Multidimensional Systems and Signal Processing
https://doi.org/10.1007/s11045-018-0605-1

Order estimation of 2-dimensional complex superimposed exponential signal model using exponentially embedded family (EEF) rule: large sample consistency properties

Anupreet Porwal[1] · Sharmishtha Mitra[1] · Amit Mitra[1]

Received: 28 November 2017 / Revised: 7 May 2018 / Accepted: 6 July 2018

Abstract
In this paper, we consider the order estimation problem of a 2-dimensional complex superimposed exponential signal model in presence of additive white noise. We use the recently proposed exponentially embedded family (EEF) rule (see Stoica and Babu in IEEE Signal Process Lett 19(9):551–554, 2012; Kay in IEEE Trans Aerosp Electron Syst 41(1):333–345, 2005) for estimating the order of the 2-dimensional signal model and prove that the EEF rule based estimator is consistent in large sample scenario. Extensive simulations are performed to ascertain the performance of the order estimation rule and also to compare the finite sample performance of EEF rule based estimator with other popular order selection rules using simulation examples.

Keywords Akaike information criterion (AIC) · Bayesian information criterion (BIC) · Two-dimensional complex superimposed exponential signal · Consistency · Exponentially embedded family (EEF) · Model order estimation

1 Introduction

The need for model order selection is of primary interest to parametric signal processing and time series problems. In this paper, we consider the problem of estimation of model order of a 2-dimensional complex superimposed exponential signal model. Detection of signal components in the presence of noise of a 2-dimensional complex exponential model is an important problem in statistical signal processing. Specifically, we consider the following signal model:

$$y(s,t) = f(s,t,\theta_m) + \epsilon(s,t);\ 1 \le s \le S;\ 1 \le t \le T; \tag{1}$$

Sharmishtha Mitra: Work partially supported by Research Project Grant MTR/2017/000208 of Science and Engineering Research Board, Government of India.

✉ Amit Mitra
amitra@iitk.ac.in

1 Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur 208016, India

Published online: 11 July 2018

$$f(s, t, \theta_m) = \sum_{k=1}^{m} \alpha_k e^{i(s\beta_k + t\omega_k)}. \tag{2}$$

Thus the signal model is

$$y(s, t) = \sum_{k=1}^{m} \alpha_k e^{i(s\beta_k + t\omega_k)} + \epsilon(s, t). \tag{3}$$

$\theta_m = (\alpha_{1_R}, \alpha_{1_C}, \beta_1, \omega_1, ..., \alpha_{m_R}, \alpha_{m_C}, \beta_m, \omega_m)'$ is a $4m$x1 vector of unknown signal parameters; α_{jR} and α_{jC} denotes the real and the imaginary parts of α_j for $j = 1, ..., m$. Let m_o denote the unknown true number of components in the the observed signal. Given a sample of size ST, $y = (y(1, 1), \ldots, y(S, 1), y(1, 2), \ldots, y(S, 2), \ldots, y(S, T))'$, the model order estimation problem is to estimate m_o.

Abundant literature on model order selection techniques exists, including some of the most popular techniques, the Bayesian Information Criteria (BIC) (Schwarz 1978; Rissanen 1978, 1982) and the Akaike Information Criteria (AIC) (Sakamoto et al. 1986; Akaike 1974). There has been some recent advancements too, where novel rules have been designed with favourable properties like Penalizing Adaptively the Likelihood (PAL) approach (Stoica and Babu 2013) and exponentially embedded family (EEF) rule (Stoica and Babu 2012; Kay 2005). Kay (2005) proposed the EEF rule for model selection. Stoica and Babu (2012) presented a GLR based derivation of the proposed EEF rule and thus bridged a connection between EEF and other conventional model order selection techniques. Recently, Zhu and Kay (2018), gives an explanation on the EEF penalty term from a Bayesian perspective, which may shed light on explaining the large sample consistency of EEF rule. An attempt to extend EEF in complex valued signal processing, specifically for estimating degree of non-circularity of complex valued signal vectors have been made by Zhu and Kay (2017). More recently, Agrawal et al. (2018) used the EEF formulation for order estimation of real valued multiple sinusoidal model and established that the EEF rule based order estimator is large sample consistent. No attempt as such in the literature is reported however, at least not known to the authors, regarding EEF approach for high dimensional complex valued signal models. The purpose of this paper is to extend the EEF rule for addressing the problem of order estimation of higher dimensional complex valued signal models and study the large sample behaviour of the estimator.

Detection of the signal component in presence of noise is an important problem in statistical signal processing. A real valued 2-dimensional sinusoidal model has received considerable attention in the signal processing literature because of its widespread applicability in the texture analysis. The problem of estimating the parameters of complex valued two-dimensional exponential signals corrupted by noise occurs in a variety of signal processing problems (see, e.g. Li et al. 1996; Rao et al. 1994; Kundu and Mitra 1996; Mitra and Stoica 2002 and the references cited therein). In this paper, we study the large sample asymptotic properties of the EEF based estimator of the model order m_o of (3). Results proved in this paper generalize existing results available in the literature for one-dimensional real valued sinusoidal model (see for example (Agrawal et al. 2018) and the references cited therein).

The rest of the paper is organised as follows. In Sect. 2, we formulate the EEF rule for model order estimation of the 2-dimensional complex exponential signal model. In Sect. 3, we study and establish the consistency property of the EEF based estimator of model order. Finally in Sect. 4, we present the finite sample simulation studies to access the performance of the EEF based order estimation method and compare it's performance with other prominent model selection rules.

2 Model order estimation of 2-dimensional complex exponential signals using EEF

In this section we formulate the EEF rule based order estimation method for estimating the order of the signal model (3). We will be using the following notations through out the paper.

$$y = (y(1,1), \dots, y(S,1), y(1,2), \dots, y(S,T))'$$
$$\epsilon = (\epsilon(1,1), \dots, \epsilon(S,1), \epsilon(1,2), \dots, \epsilon(S,T))'$$
$$f(\theta_m) = (f(1,1,\theta_m), \dots, f(S,1,\theta_m), \dots, f(S,T,\theta_m))'$$

We make the following assumptions for the model (3);

- Assumption A1: $\epsilon(s,t)$ are i.i.d complex valued gaussian with zero mean such that
$$\epsilon(s,t) = \epsilon_R(s,t) + i\,\epsilon_C(s,t),$$
$$\epsilon_R(s,t) \sim \mathcal{N}(0, \sigma^2/2),$$
$$\epsilon_C(s,t) \sim \mathcal{N}(0, \sigma^2/2).$$
Further, $\epsilon_R(s,t)$ and $\epsilon_C(s,t)$ are independent.
- Assumption A2: $\forall k = 1, 2, ..., m_o$, $(\beta_k, \omega_k) \in (0, 2\pi)\text{x}(0, 2\pi)$; where (β_k, ω_k) are pairwise different i.e. $\omega_j \neq \omega_k$ or $\beta_j \neq \beta_k, \forall j \neq k$. Furthermore, $\forall k = 1, 2..., m_o$, α_ks are bounded.
- Assumption A3: The true model parameter vector θ_{m_o} is an interior point in the parameter space $\Theta \subset \mathbb{R}^{4m_o}$.

