^2})} + \frac{360c\sigma^2}{(A_1^{0^2} + B_1^{0^2})}, \quad (14)$$

$$AsymVar(\tilde{\beta}) = \frac{360c\sigma^2}{(A_1^{0^2} + B_1^{0^2})}, \quad (15)$$

$$AsymVar(\tilde{A}_k) = \frac{2c\sigma^2}{(A_k^{0^2} + B_k^{0^2})} \left[(A_k^{0^2} + 4B_k^{0^2}) + \frac{5B_k^{0^2} (A_k^{0^2} + B_k^{0^2})}{(A_1^{0^2} + B_1^{0^2})} \right], \quad (16)$$

$$AsymVar(\tilde{B}_k) = \frac{2c\sigma^2}{(A_k^{0^2} + B_k^{0^2})} \left[(4A_k^{0^2} + B_k^{0^2}) + \frac{5A_k^{0^2}(A_k^{0^2} + B_k^{0^2})}{(A_1^{0^2} + B_1^{0^2})} \right]. \quad (17)$$

Asymptotic covariance matrix $2c\sigma^2 \Sigma_{1k}$ between $\tilde{\theta}_1$ and $\tilde{\theta}_k$ is given by

$$\frac{2c\sigma^2}{(A_1^{0^2} + B_1^{0^2})} \begin{bmatrix} 5B_1^0 B_k^0 & -5B_1^0 A_k^0 & -30B_1^0 \\ -5A_1^0 B_k^0 & 5A_1^0 A_k^0 & 30A_1^0 \\ -30B_k^0 & 30A_k^0 & 180 \\ 30B_k^0 & -30A_k^0 & -180 \end{bmatrix}.$$

Asymptotic covariance between k^{th} and j^{th} components estimators for $1 < k < j \leq p$, can be obtained by using results provided in the supplementary material (page 3). ■

A short comparison between asymptotic variance of sequential plugin and sequential combined estimators of α_k^0 , A_k^0 and B_k^0 , for $k \geq 2$ has also been put in the supplementary material (page 3) for the sake of completeness of the study.

REFERENCES

- [1] Aifang, L., Xiaohua, Z., Jinhui, L., & Zhong, L. (2003). The ISAR range profile compensation of fast-moving target using the dechirp method. In IEEE International Conference on Neural Networks and Signal Processing, 2003, December. Proceedings of the 2003 (Vol. 2, pp. 1619-1623).
- [2] Barbarossa, S., Scaglione, A., & Giannakis, G. B. (1998). Product high-order ambiguity function for multicomponent polynomial-phase signal modeling. IEEE Transactions on Signal Processing, 46(3), 691-708.
- [3] Barbarossa, S. (1995). Analysis of multicomponent LFM signals by a combined Wigner-Hough transform. IEEE Transactions on Signal Processing, 43(6), 1511-1515.
- [4] Djurić, P. M. (1996). A model selection rule for sinusoids in white Gaussian noise. IEEE Transactions on Signal Processing, 44(7), 1744-1751.
- [5] Fuller, W. A. (2009). Introduction to statistical time series. John Wiley & Sons.
- [6] Grover, R., Kundu, D., & Mitra, A. (2018). Approximate least squares estimators of a two-dimensional chirp model and their asymptotic properties. Journal of Multivariate Analysis, 168, 211-220.
- [7] Grover, R., Kundu, D., & Mitra, A. (2021). An efficient methodology to estimate the parameters of a two-dimensional chirp signal model. Multidimensional Systems and Signal Processing, 32(1), 49-75.
- [8] Kundu, D., & Nandi, S. (2008). Parameter estimation of chirp signals in presence of stationary noise. Statistica Sinica, 187-201.
- [9] Liu, A., Zhu, X., & Lin, J. (2005). Application of the fractional Fourier transform to ISAR range profile compensation of fast-moving target. In IEEE/SP 13th Workshop on Statistical Signal Processing, 2005, July (pp. 950-955).
- [10] Lahiri, A., Kundu, D., & Mitra, A. (2015). Estimating the parameters of multiple chirp signals. Journal of Multivariate Analysis, 139, 189-206.
- [11] Ma, N., & Goh, J. T. (2004). DOA estimation for broadband chirp signals. In IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, May (Vol. 2, pp. ii-261).
- [12] Ma, N., & Goh, J. T. (2006). Ambiguity-function-based techniques to estimate DOA of broadband chirp signals. IEEE Transactions on signal Processing, 54(5), 1826-1839.
- [13] Nandi, S., & Kundu, D. (2020). Statistical signal processing. Springer Singapore.
- [14] Nandi, S., Grover, R., & Kundu, D. (2021). Estimation of parameters of multiple chirp signal in presence of additive alpha-stable errors. Signal Processing, 189, 108232.
- [15] Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. The computer journal, 7(4), 308-313.
- [16] Ozdemir, C. (2012). Inverse synthetic aperture radar imaging with MATLAB algorithms (Vol. 210). John Wiley & Sons.
- [17] Prasad, A., Kundu, D., & Mitra, A. (2008). Sequential estimation of the sum of sinusoidal model parameters. Journal of Statistical Planning and Inference, 138(5), 1297-1313.
- [18] Rao, W., Li, G., Wang, X., & Xia, X. G. (2014). Parametric sparse representation method for ISAR imaging of rotating targets. IEEE Transactions on Aerospace and Electronic Systems, 50(2), 910-919.
- [19] Richards, F. S. (1961). A method of maximum-likelihood estimation. Journal of the Royal Statistical Society: Series B (Methodological), 23(2), 469-475.
- [20] Ruan, H., Wu, Y. H., & Ye, W. (2013). A range profile motion compensation method for inverse synthetic aperture Ladar. In IET International Radar Conference, 2013, April (pp. 1-5).
- [21] Scrucca, L. (2013). GA: a package for genetic algorithms in R. Journal of Statistical Software, 53, 1-37.
- [22] Tian, B., Chen, Z., Xu, S., & Liu, Y. (2013). ISAR imaging compensation of high speed targets based on integrated cubic phase function. In MIPPR : Multispectral Image Acquisition, Processing, and Analysis, 2013, October (Vol. 8917, p. 89170B). International Society for Optics and Photonics.
- [23] Wang, P., & Yang, J. (2006). Multicomponent chirp signals analysis using product cubic phase function. Digital Signal Processing, 16(6), 654-669.
- [24] Wang, P., Li, H., & Himed, B. (2008). Parameter estimation of linear frequency-modulated signals using integrated cubic phase function. In IEEE 42nd Asilomar Conference on Signals, Systems and Computers, 2008, October (pp. 487-491).
- [25] Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. The Annals of Statistics, 9(3), 501-513.
- [26] Wu, Y., Ruan, H., & Yu, D. (2014). Inverse synthetic aperture laser radar imaging algorithm for maneuvering targets. In IEEE 7th International Congress on Image and Signal Processing, 2014, October (pp. 569-574).
- [27] Xu, F., Bao, Q., Chen, Z., Pan, S., & Lin, C. (2018). Parameter estimation of multi-component LFM signals based on STFT+ Hough transform and fractional fourier transform. In 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2018, May (pp. 839-842).
- [28] Xue, J., Han, X., & Zhang, Q. (2017). A symmetric accumulated cross-correlation method of parameter estimation based on fractional Fourier transform for ISAR motion compensation. In IEEE 9th International Conference on Wireless Communications and Signal Processing (WCSP), 2017, October (pp. 1-6).
- [29] Yang, P., Liu, Z., & Jiang, W. L. (2015). Parameter estimation of multi-component chirp signals based on discrete chirp Fourier transform and population Monte Carlo. Signal, Image and Video Processing, 9(5), 1137-1149.

