MISSPECIFIED CRAMER-RAO BOUND OF SEMI-BLIND CHANNEL ESTIMATION WITH SIDE INFORMATION

Le Trung Thanh[⋆,†], Karim Abed-Meraim[⋆], Nguyen Linh Trung[†]

[⋆]PRISME Laboratory, University of Orléans, Orléans, France
[†]AVITECH Institute, VNU University of Engineering and Technology, Hanoi, Vietnam

***Abstract*—In this paper, we derive performance lower bounds for semi-blind channel estimation in MIMO-OFDM systems with prior propagation channel information in the presence of channel order misspecification. We propose to use misspecified Cramer-Rao bound (MCRB), which is an extension of CRB for misspecification models, to analyze the theoretical performance limit of channel estimators. Two closed-form expressions of the MCRB are derived for unbiased training-based estimators and semi-blind estimators. Experimental results indicate that such prior information will gain the performance limit of channel estimators even when the order is underestimated. Besides, the proposed MCRBs become the classical CRBs when the channel order is accurately estimated.**

***Index Terms*—Peformance lower bounds, Misspecified Cramer-Rao bound, Misspecification, MIMO, Channel order, Propagation channel.**

I. Introduction

Channel estimation is an essential problem in MIMO-OFDM communication systems [1]. Methods for channel estimation can broadly be categorized into three main classes: blind, training-based, and semi-blind. Blind methods can estimate the channel directly from received signals, but often require a large number of transmitted data. Training-based methods exploit known pilot sequences at the receiver to identify the channel. They provide better estimation accuracy but result in lower spectral efficiency than blind methods. Semi-blind methods would be preferable when information of pilots and unknown data are both exploited. This approach not only yields a competitive performance but also gives rise to the spectral efficiency and throughput of communications systems. Also, incorporating prior and side information on signals, channels, and systems can be tractable for improving the quality of the estimates.

In estimation, Cramer-Rao bound (CRB) is often used as a benchmark for parameter estimators as it provides a lower bound on their variance [2]. Many studies have derived analytical expressions of the CRB for channel estimators in general and semi-blind estimators in particular, see [3]–[8] for examples. In the presence of side information, some efforts have been conducted to qualify the gain via the lens of CRB such as in [9] and [10]. However, these CRB bounds are appropriate only for perfect specification models. Very recently, our companion works on blind channel estimation have introduced misspecified CRBs for analyzing performance limits of blind methods when the channel order information is imperfect [11], [12]. The misspecified bounds are, however, designed only for blind channel estimators. Performance lower bounds for semi-blind channel estimation under model misspecification remain to date unexplored.

To fill this gap, in this paper, we propose to use the misspecified CRB (MCRB), which is a generalization of the classical CRB [13]–[15], to analyze the performance limit of semi-blind estimators in the presence of side information when the channel order is misspecified. Here, the specular channel model is investigated where the fading, time delay, and DOA parameters are considered. Two closed-form expressions of the MCRB are then derived for unbiased training-based and semi-blind estimators. The proposed MCRBs, for the first time, provide performance lower bounds for semi-blind channel estimation techniques under channel order misspecification.

II. Misspecified Cramer-Rao Bound

We briefly review the MCRB which is an extension of CRB for dealing with misspecification models [15]. Assume that data samples are i.i.d. derived from the true $f(\boldsymbol{y})$. In mismatched estimation, instead of $f(\boldsymbol{y})$, users adopt a different pdf $\tilde{f}(\boldsymbol{y}|\boldsymbol{\theta})$ where $\tilde{f}(\boldsymbol{y}|\boldsymbol{\theta}) \neq f(\boldsymbol{y}) \forall \boldsymbol{\theta}$ is allowed. For the users, the problem of interest is now to estimate $\boldsymbol{\theta}$.