Under the stated assumptions on the complex noise random variables, the likelihood function for an m-component model is given by

$$f_m(y, \theta_m^*) = \frac{1}{\left(\frac{2\pi\sigma^2}{2}\right)^{ST}} e^{\frac{-(y-f(\theta_m))^H (y-f(\theta_m))}{\sigma_m^2}}. \tag{4}$$

For solving the order estimation problem, we consider the set of $\tilde{m}$ nested models given by $\{M_m\}_{m=1}^{\tilde{m}}$, where M_m is the m component 2-dimensional signal model with parameter vector θ_m^*. We assume that the true model M_{m_o} is contained in this set, i.e., $m_o \leq \tilde{m}$ and y is not a purely white noise process i.e., $M_0 \neq M_{m_o}$.

Let us consider the Generalized likelihood ratio,

$$\hat{r}_{m+1} = 2\ln\left[\frac{f_m(y, \hat{\theta}_m^*)}{f_0(y, \hat{\theta}_0^*)}\right], \tag{5}$$

where $\hat{\theta}_k^*$ is the maximum likelihood estimate of the underlying signal and noise parameter vector θ_k^* and $f_0(y, \hat{\theta}_0^*)$ denotes the p.d.f of y when M_0 is the model.

The EEF rule based criterion function for the 2-dimensional complex exponential signal model is formulated as

$$EEF(m) = \begin{cases} \hat{r}_{m+1} - (4m+1)\left[1 + \ln\left(\frac{\hat{r}_{m+1}}{4m+1}\right)\right], & \text{if } \hat{r}_{m+1} > 4m+1. \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

EEF based estimator of the model order is obtained by maximising the EEF statistic over $m = 1, 2, \dots, \tilde{m}$, i.e.

$$\hat{m} = \underset{m \in \{1,2,\dots,\tilde{m}\}}{\arg\max} \; EEF(m). \tag{7}$$

Remark The above formulation of EEF rule based estimator for the 2-dimensional complex exponential model is a generalization of the formulation introduced in (Stoica and Babu 2012; Kay 2005; Agrawal et al. 2018) applicable for one-dimensional real valued sinusoidal model.

3 Consistency of EEF rule

In this section, we establish the main large sample asymptotic results of the paper. We write the model (3) as:

$$y = D_m(\beta, \omega)a_m + \epsilon, \tag{8}$$

where, $\forall j = 1, 2, ..., m$,

$$\begin{aligned} e_j &= (e^{i(\beta_k+\omega_k)}, \dots, e^{i(S\beta_k+\omega_k)}, e^{i(S\beta_k+T\omega_k)})', \\ \beta &= (\beta_1, \beta_2, ..., \beta_m)', \\ \omega &= (\omega_1, \omega_2, ..., \omega_m)', \\ D_m(\beta, \omega) &= (e_1, e_2, ..., e_m)_{ST\text{x}m}, \\ a_m &= (\alpha_1, \alpha_2, ..., \alpha_m)'_{m\text{x}1}. \end{aligned}$$

Throughout the paper, we would write D_m for $D_m(\beta, \omega)$ and $\hat{D_m}$ for $D_m(\hat{\beta}, \hat{\omega})$ for notational simplicity. In the derivations and results that follow, we will use the following notations.

- We use $O(.)$ and $o(.)$ to denote either the order in probability or deterministic order depending on the context.
- Let $\{\Psi_i\}$ be a sequence of rectangles such that $\Psi_i = \{(s, t) \in \mathbb{Z}^2 | 1 \leq s \leq S_i, 1 \leq t \leq T_i\}$. Then, the sequence of subsets $\{\Psi_i\}$ will be said to tend to infinity as $i \to \infty$ if $\lim_{i\to\infty} min(S_i, T_i) \to \infty$ and $0 < \lim_{i\to\infty}\left(\frac{S_i}{T_i}\right) < \infty$. For notational convenience, we omit subscript i. Thus, $\Psi(S, T) \to \infty$ implies both S and T tend to infinity as a function of i, and roughly at the same rate.

We need the following lemmas to prove the main result.

Lemma 3.1 *Under the assumptions A1-A3,* $\forall m \leq m_o$,

$$\hat{\sigma}_m^2 = \sigma^2 + \sum_{j=1}^{m_o} \alpha_j^H \alpha_j - \sum_{j=1}^{m} \hat{\alpha}_j^H \hat{\alpha}_j + o(1), \tag{9}$$

a.s. as $\Psi(S, T) \to \infty$.

Proof The estimated variance of noise for (8) given by

$$\hat{\sigma}_m^2 = \frac{y^H(I_{ST} - P_m(\hat{\beta}, \hat{\omega}))y}{ST}, \tag{10}$$

where, $P_m(\hat{\beta}, \hat{\omega}) = \hat{D_m}(\hat{D_m}^H \hat{D_m})^{-1}\hat{D_m}^H$ is the projection matrix. Note that for the true model $y = D_{m_o}a_{m_o} + \epsilon$ and hence

$$\begin{aligned} \hat{\sigma}_m^2 =& \frac{1}{ST}\Big[\epsilon^H\epsilon + 2\epsilon^H D_{m_o}a_{m_o} + a_{m_o}^H D_{m_o}^H D_{m_o}a_{m_o} \\ & - \hat{a}_{m_o}^H \hat{D}_{m_o}^H \hat{D}_{m_o}\hat{a}_{m_o}\Big], \end{aligned} \tag{11}$$

where, $\hat{a}_m = (\hat{D_m}^H \hat{D_m})^{-1} \hat{D_m}^H y$. Since $\epsilon_R(s,t)$ and $\epsilon_C(s,t)$ are independent, by Kolmogorov's Strong law of Large Numbers (see e.g. Chung 2001) we have