Abhineek Shukla completed B.Sc Hons. degree in Statistics and M.sc degree in Statistics from Banaras Hindu University in the period of 2013-2018. He joined IIT Kanpur, India, as a research scholar under the supervision of Prof. Amit Mitra and Prof. Debasis Kundu. His current research interests are parameter estimation and detection of polynomial phase signal models. He has also worked in improving inference based on averaged stochastic gradient descent estimators.

Debasis Kundu received his Ph.D. degree from the Pennsylvania State University, Pennsylvania, PA, USA, in 1989, under the guidance of Prof. C. R. Rao. After finishing his Ph.D., he joined The University of Texas at Dallas, Dallas, TX, USA, as an Assistant Professor, before joining the Indian Institute of Technology Kanpur, Kanpur, India, in 1990. From 2011 to 2014, he was the Head of the Department of Mathematics and Statistics, IIT Kanpur, and since 2008, an Endowed Chair Professor, and since 2019, the Dean of Faculty Affairs of the Institute. He has published more than 300 research papers in different refereed journals. He has coauthored two research monographs on Statistical Signal Processing and on Step-Stress Models, and coedited a book on Statistical Computing. He works on different areas of statistics. His major research interests include distribution theory, lifetime data analysis, censoring, statistical signal processing, and statistical computing. Prof. Kundu is the recipient of the Chandana Award from the Canadian Mathematical Society, Distinguished Statistician Award from the Indian Society of Probability and Statistics and the P. C. Mahalanobis Award from the Indian Society of Operation Research. He is currently the Editor-in-Chief of the Journal of the Indian Society of Probability and Statistics, and a member of the editorial boards of Communications in Statistics—Theory and Methods, Communications in Statistics—Simulation and Computation, and Sankhya, Series B. He was a member of the editorial boards of the IEEE TRANSACTIONS ON RELIABILITY, Journal of Statistical Theory and Practice, and Journal of Distribution Theory. He is a Fellow of the National Academy of Sciences, India and a Fellow of the Royal Statistical Society London.

Amit Mitra received his B.Sc. degree in Statistics from University of Calcutta. He obtained his M.Sc. degree and Ph.D. from Indian Institute of Technology Kanpur, both in Statistics. He is a Professor of Statistics at the Department of Mathematics and Statistics, Indian Institute of Technology Kanpur. His research interests are statistical signal processing, data mining of financial and economic time series and parameter estimation of non-linear time series models.

Rhythm Grover received her B.Sc. degree from Delhi University. She obtained her M.Sc. and Ph.D. from Indian Institute of Technology Kanpur. She was a postdoctoral researcher at Indian Statistical Institute in the Theoretical Statistics and Mathematics Unit. She joined Mehta Family School of Data Science and Artificial Intelligence, IIT Guwahati as an Assistant Professor in 2021. Her research interests lie at the intersection of statistics and signal processing, particularly in development of efficient algorithm for parameter estimation of signal processing models, derivation of statistical properties of classical parameter estimation methods and robust estimation of model parameters in the presence of outliers.