In the context of MCRB, Kullback-Leibler (KL) divergence is used to determine the "best" performance unbiased estimators can achieve under model misspecification, which is defined as

$$\mathrm{KL}\big(f_{\mathbf{y}} \parallel \tilde{f}_{\mathbf{y}|\boldsymbol{\theta}}\big) \triangleq \mathbb{E}_f\big\{\log f(\boldsymbol{y})\big\} - \mathbb{E}_f\big\{\log \tilde{f}(\boldsymbol{y}|\boldsymbol{\theta})\big\}. \quad (1)$$

The unique minimizer of $\mathrm{KL}(f_{\mathbf{y}} \parallel \tilde{f}_{\mathbf{y}|\boldsymbol{\theta}})$ is called the *pseudo-true* parameter, $\boldsymbol{\theta}_{pt}$. In fact, minimizing (1) is equivalent to maximizing

$$\boldsymbol{\theta}_{pt} \triangleq \underset{\boldsymbol{\theta}}{\mathrm{argmin}}\, \mathrm{KL}\big(f_{\mathbf{y}} \parallel \tilde{f}_{\mathbf{y}|\boldsymbol{\theta}}\big) = \underset{\boldsymbol{\theta}}{\mathrm{argmax}}\, \mathbb{E}_f\big\{\ell(\boldsymbol{y}|\boldsymbol{\theta})\big\}. \quad (2)$$

Accordingly, the maximum likelihood estimator (MLE) converges in probability to $\boldsymbol{\theta}_{pt}$ [13], [16].

Let $\hat{\boldsymbol{\theta}}$ be an estimator derived under the misspecified model $\tilde{f}(\boldsymbol{y}|\boldsymbol{\theta})$ from the output samples. We call $\hat{\boldsymbol{\theta}}$ as misspecified (MS)-unbiased estimator if and only if

$$\mathbb{E}_f\big\{\hat{\boldsymbol{\theta}}(\boldsymbol{y})\big\} = \int \hat{\boldsymbol{\theta}}(\boldsymbol{y}) f(\boldsymbol{y}) d\boldsymbol{y} = \boldsymbol{\theta}_{pt}. \quad (3)$$

Covariance of any MS-unbiased estimator $\hat{\boldsymbol{\theta}}$ is bounded by[1]

$$\mathrm{VAR}\big(\hat{\boldsymbol{\theta}}\big) \geq \mathrm{MCRB}(\boldsymbol{\theta}_{pt}) \triangleq \boldsymbol{A}^{\#}(\boldsymbol{\theta}_{pt}) \boldsymbol{J}(\boldsymbol{\theta}_{pt}) \boldsymbol{A}^{\#}(\boldsymbol{\theta}_{pt}), \quad (4)$$

[1]In regular problems, the generalized interpretation of the MCRB in (4) boils down to the usual one

$$\mathrm{MCRB}(\boldsymbol{\theta}_{pt}) \triangleq \boldsymbol{A}^{-1}(\boldsymbol{\theta}_{pt}) \boldsymbol{J}(\boldsymbol{\theta}_{pt}) \boldsymbol{A}^{-1}(\boldsymbol{\theta}_{pt}).$$

where $(.)^{\#}$ denotes the pseudo-inverse operator and the two matrices $\boldsymbol{J}(\boldsymbol{\theta})$ and $\boldsymbol{A}(\boldsymbol{\theta})$ are defined as

$$\boldsymbol{J}(\boldsymbol{\theta}) = \mathbb{E}_f\left\{\frac{\partial \ell}{\partial \boldsymbol{\theta}^*}\left(\frac{\partial \ell}{\partial \boldsymbol{\theta}^*}\right)^H\right\},\ \boldsymbol{A}(\boldsymbol{\theta}) = \mathbb{E}_f\left\{\frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial \ell}{\partial \boldsymbol{\theta}^*}\right)\right\}. \quad (5)$$

We refer the reader to [12]–[15] for further details.

III. System Model

In this work, we consider a MIMO-OFDM system with N_t transmit antennas and N_r receive antennas using K subcarriers. Suppose that the MIMO channels are modeled as finite impulse responses (FIR) and remain constant during the transmission period. Each OFDM symbol contains K data samples and a cyclic prefix (CP) to avoid intersymbol interference. Particularly, the length of CP is required to be greater than the maximum channel delay L.