$$\begin{aligned}\frac{\epsilon^H \epsilon}{ST} &= \frac{\sum_{s=1}^S \sum_{t=1}^T \epsilon_R^2(s,t)}{ST} + \frac{\sum_{s=1}^S \sum_{t=1}^T \epsilon_C^2(s,t)}{ST}, \\ &\to \frac{\sigma^2}{2} + \frac{\sigma^2}{2} = \sigma^2,\end{aligned} \tag{12}$$

a.s. as $\Psi(S,T) \to \infty$. Realize that

$$\frac{\epsilon^H D_{m_o} a_{m_o}}{ST} = \frac{1}{ST}\left(\sum_{k=1}^{m_o} \alpha_k \left(\sum_{l=1}^{S}\sum_{j=1}^{T} \bar{\epsilon}(l,j) e^{i(l\beta_k + j\omega_k)}\right)\right). \tag{13}$$

Lemma 2 of Kundu and Mitra (1999; 1996) implies

$$\frac{\epsilon^H D_{m_o} a_{m_o}}{ST} = o(1) \text{ a.s.} \tag{14}$$

as $\Psi(S,T) \to \infty$. Further, using the fact that

$$\lim_{n\to\infty} \frac{\sum_{t=1}^n e^{it\tilde{\omega}}}{n} = o(1) \text{ for some fixed } \tilde{\omega} \in (0, 2\pi), \tag{15}$$

it can be shown that

$$\frac{1}{ST} a_{m_o}^H D_{m_o}^H D_{m_o} a_{m_o} = \sum_{j=1}^{m_o} \alpha_j^H \alpha_j + o(1) \text{ a.s.,} \tag{16}$$

and

$$\frac{1}{ST} \hat{a}_{m_o}^H \hat{D}_{m_o}^H \hat{D}_{m_o} \hat{a}_{m_o} = \sum_{j=1}^{m} \hat{\alpha}_j^H \hat{\alpha}_j + o(1) \text{ a.s.} \tag{17}$$

as $\Psi(S,T) \to \infty$. Thus using (12), (14), (16) and (17), we have $\forall m \le m_o$

$$\hat{\sigma}_m^2 = \sigma^2 + \sum_{j=1}^{m_o} \alpha_j^H \alpha_j - \sum_{j=1}^{m} \hat{\alpha}_j^H \hat{\alpha}_j + o(1) \text{ a.s. as } \Psi(S,T) \to \infty. \tag{18}$$

□

Lemma 3.2 *Under A1-A3, for any integer $k \ge 1$*

$$\hat{\sigma}_{m_o+k}^2 = \hat{\sigma}_{m_o}^2 - \frac{G_k}{ST} + o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right) \tag{19}$$

a.s. as $\Psi(S,T) \to \infty$, where $G_k = \sum_{j=1}^k I_\epsilon(\hat{\beta}_{m_o+j}, \hat{\omega}_{m_o+j})$
and

$$I_\epsilon(\beta,\omega) = \frac{1}{ST}\left|\sum_{s=1}^{S}\sum_{t=1}^{T} \epsilon(s,t) e^{-i(s\beta + t\omega)}\right|^2. \tag{20}$$

$I_\epsilon(\beta,\omega)$ corresponds to periodogram of underlying white noise process and $(\hat{\beta}_{m_o+1}, \hat{\omega}_{m_o+1})$, ... , $(\hat{\beta}_{m_o+k}, \hat{\omega}_{m_o+k})$ are frequencies that correspond to the k largest peaks of $I_\epsilon(\beta,\omega)$. Thus, G_k is the sum of k largest elements of the periodogram of noise.

Proof Note that

$$\hat{\alpha}_{m_o+l} = \frac{1}{ST}\sum_{s=1}^{S}\sum_{t=1}^{T}\epsilon(s,t)e^{-i(s\hat{\beta}_{m_o+l}+t\hat{\omega}_{m_o+l})},$$

$$\hat{D}_{m_o+k} = (\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_{m_o+k}),$$

$$= (\hat{D}_{m_o}, \hat{e}_{m_o+1}, \hat{e}_{m_o+2}, \ldots, \hat{e}_{m_o+k}).$$

Similarly,

$$\hat{a}_{m_o+k} = (\hat{a}'_{m_o}, \hat{\alpha}_{m_o+1}, \hat{\alpha}_{m_o+2}, \ldots, \hat{\alpha}_{m_o+k})'.$$

Using Theorem 2 of Francos and Kliger (2005) (see also Kliger and Francos 2008; Prasad et al. 2008), we note that in case of overestimation, MLE of the overestimated parameter vector contains a subvector equal in dimension to the true model order, that converges almost surely to the true parameter vector. Further, the frequencies of the overestimated components correspond to those which sequentially maximise the noise periodogram. The variance of the overestimated model is given by,

$$\hat{\sigma}^2_{m_o+k} = \frac{y^H(I_{ST} - P_{m_o+k}(\hat{\beta}_{m_o+k}, \hat{\omega}_{m_o+k}))y}{ST}. \tag{21}$$

Substituting the values of $\hat{a}_{m_o+k}$ and $\hat{D}_{m_o+k}$ we get

$$\hat{\sigma}^2_{m_o+k} = \hat{\sigma}^2_{m_o} - T_1 - T_2, \tag{22}$$

where,

$$\hat{\sigma}^2_{m_o} = \frac{1}{ST}(y^H y - \hat{a}^H_{m_o}\hat{D}^H_{m_o}\hat{D}_{m_o}\hat{a}_{m_o}),$$

$$T_1 = 2\sum_{k=1}^{k} T_{1_j} \text{ and } T_2 = \sum_{j=1}^{k} T_{2_j} + 2\sum_{j=1}^{k}\sum_{\substack{\tilde{j}=1\\ j\neq\tilde{j}}}^{k} T_{2(j,\tilde{j})},$$

$$T_{1_j} = \frac{1}{ST}\delta_j\Delta_j^H\hat{D}_{m_o}\hat{a}_{m_o},$$

$$T_{2_j} = \frac{1}{ST}\delta_j\Delta_j^H\Delta_j\delta_j^H,$$

$$T_{2(j,\tilde{j})} = \frac{1}{ST}\delta_j\Delta_j^H\Delta_{\tilde{j}}\delta_{\tilde{j}}^H,$$

and $\delta_j = \hat{\alpha}_{m_o+j}$ and $\Delta_j = \hat{e}_{m_o+j}$. If $\epsilon(s,t)$ is 2-dimensional array of circularly symmetric Gaussian with zero mean and finite variance then, $E(\epsilon_R(1,1)^2 log|\epsilon_R(1,1)|) < \infty$ and $E(\epsilon_C(1,1)^2 log|\epsilon_C(1,1)|) < \infty$ (He 1995) and hence it follows from Theorem 2.2 of He (1995) that

$$\limsup_{\Psi(S,T)\to\infty} \frac{\sup_\omega I_\epsilon(\omega)}{\sigma^2 \ln(ST)} \leq 8 \text{ a.s.} \tag{23}$$