After removing the CP and taking the K-point FFT, the output signal $\boldsymbol{y}_r$ at the r-th receiver is given by

$$\boldsymbol{y}_r = \sum_{j=1}^{N_t} \boldsymbol{F}\mathcal{T}(\boldsymbol{h}_{r,j})\frac{\boldsymbol{F}}{K}\boldsymbol{x}_j + \boldsymbol{v}_r, \quad (6)$$

where $\boldsymbol{x}_j$ is the j-th OFDM symbol; $\boldsymbol{v}_r$ is a $N_r \times 1$ additive noise vector drawn from an i.i.d. circular complex Guassian distribution $\mathcal{CN}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}_{N_r})$, $\mathbb{E}\{\boldsymbol{v}_r \boldsymbol{v}_r^\top\} = \boldsymbol{0}$; $\boldsymbol{F}$ is the K-point discrete Fourier matrix; $\mathcal{T}(\boldsymbol{h}_{r,j})$ is a circulant matrix formed by the channel $\boldsymbol{h}_{r,j}$ of order $L-1$ between the r-th receiver and the j-th transmitter which is defined as

$$h_{r,j}(l) = \sum_{m=1}^{M} \beta_{m,j}\,\mathrm{sinc}(l - \tau_{m,j}) e^{-i2\pi\frac{d}{\lambda}(r-1)\sin(\alpha_{m,j})}, \quad (7)$$

for $l = 0, \dots, L-1$, where M is the number of multipaths; $\beta_{m,j}$, $\tau_{m,j}$, and $\alpha_{m,j}$ are the fading, path delay and DOA respectively; λ and d denote the wave length and the distance between two adjacent receive antennas.

Within a block data transmission, collecting N_r output vectors $\{\boldsymbol{y}_r\}_{r=1}^{N_r}$ into a single vector $\boldsymbol{y} = \left[\boldsymbol{y}_1^\top, \boldsymbol{y}_2^\top, \dots, \boldsymbol{y}_{N_r}^\top\right]^\top$, we can recast the model (6) into the following expression

$$\boldsymbol{y} = \boldsymbol{\lambda}\boldsymbol{x} + \boldsymbol{v}, \quad (8)$$

where the input vector $\boldsymbol{x}$ and noise $\boldsymbol{v}$ are given by

$$\boldsymbol{x} = \left[\boldsymbol{x}_1^\top, \boldsymbol{x}_2^\top, \dots, \boldsymbol{x}_{N_t}^\top\right]^\top, \quad \boldsymbol{v} = \left[\boldsymbol{v}_1^\top, \boldsymbol{v}_2^\top, \dots, \boldsymbol{v}_{N_r}^\top\right]^\top, \quad (9)$$

and the system matrix $\boldsymbol{\lambda} \in \mathbb{C}^{N_r K \times N_t K}$ is given by

$$\boldsymbol{\lambda} = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_{N_t}] \text{ with } \boldsymbol{\lambda}_j = [\boldsymbol{\lambda}_{1,j}, \boldsymbol{\lambda}_{2,j}, \dots, \boldsymbol{\lambda}_{N_r,j}]^\top, \quad (10)$$

where $\boldsymbol{\lambda}_{r,j} = \mathrm{diag}(\boldsymbol{F}_L \boldsymbol{h}_{r,j})$, and $\boldsymbol{F}_L$ is the matrix containing the L first columns of $\boldsymbol{F}$.

The expression (8) is equivalent to

$$\boldsymbol{y} = \boldsymbol{\mathcal{X}}\boldsymbol{h} + \boldsymbol{v}, \quad (11)$$

where the channel vector $\boldsymbol{h}$ is given by

$$\boldsymbol{h} = \left[\boldsymbol{h}_1^\top, \boldsymbol{h}_2^\top, \dots, \boldsymbol{h}_{N_r}^\top\right]^\top, \boldsymbol{h}_r = \left[\boldsymbol{h}_{r,1}^\top, \boldsymbol{h}_{r,2}^\top, \dots, \boldsymbol{h}_{r,N_t}^\top\right]^\top, \quad (12)$$

and the input matrix $\boldsymbol{\mathcal{X}}$ is defined as $\boldsymbol{\mathcal{X}} = \boldsymbol{I}_{N_r} \otimes \boldsymbol{X}$ where

$$\boldsymbol{X} = \left[\mathrm{diag}(\boldsymbol{x}_1)\boldsymbol{F}_L, \mathrm{diag}(\boldsymbol{x}_2)\boldsymbol{F}_L, \dots, \mathrm{diag}(\boldsymbol{x}_{N_t})\boldsymbol{F}_L\right]. \quad (13)$$

Without loss of generality, we consider the block-type pilot arrangement. With this arrangement, all subcarriers are used for pilots and data transmission. Data symbols are assumed to be i.i.d. circular Gaussian variables with zero mean and covariance matrix $\boldsymbol{C}_x = \mathrm{diag}\left([\sigma_{x_1}^2, \sigma_{x_2}^2, \dots, \sigma_{x_{N_t}}^2]\right)$ where $\sigma_{x_j}^2$ denotes the transmit power of the j-th user. Suppose that we have N_p pilot symbols and N_d data symbols in total.

Now let us denote $\boldsymbol{y}^{(p)} = \left[\boldsymbol{y}_1^{(p)\top}, \boldsymbol{y}_2^{(p)\top}, \dots, \boldsymbol{y}_{N_p}^{(p)\top}\right]^\top$ and $\boldsymbol{y}^{(d)} = \left[\boldsymbol{y}_1^{(d)\top}, \boldsymbol{y}_2^{(d)\top}, \dots, \boldsymbol{y}_{N_d}^{(d)\top}\right]^\top$ the output vectors corresponding to the N_p pilot symbols and N_d data symbols respectively. The true joint pdf function $f\left(\boldsymbol{y}^{(p)}, \boldsymbol{y}^{(d)} | \boldsymbol{\theta}\right)$ can be expressed as follows

$$\begin{aligned} f\left(\boldsymbol{y}^{(p)}, \boldsymbol{y}^{(d)} | \boldsymbol{\theta}\right) &= \prod_{n=1}^{N_p} \frac{1}{(\pi\sigma_v^2)^{N_r K}} \exp\left(-\frac{1}{\sigma_v^2}\left\|\boldsymbol{y}_n^{(p)} - \boldsymbol{m}_n^{(p)}\right\|_2^2\right) \\ &\times \prod_{n=1}^{N_d} \frac{1}{\pi^{N_r K}\det \boldsymbol{C}} \exp\left(-\boldsymbol{y}_n^{(d)H} \boldsymbol{C}^{-1} \boldsymbol{y}_n^{(d)}\right), \quad (14) \end{aligned}$$

where the mean $\boldsymbol{m}_n^{(p)} = \boldsymbol{\lambda}\boldsymbol{x}_n^{(p)} = \boldsymbol{\mathcal{X}}_n^{(p)}\boldsymbol{h}$ and the covariance matrix is given by $\boldsymbol{C} = \sum_{j=1}^{N_t} \sigma_{x_j}^2 \boldsymbol{\lambda}_j \boldsymbol{\lambda}_j^H + \sigma_v^2 \boldsymbol{I}_{KN_r}$, and the vector of unknown parameters is given by

$$\boldsymbol{\theta} = \left[\boldsymbol{\beta}^\top, \boldsymbol{\alpha}^\top, \boldsymbol{\tau}^\top, \sigma_v^2, \boldsymbol{\sigma}_x^2\right]. \quad (15)$$

More specifically, the parameter vector $\boldsymbol{\psi} = \left[\boldsymbol{\beta}^\top, \boldsymbol{\alpha}^\top, \boldsymbol{\tau}^\top\right]^\top$ is of interest with $\boldsymbol{\beta} = [\beta_{1,1}, \dots, \beta_{m,j}]^\top$, $\boldsymbol{\tau} = [\tau_{1,1}, \dots, \tau_{m,j}]^\top$, and $\boldsymbol{\alpha} = [\alpha_{1,1}, \dots, \alpha_{m,j}]^\top$, while the signal and noise powers are considered as nuisance parameters.

In practice, the true channel order L is not always specified or estimated correctly while it strongly impacts the performance of channel estimators [12]. In such case, the true pdf (14) is not accessible. Accordingly, classical performance lower bounds on the mean square estimation like CRB cannot be directly applied. In this paper, we fill this lack by using the MCRB introduced in [13], [15].