Also, note that $\forall\omega \in (0, 2\pi)$

$$\frac{1}{n}\sum_{t=1}^{n} e^{it\omega} = o\left(\frac{\ln n}{n}\right)^{1/2}. \tag{24}$$

Observe that $\forall j = 1, 2, \ldots, k$

$$|\hat{\alpha}_{m_o+j}|^2 = \frac{1}{ST} I_\epsilon(\widehat{\beta_{m_o+j}}, \widehat{\omega_{m_o+j}}), \tag{25}$$

$$|\hat{\alpha}_{m_o+j}| = o\left(\frac{\ln(ST)}{ST}\right)^{1/2} \text{ a.s..} \tag{26}$$

Using (23),(24) and (25), $\forall j = 1, 2, \ldots, m_o$ as $\Psi(S, T) \to \infty$

$$\frac{1}{ST}\hat{\alpha}_{m_o+1}\hat{e}^H_{m_o+1}\hat{e}_j\hat{\alpha}^H_j = o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right).$$

Thus,

$$T_{1_1} = o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right) \text{ a.s.}$$

Further, $\forall j = 1, 2, \ldots, k$

$$T_{1_j} = o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right) \text{ a.s.,}$$

$$T_1 = 2\sum_{k=1}^{k} T_{1_j} = o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right) \text{a.s.,}$$

as $\Psi(S, T) \to \infty$. Similarly, $\forall j \neq \tilde{j}$

$$T_{2(j,\tilde{j})} = o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right) \text{a.s. as } \Psi(S, T) \to \infty$$

Also, $\forall 1 \leq j \leq k$,

$$T_{2_j} = \frac{1}{ST}\hat{\alpha}_{m_o+j}\hat{e}^H_{m_o+j}\hat{e}_{m_o+j}\hat{\alpha}^H_{m_o+j},$$

$$= |\hat{\alpha}_{m_o+j}|^2 = \frac{1}{ST} I_\epsilon(\widehat{\beta_{m_o+j}}, \widehat{\omega_{m_o+j}}).$$

Thus finally we have as $\Psi(S, T) \to \infty$

$$\hat{\sigma}^2_{m_o+k} = \hat{\sigma}^2_{m_o} - \frac{G_k}{ST} + o\left(\frac{(\ln(ST)\ln S \ln T)^{1/2}}{ST}\right), \tag{27}$$

almost surely, where $G_k = \sum_{j=1}^{k} I_\epsilon(\hat{\beta}_{m_o+j}, \hat{\omega}_{m_o+j})$. □

Lemma 3.3 *Under assumptions A1-A3, r_m as defined before in (5) satisfies*

$$r_m = \begin{cases} 0, & m = 1, \\ O(ST), & 2 \leq m \leq \tilde{m}. \end{cases} \tag{28}$$

Proof From (5), we get

$$r_m = 2ST\ln\left(\frac{\hat{\sigma}^2_0}{\hat{\sigma}^2_{m-1}}\right), \tag{29}$$

where, $\hat{\sigma}^2_0 = \frac{1}{ST}\sum_{s=1}^{S}\sum_{t=1}^{T} y(s,t)^H y(s,t)$ as $\Psi(S, T) \to \infty$. It is easy to see that for $m = 1, r_1 = 0$. Under the assumptions on the 2-dimensional complex array of noise random variables, MLE $\hat{\theta}_k$ of θ_k is same as the nonlinear least square estimate and we have using results from Kliger and Francos (2008), Prasad et al. (2008) that as $\Psi(S, T) \to \infty$,

$$\hat{\alpha}_j \to \alpha_j \text{ for j} = 1, 2, \ldots, \text{m}; m \leq m_o. \tag{30}$$

We consider the overestimation and underestimation cases separately.
Case 1: $m \leq m_o$(underestimation)
$\forall m \leq m_o$, we have

$$\hat{\sigma}_m^2 \to \sigma^2 + \sum_{j=m+1}^{m_o} \alpha_j^H \alpha_j \text{a.s. and } \hat{\sigma}_{m_o}^2 \to \sigma^2 \text{ a.s.} \tag{31}$$

as $\Psi(S, T) \to \infty$. Thus for $2 \leq m \leq m_o$

$$\frac{r_m}{ST} = 2 \ln\left(1 + \frac{\sum_{j=1}^{m-1} \alpha_j^H \alpha_j}{\sigma^2 + \sum_{j=m}^{m_o} \alpha_j^H \alpha_j}\right) \text{ a.s.} \tag{32}$$

as $\Psi(S, T) \to \infty$. Note that the r.h.s. is bounded by model assumptions and is a strictly positive quantity. Hence, we have $r_m = O(n)$ a.s. for all $m \leq m_o$.
Case 2: $m > m_o$(overestimation)
For $m > m_o$, using 3.1 and Lemma 3.2 we have

$$\frac{r_m}{ST} = 2 \ln\left(\frac{\sigma^2 + \sum_{j=1}^{m_o} \alpha_j^H \alpha_j}{\sigma^2 + \frac{G_{m-m_o}}{ST} + o\left(\frac{(\ln(ST) \ln S \ln T)^{1/2}}{ST}\right)}\right) \text{ a.s.} \tag{33}$$

as $\Psi(S, T) \to \infty$. Since $\forall k = 1, 2, \ldots, \tilde{m} - m_o$, $G_k = \sum_{j=1}^{k} I_\epsilon(\beta_{m_o+j}, \omega_{m_o+j})$. Using (23), $G_k = O(\ln(ST))$

$$\frac{G_{m-m_o}}{ST} \to 0 \text{ a.s and } o\left(\frac{(\ln(ST) \ln S \ln T)^{1/2}}{ST}\right) \to 0, \tag{34}$$

as $\Psi(S, T) \to \infty$. Combining this we have for all $m > m_o$

$$\frac{r_m}{ST} = 2 \ln\left(1 + \frac{\sum_{j=1}^{m_o} \alpha_j^H \alpha_j}{\sigma^2}\right) = O(1) \text{ a.s. as } \Psi(S, T) \to \infty. \tag{35}$$

Hence, we have the desired result. □

Remark Note that Lemma 3.1, Lemma 3.2, Lemma 3.3 and Lemma 3.4 are generalizations of the lemmas presented in Agrawal et al. (2018) for one-dimensional real valued multiple sinusoidal model. The novelty in the proofs of the above lemmas and also in the proof of the main consistency result that follows, as compared to those in Agrawal et al. (2018), lie in the treatment of 2-dimensional array of circularly symmetric Gaussian noise and resultant 2-dimensional array of complex random signals for deriving the asymptotic results. The lemmas in Agrawal et al. (2018) follow from the present lemmas as special cases. The lemmas presented above allows us to generalize the large sample consistency results of EEF rule based estimator for one-dimensional real valued sinusoidal signal model to the case of two-dimensional complex exponential signal model. We now state and prove the main consistency result of the paper.

Theorem *Under the Assumptions A1-A3, with m_o as the true model order and if $\hat{m}$ is the estimated model order using EEF rule, then*

$$P(\hat{m} \neq m_o) \to 0 \; as \;\; \Psi(S, T) \to \infty. \tag{36}$$

Proof In the proof, we consider the overestimation and underestimation cases separately and prove that in each of the cases, probability of wrong estimation of model order goes to zero as $\Psi(S,T) \to \infty$.

Case I: $m \le m_o$ **(underestimation)**

Subcase I: $\hat{r}_{m+1} > 4m+1$;$\hat{r}_{m_o+1} > 4m_o+1$

It is easy to see that

$$EEF(m)-EEF(m_o) = \hat{r}_{m+1}-\hat{r}_{m_o+1}-(4m+1)\ln(\hat{r}_{m+1})+(4m_o+1)\ln(\hat{r}_{m_o+1})+k, \quad (37)$$

where,

$$k = 4(m_o-m)+(4m+1)\ln(4m+1)-(4m_o+1)\ln(4m_o+1). \quad (38)$$

Also since

$$\hat{r}_{m+1} = 2ST\ln\left(\frac{\hat{\sigma_o^2}}{\hat{\sigma_m^2}}\right), \quad (39)$$

we have,

$$EEF(m)-EEF(m_o) = -2ST\ln\left(\frac{\hat{\sigma_m^2}}{\hat{\sigma_{m_o}^2}}\right)-(4m+1)\ln(\hat{r}_{m+1})+(4m_o+1)\ln(\hat{r}_{m_o+1})+k. \quad (40)$$

Using Lemma 3, we get $\hat{r}_{m+1} = O(ST)\ \forall m \le m_o$ and using Lemma 1, we have $\forall m \le m_o$

$$\hat{\sigma_m^2} \to \sigma^2 + \sum_{k=m+1}^{m_o} \alpha_k^H \alpha_k \text{ a.s.,}$$

$$\text{and } \hat{\sigma_{m_o}^2} \to \sigma^2,$$

a.s. as $\Psi(S,T) \to \infty$. Thus we have $\forall m \le m_o$

$$\ln\left(\frac{\hat{\sigma_m^2}}{\hat{\sigma_{m_o}^2}}\right) \to \ln\left(1+\frac{\sum_{j=m+1}^{m_o}\alpha_j^H\alpha_j}{\sigma^2}\right)\text{a.s. as } \Psi(S,T)\to\infty, \quad (41)$$

the r.h.s. above is strictly positive and a bounded quantity. Using this in (40), we get

$$EEF(m)-EEF(m_o) = O(ST)-(4m+1)O(\ln(ST))+(4m_o+1)O(\ln(ST))+k \text{ a.s..} \quad (42)$$

Subsequently, we have that

$$\frac{EEF(m)-EEF(m_o)}{ST} \to -2\ln\left(1+\frac{\sum_{j=m+1}^{m_o}\alpha_j^H\alpha_j}{\sigma^2}\right), \quad (43)$$

a.s. as $\Psi(S,T) \to \infty$. Therefore, we have $EEF(m) < EEF(m_o)$ with probability 1.

Subcase 2: $\hat{r}_{m+1} < 4m+1$;$\hat{r}_{m_o+1} > 4m_o+1$ We have

$$EEF(m) = 0,$$

$$EEF(m_o) = \hat{r}_{m_o+1}-(4m_o+1)\left[1+\ln\left(\frac{\hat{r}_{m_o+1}}{4m_o+1}\right)\right].$$

Note that $g(x) = x - \ln x - 1$ has a unique minimum value of 0 at $x = 1$, hence $EEF(m_o) > 0$ for $\hat{r}_{m_o+1} \ne 3m_o+1$. Thus $EEF(m_o) > EEF(m), \forall m \le m_o$ and hence underestimation is not possible in this case.

Subcase 3: $\hat{r}_{m+1} > 4m+1$;$\hat{r}_{m_o+1} < 4m_o+1$

We know from Stoica and Babu (2012), that for large sample size $EEF(m) \cong \hat{r}_{m+1} - (4m+1)\ln n$. Thus $\hat{r}_{m_o+1} < 4m_o + 1$ and $EEF(m_o) = 0$ and hence we have

$$\frac{EEF(m) - EEF(m_o)}{ST\ln(ST)} = \frac{\hat{r}_{m+1}}{ST\ln(ST)} - \frac{3m+1}{ST} \tag{44}$$

as $\Psi(S,T) \to \infty$. By Lemma 3, $\hat{r}_{m+1} = O(ST)$ and therefore

$$\frac{EEF(m) - EEF(m_o)}{ST\ln(ST)} \to 0 \text{ as } \Psi(S,T) \to \infty,$$
$$\text{or, } P\Big(\frac{EEF(m) - EEF(m_o)}{ST\ln(ST)} > 0\Big) \to 0 \text{ as } \Psi(S,T) \to \infty.$$

Thus for all the subcases discussed above,

$$\begin{aligned} &P(\hat{m} < m_o) \\ &\quad = P(EEF(m) > EEF(m_o) \text{ for some } m < m_o) \\ &\quad \to 0 \text{ as } \Psi(S,T) \to \infty. \end{aligned}$$