IV. Proposed MCRB for Specular Channel Estimation

In this section, we derive the analytical derivation of the MCRB to analyze the performance limit of unbiased channel estimators when the channel order is misspecified.

Under the channel order misspecification, the users use the following assumed pdf function

$$\begin{aligned} \tilde{f}\left(\boldsymbol{y}^{(p)}, \boldsymbol{y}^{(d)} | \boldsymbol{\theta}\right) &= \prod_{n=1}^{N_p} \frac{1}{(\pi\sigma_v^2)^{N_r K}} \exp\left(-\frac{1}{\sigma_v^2}\left\|\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\right\|_2^2\right) \\ &\times \prod_{n=1}^{N_d} \frac{1}{\pi^{N_r K}\det \tilde{\boldsymbol{C}}} \exp\left(-\boldsymbol{y}_n^{(d)H} \tilde{\boldsymbol{C}}^{-1} \boldsymbol{y}_n^{(d)}\right) \\ &\triangleq \tilde{f}_p(\boldsymbol{y}^{(p)} | \boldsymbol{\theta}) \tilde{f}_d(\boldsymbol{y}^{(d)} | \boldsymbol{\theta}), \quad (16) \end{aligned}$$

where the true mean $\boldsymbol{m}_n^{(p)}$ turns out $\tilde{\boldsymbol{m}}_n^{(p)} = \tilde{\boldsymbol{\mathcal{X}}}_n^{(p)}\tilde{\boldsymbol{h}} = \tilde{\boldsymbol{\lambda}}\boldsymbol{x}_n^{(p)}$, the new channel parameter $\tilde{\boldsymbol{h}}$ is formed by (12) but with $\tilde{\boldsymbol{h}}_{r,j} =$

$[\tilde{h}_{r,j}(0),\ldots,\tilde{h}_{r,j}(\tilde{L}-1)]$, $\tilde{\mathcal{X}}_n^{(p)}$ is supposed to be formed as $\tilde{\mathcal{X}}_n^{(p)} = \boldsymbol{I}_{N_r} \otimes \tilde{\boldsymbol{X}}_n^{(p)}$ with

$$\tilde{\boldsymbol{X}}_n^{(p)} = \left[\mathrm{diag}(\boldsymbol{x}_{n,1}^{(p)})\boldsymbol{F}_{\tilde{L}}, \mathrm{diag}(\boldsymbol{x}_{n,2}^{(p)})\boldsymbol{F}_{\tilde{L}}, \ldots, \mathrm{diag}(\boldsymbol{x}_{n,N_t}^{(p)})\boldsymbol{F}_{\tilde{L}}\right],$$

and the covariance matrix $\boldsymbol{C}$ becomes

$$\tilde{\boldsymbol{C}} = \sum_{j=1}^{N_t} \sigma_{x_j}^2 \tilde{\boldsymbol{\lambda}}_j \tilde{\boldsymbol{\lambda}}_j^H + \sigma_v^2 \boldsymbol{I}_{KN_r}, \tag{17}$$

where the system matrix $\tilde{\boldsymbol{\lambda}}$ is defined as

$$\tilde{\boldsymbol{\lambda}} = \left[\tilde{\boldsymbol{\lambda}}_1, \tilde{\boldsymbol{\lambda}}_2, \ldots, \tilde{\boldsymbol{\lambda}}_{N_t}\right], \quad \tilde{\boldsymbol{\lambda}}_j = \left[\tilde{\boldsymbol{\lambda}}_{1,j}, \tilde{\boldsymbol{\lambda}}_{2,j}, \ldots, \tilde{\boldsymbol{\lambda}}_{N_r,j}\right]^\top, \tag{18}$$

with $\tilde{\boldsymbol{\lambda}}_{r,j} = \mathrm{diag}(\boldsymbol{F}_{\tilde{L}} \tilde{\boldsymbol{h}}_{r,j})$.
The misspecified log-likehood function is then given by

$$\ell\big(\boldsymbol{y}_p, \boldsymbol{y}_d|\boldsymbol{\theta}\big) \triangleq \log\left(\tilde{f}(\boldsymbol{y}^{(p)}, \boldsymbol{y}^{(d)}|\boldsymbol{\theta})\right) = \ell_p(\boldsymbol{y}_p|\boldsymbol{\theta}) + \ell_d\big(\boldsymbol{y}_d|\boldsymbol{\theta}\big),$$