Case 2: $m > m_o$ **(overestimation)**
Subcase 1: $\hat{r}_{m+1} > 4m + 1$;$\hat{r}_{m_o+1} > 4m_o + 1$
We have

$$EEF(m) - EEF(m_o) = -2ST\ln\Big(\frac{\hat{\sigma}^2_m}{\hat{\sigma}^2_{m_o}}\Big) - (4m+1)\ln(\hat{r}_{m+1}) + (4m_o+1)\ln(\hat{r}_{m_o+1}) + k. \tag{45}$$

Since $m > m_o$, $\hat{r}_{m+1} = 2\ln\Big(\frac{\hat{f}_m}{\hat{f}_o}\Big) > 2\ln\Big(\frac{\hat{f}_{m_o}}{\hat{f}_o}\Big) = \hat{r}_{m_o+1}$, using which we have

$$EEF(m) - EEF(m_o) \le -2ST\ln\Big(\frac{\hat{\sigma}^2_m}{\hat{\sigma}^2_{m_o}}\Big) + k + 4(m_o - m)\ln(\hat{r}_{m+1}). \tag{46}$$

It follows from the asymptotic theory of likelihood ratios (see Wilks 1946; Lehmann and Romano 2005) that,

$$2ST\ln\Big(\frac{\hat{\sigma}^2_m}{\hat{\sigma}^2_{m_o}}\Big) \sim \chi^2_{4(m-m_o)}. \tag{47}$$

Thus, $\frac{1}{\ln(ST)}\Big(2ST\ln\Big(\frac{\hat{\sigma}^2_m}{\hat{\sigma}^2_{m_o}}\Big)\Big) = o_p(1)$. Since k is independent of n and $m > m_o$, and from Lemma 3 we have $\hat{r}_{m+1} = O(ST)$, we get

$$P\Big(\frac{1}{\ln(ST)}\Big(-2ST\ln\Big(\frac{\hat{\sigma}^2_m}{\hat{\sigma}^2_{m_o}}\Big) + k + 4(m_o - m)\ln(\hat{r_{m+1}})\Big) < 0\Big) \to 1 \text{ as } \Psi(S,T) \to \infty. \tag{48}$$

Hence we have from (46) that

$$P\Big(\frac{EEF(m) - EEF(m_o)}{\ln(ST)} < 0\Big) \to 1 \text{ as } \Psi(S,T) \to \infty. \tag{49}$$

Therefore,

$$\begin{aligned} &P(\hat{m} > m_o) \\ &\quad = P(EEF(m) - EEF(m_o) > 0 \text{ for some } m > m_o) \\ &\quad = P\Big(\frac{EEF(m) - EEF(m_o)}{\ln(ST)} > 0\Big) \to 0 \text{ as } \Psi(S,T) \to \infty. \end{aligned}$$

Subcase 2 : $\hat{r}_{m+1} < 4m + 1$;$\hat{r}_{m_o+1} > 4m_o + 1$
Under this subcase,

$$EEF(m) = 0$$
$$\text{and } EEF(m_o) = \hat{r}_{m_o+1} - (4m_o + 1)\left[1 + \ln\left(\frac{\hat{r}_{m_o+1}}{4m_o + 1}\right)\right].$$

Using the same argument as in Subcase 2 of Case 1, we conclude that overestimation is not possible in this case either.
Subcase 3 : $\hat{r}_{m+1} > 4m + 1$; $\hat{r}_{m_o+1} < 4m_o + 1$
Similar to the underestimation case, we know from Stoica and Babu (2012), that for large samples $EEF(m) \cong \hat{r}_{m+1} - (4m + 1)\ln(ST)$. Note that $\hat{r}_{m_o+1} < 4m_o + 1$, $EEF(m_o) = 0$ and hence we have

$$\frac{EEF(m) - EEF(m_o)}{ST\ln(ST)} = \frac{\hat{r}_{m+1}}{ST\ln(ST)} - \frac{3m + 1}{ST}$$

as $\Psi(S, T) \to \infty$. From Lemma 3, $\hat{r}_{m+1} = O(ST)$ and hence

$$P\left(\frac{EEF(m) - EEF(m_o)}{ST\ln(ST)} < 0\right) \to 1 \text{ as } \Psi(S, T) \to \infty. \tag{50}$$

Thus for all the subcases discussed above,

$$\begin{aligned} P(\hat{m} < m_o) &= P(EEF(m) < EEF(m_o) \text{ for some } m < m_o) \\ &\to 0 \text{ as } \Psi(S, T) \to \infty \end{aligned} \tag{51}$$

Hence, from the two cases of over and underestimation we have that

$$P(\hat{m} \neq m_o) \to 0 \text{ as } \Psi(S, T) \to \infty. \tag{52}$$

□

4 Numerical simulations

In this section we present the finite sample simulation studies to investigate the performance of the EEF based order estimation rule for estimating the order of 2-dimensional complex exponential signal model and to compare its performance with other popular model order estimation rules. We compare the performance of the EEF based order estimates with AIC, BIC and PAL based approach for model order selection. We consider the following simulation signal model

$$y(s, t) = \sum_{k=1}^{2} \alpha_k e^{i(s\beta_k + t\omega_k)} + \epsilon(s, t).$$

with $\alpha_1 = 1 + \sqrt{2}i$, $\beta_1 = 0.26\pi$, $\omega_1 = 0.26\pi$; $\alpha_2 = 2 + 2i$, $\beta_2 = 0.62\pi$, $\omega_2 = 0.62\pi$.
$\epsilon(s, t)$ is taken as complex valued Gaussian as defined in Assumption A1 (Fig. 1). We vary the value of σ^2 from 0.5 to 15, S and T are varied from 9 to 20 and compute the estimate of probability of correct estimation of model order over 500 simulation runs and compare the performances with the rules listed below.

$$PAL(m) = -2\ln(f_m(y, \hat{\theta}_m^*)) + (4m + 1)\ln(4\tilde{m} + 1)\frac{\ln(r_m + 1)}{\ln(\rho_m + 1)},$$