where

$$\ell_p(\boldsymbol{y}_p|\boldsymbol{\theta}) = -N_r N_p K \log(\pi\sigma_v^2) - \sum_{n=1}^{N_p} \frac{1}{\sigma_v^2} \big\|\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big\|_2^2, \tag{19}$$

$$\ell_d\big(\boldsymbol{y}_d|\boldsymbol{\theta}\big) = -\sum_{n=1}^{N_d} \log\left(\pi^{PN} \det \tilde{\boldsymbol{C}}\right) - \sum_{n=1}^{N_d} \boldsymbol{y}_n^{(d)H} \tilde{\boldsymbol{C}}^{-1} \boldsymbol{y}_n^{(d)}. \tag{20}$$

Since the *pseudo-true* parameter $\boldsymbol{\theta}_{pt}$ is derived from $\max_{\boldsymbol{\theta}} \mathbb{E}_f\big\{\ell\big(\boldsymbol{y}_p, \boldsymbol{y}_d|\boldsymbol{\theta}\big)\big\}$, we obtain

$$\mathbb{E}_f\big\{\nabla\ell_p\big(\boldsymbol{y}_p|\boldsymbol{\theta}\big)\big\}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{pt}} = -\mathbb{E}_f\big\{\nabla\ell_d\big(\boldsymbol{y}_d|\boldsymbol{\theta}\big)\big\}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_{pt}}. \tag{21}$$

As a result, the two factor matrices of the MCRB can be decomposed into two parts as follows

$$\begin{aligned}\boldsymbol{J}_{\mathrm{SB}}(\boldsymbol{\theta}) &= \mathbb{E}_f\left\{\frac{\partial \ell_p}{\partial \boldsymbol{\theta}^*}\left(\frac{\partial \ell_p}{\partial \boldsymbol{\theta}^*}\right)^H\right\} + \mathbb{E}_f\left\{\frac{\partial \ell_d}{\partial \boldsymbol{\theta}^*}\left(\frac{\partial \ell_d}{\partial \boldsymbol{\theta}^*}\right)^H\right\} \\ &\triangleq \boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta}) + \boldsymbol{J}_{\mathrm{D}}(\boldsymbol{\theta}),\end{aligned} \tag{22}$$

$$\begin{aligned}\boldsymbol{A}_{\mathrm{SB}}(\boldsymbol{\theta}) &= \mathbb{E}_f\left\{\frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial \ell_p}{\partial \boldsymbol{\theta}^*}\right)\right\} + \mathbb{E}_f\left\{\frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{\partial \ell_d}{\partial \boldsymbol{\theta}^*}\right)\right\} \\ &\triangleq \boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta}) + \boldsymbol{A}_{\mathrm{D}}(\boldsymbol{\theta}),\end{aligned} \tag{23}$$

where $\boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta})$ and $\boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta})$ are dedicated to $\mathrm{MCRB}_{\mathrm{P}}(\boldsymbol{\theta})$ for the pilot symbols, while $\boldsymbol{J}_{\mathrm{D}}(\boldsymbol{\theta})$ and $\boldsymbol{A}_{\mathrm{D}}(\boldsymbol{\theta})$ concern the misspecified lower bound corresponding to the data symbols.

A. Pilot-based MCRB

Under the assumption that the noise is supposed to be i.i.d. random variables, $\boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta})$ and $\boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta})$ can be expressed as

$$\boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta}) = \sum_{n=1}^{N_p} \boldsymbol{J}_n(\boldsymbol{\theta}), \text{ and } \boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta}) = \sum_{n=1}^{N_p} \boldsymbol{A}_n(\boldsymbol{\theta}), \tag{24}$$

where $\boldsymbol{J}_n(\boldsymbol{\theta})$ and $\boldsymbol{A}_n(\boldsymbol{\theta})$ are corresponding to the n-th pilot OFDM symbol. Moreover, the misspecified log-likelihood function $\ell_p(\boldsymbol{y}_p|\boldsymbol{\theta})$ is also decomposed into

$$\ell_p(\boldsymbol{y}_p|\boldsymbol{\theta}) = \sum_{n=1}^{N_p} \ell_n(\boldsymbol{y}_n^{(p)}|\boldsymbol{\theta}), \text{ where} \tag{25}$$

$$\ell_n(\boldsymbol{y}_n^{(p)}|\boldsymbol{\theta}) = -N_r K \log(\pi\sigma_v^2) - \frac{1}{\sigma_v^2}\big\|\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big\|_2^2, \tag{26}$$