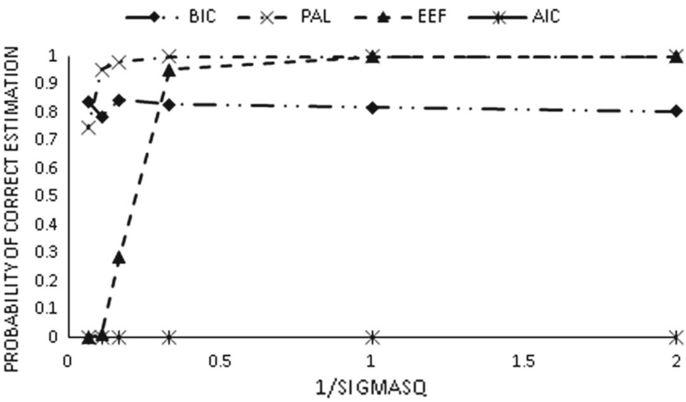

Fig. 1 Probability of correct order estimation; $S, T = 9$

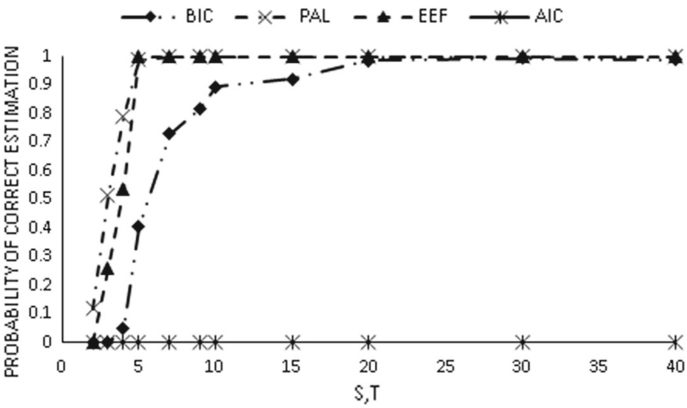

Fig. 2 Probability of correct order estimation; $\sigma^2 = 1$

$$BIC(m) = -2\ln(f_m(y, \hat{\theta}_m^*)) + (4m + 1)\ln(ST),$$
$$AIC(m) = -2\ln(f_m(y, \hat{\theta}_m^*)) + 2(4m + 1).$$

Where, ρ_m in PAL(m) is given by $\rho_m = 2\ln\left[\frac{f_{\tilde{m}}(y,\hat{\theta}_{\tilde{m}}^*)}{f_{m-1}(y,\hat{\theta}_{m-1}^*)}\right]$.

From the simulation studies we observe that the EEF rules performs well for low noise variance and large sample sizes of S and T. Probability of correct estimation of model order at $S, T = 9$ reaches 1 even for high sigma square values of 3. Its performance is better than AIC and BIC for small values of σ^2 and its comparable to PAL rule. However for larger values of noise variance, the probability of correct estimation takes large number of samples for attaining the limiting probability of 1. Simulation results validate the asymptotic consistency results for EEF established in the paper. We further observe that, except AIC, for all the other methods, the probability of correct estimation increases as S and/or T increases or as noise variance decreases (Figs. 1, 2, 3, 4, 5 and 6).

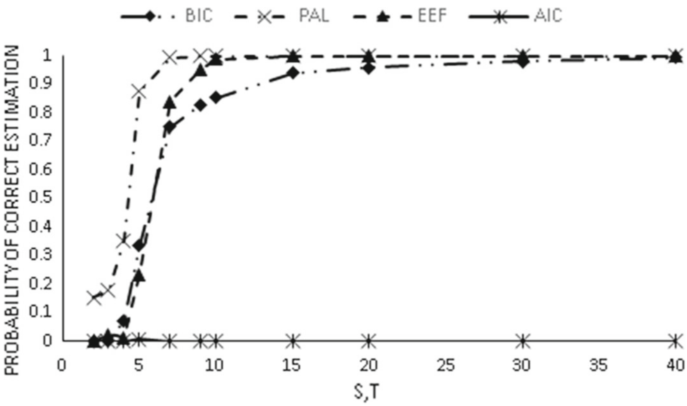

Fig. 3 Probability of correct order estimation; $\sigma^2 = 3$

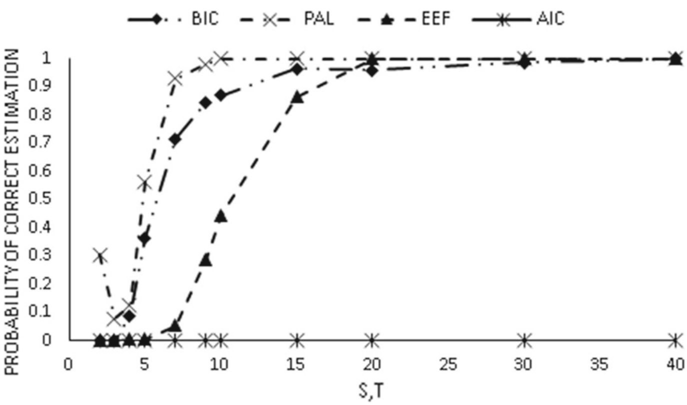

Fig. 4 Probability of correct order estimation; $\sigma^2 = 6$

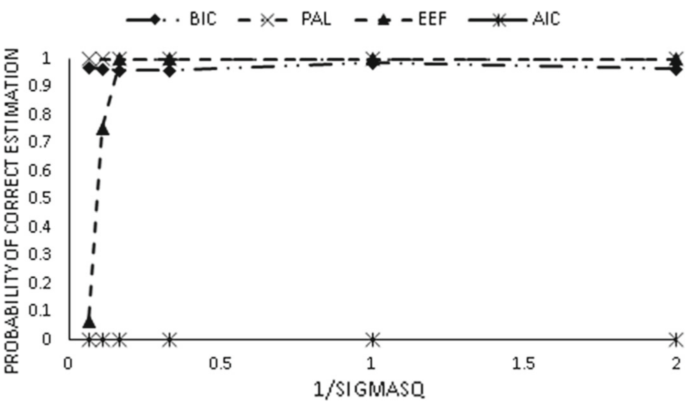

Fig. 5 Probability of correct order estimation; $S, T = 20$

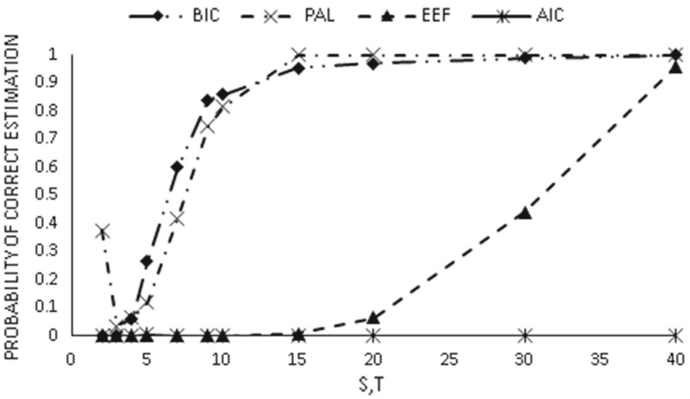

Fig. 6 Probability of correct order estimation; $\sigma^2 = 15$

5 Conclusions

In this paper, we framed EEF based model order estimation of the order of a 2-dimensional superimposed complex exponential signals model and study the large sample asymptotic statistical properties of the estimator of model order. We established that the estimator is large sample consistent. Numerical simulations are performed to ascertain the performance of the order estimation method and to compare its performance with other popular order estimation methods. Simulation results validate the theoretical asymptotic results and show satisfactory performance of the EEF rule based method.

References

Agrawal, Shruti, Mitra, Sharmishtha, & Mitra, Amit. (2018). Estimating the order of multiple sinusoids model using exponentially embedded family rule: Large sample consistency. *Signal Processing*, *147*, 54–59.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Chung, K. L. (2001). *A course in probability theory*. Cambridge: Academic Press.

Francos, J.M., & Kliger, M. (2005). Strong consistency of the over-and under-determined LSE of 2-D exponentials in white noise. In *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 4. IEEE.

He, S. (1995). Uniform convergence of weighted periodogram of stationary linear random fields. *Chinese Annals of Mathematics*, *16*, 331–340.

Kay, S. (2005). Exponentially embedded families-new approaches to model order estimation. *IEEE Transactions on Aerospace and Electronic Systems*, *41*(1), 333–345.

Kliger, M., & Francos, J. M. (2008). Strong consistency of a family of model order selection rules for estimating 2D sinusoids in noise. *Statistics & Probability Letters*, *78*(17), 3075–3081.

Kundu, D., & Mitra, A. (1996). Asymptotic properties of the least squares estimates of 2-D exponential signals. *Multidimensional Systems & Signal Processing*, *7*, 135–150.

Kundu, D., & Mitra, A. (1996). Asymptotic properties of the least squares estimates of 2-D exponential signals. *Multidimensional Systems and Signal Processing*, *7*(2), 135–150.

Kundu, D., & Mitra, A. (1999). On asymptotic properties and confidence intervals for exponential signals. *Signal Processing*, *72*(2), 129–139.

Lehmann, E. L., & Romano, J. P. (2005). *Testing statistical hypotheses*. Berlin: Springer.

Li, J., Stoica, P., & Zheng, D. (1996). An efficient algorithm for two-dimensional frequency estimation. *Multidimensional Systems and Signal Processing*, *7*, 151?178.

Mitra, A., & Stoica, P. (2002). The Asymptotic Cramer-Rao bound for 2-D superimposed exponential signals. *Multidimensional Systems & Signal Processing*, *13*, 317–331.
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.
Rao, C. R., Zhao, L., & Zhou, B. (1994). Maximum likelihood estimation of 2-D superimposed exponential signals. *IEEE Transactions on Signal Processing*, *42*(7), 1795?1802.
Rissanen, J. (1982). Estimation of structure by minimum description length. In *Rational Approximation in Systems Engineering*. Birkhäuser Boston, pp. 395–406.
Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, *14*(5), 465–471.
Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). Akaike information criterion statistics. Dordrecht: D. Reidel.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.
Stoica, P., & Babu, P. (2012). On the exponentially embedded family (EEF) rule for model order selection. *IEEE Signal Processing Letters*, *19*(9), 551–554.
Stoica, P., & Babu, P. (2013). Model order estimation via penalizing adaptively the likelihood (PAL). *Signal Processing*, *93*(11), 2865–2871.
Wilks, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *The Annals of Mathematical Statistics*, *17*(3), 257–281.
Zhu, Z., & Kay, S. (2017). Estimate the degree of noncircularity of complex-valued vectors via exponentially embedded family. In *Proceedings of IEEE Radar Conference*, Seattle.
Zhu, Z., & Kay, S. (2018). On Bayesian exponentially embedded family for model order selection. *IEEE Transactions on Signal Processing*, *66*(4), 933–943.

Anupreet Porwal is an analyst working with North American liability strategies team at Deutsche Bank since 2017. Prior to joining Deutsche Bank, he received his B.S. and M.S. in Mathematics and Scientific Computing from Indian Institute of Technology Kanpur. Selected honors include B.D. Sanghi Memorial gold medal and Prof. Burton J. Moyer memorial gold medal for excellence in performance during undergraduate and graduate programs at IIT Kanpur. His broad research interests span statistical modelling techniques and probabilistic machine learning. He will be joining the Ph.D. program in Statistics at the University of Washington Seattle in Fall 2018.

Sharmishtha Mitra completed Bachelor of Science (Honours) in Statistics from Presidency College, Kolkata, and then did her Master of Science and Ph.D. from the Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur. She began her professional career in the Department of Statistical Analysis and Computer Services (now, Department of Systems and Information Management) of the Reserve Bank of India, Mumbai, in 1996. In 2005 she joined her alma mater, IIT Kanpur as a Visiting Faculty in the Department of Mathematics and Statistics, and at present, she is an Associate Professor there. Her research interests include Order Statistics, Reliability, Econometrics and Time Series Models. She has 27 research publications in peer-reviewed journals.

Amit Mitra completed Masters and Ph.D. in Statistics from Indian Institute of Technology, Kanpur. He did his post-doctoral research at the Department of Systems and Control, Uppsala University, Sweden. At present, he is a Professor in the Department of Mathematics and Statistics, Indian Institute of Technology Kanpur. Prior to joining IIT Kanpur in 2005, he worked as Assistant Professor at IIT Bombay and as Assistant Adviser, Department of Statistical Analysis and Computer Services, Reserve Bank of India, Mumbai. His research interests include statistical signal processing; datamining of financial and economic time series; parameter estimation of non-linear time series models. He has been the recipient of several awards and recognition, which include the “M. N. Murthy Award” from the Indian Statistical Institute for “Best Research Work done in India by a Young Scientist in the field of Applied Statistics”, the research award from Swedish Foundation for Strategic Research (SSF) and the “Distinguished Teacher Award” of the Indian Institute of Technology Kanpur. He has 49 research publications in international journals in Statistics and Signal Processing.