In this case, the vector of unknown parameters is given by $\boldsymbol{\theta} = \left[\boldsymbol{\psi}^\top, \sigma_v^2\right]^\top$. Accordingly, we can express

$$\boldsymbol{J}_n(\boldsymbol{\theta}) = \begin{bmatrix}\boldsymbol{J}_n(\boldsymbol{\psi},\boldsymbol{\psi}) & \boldsymbol{J}_n(\boldsymbol{\psi},\sigma_v^2) \\ \boldsymbol{J}_n(\sigma_v^2,\boldsymbol{\psi}) & \boldsymbol{J}_n(\sigma_v^2,\sigma_v^2)\end{bmatrix}, \tag{27}$$

$$\boldsymbol{A}_n(\boldsymbol{\theta}) = \begin{bmatrix}\boldsymbol{A}_n(\boldsymbol{\psi},\boldsymbol{\psi}) & \boldsymbol{A}_n(\boldsymbol{\psi},\sigma_v^2) \\ \boldsymbol{A}_n(\sigma_v^2,\boldsymbol{\psi}) & \boldsymbol{A}_n(\sigma_v^2,\sigma_v^2)\end{bmatrix}. \tag{28}$$

The derivative $\ell_n(\boldsymbol{y}_p|\boldsymbol{\theta})$ w.r.t. $\boldsymbol{\psi}$ and σ_v^2 are given by

$$\frac{\partial \ell_n}{\partial \boldsymbol{\psi}} = \frac{-1}{\sigma_v^2}\left(\frac{\partial \tilde{\boldsymbol{m}}_n^{(p)}}{\partial \boldsymbol{\psi}}\right)^H \big(\boldsymbol{y} - \tilde{\boldsymbol{m}}_n^{(p)}\big), \tag{29}$$

$$\frac{\partial \ell_n}{\partial \sigma_v^2} = \frac{1}{\sigma_v^4}\big\|\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big\|_2^2 - \frac{N_r K}{\sigma_v^2}. \tag{30}$$

To facilitate the calculation of $\boldsymbol{J}_n(\boldsymbol{\theta})$ and $\boldsymbol{A}_n(\boldsymbol{\theta})$, let us take a closer look at the derivation of $\tilde{\boldsymbol{m}}_n^{(p)}$ w.r.t. $\boldsymbol{\psi}$

$$\frac{\partial \tilde{\boldsymbol{m}}_n^{(p)}}{\partial \boldsymbol{\psi}} = \mathcal{X}_n^{(p)}\left[\frac{\partial \tilde{\boldsymbol{h}}}{\partial \boldsymbol{\beta}}, \frac{\partial \tilde{\boldsymbol{h}}}{\partial \boldsymbol{\tau}}, \frac{\partial \tilde{\boldsymbol{h}}}{\partial \boldsymbol{\alpha}}\right] \triangleq \mathcal{X}_n^{(p)} \boldsymbol{G}. \tag{31}$$

More specifically, we have

$$\frac{\partial \tilde{\boldsymbol{h}}}{\partial \boldsymbol{\beta}} = \left[\boldsymbol{B}_1^\top, \quad \boldsymbol{B}_2^\top, \quad \ldots, \quad \boldsymbol{B}_{N_r}^\top\right]^\top, \tag{32}$$

$$\boldsymbol{B}_r = \mathrm{diag}\left(\left[\boldsymbol{B}_{r,1}, \ \boldsymbol{B}_{r,2}, \ \ldots, \ \boldsymbol{B}_{r,N_t}\right]\right), \tag{33}$$

$$\boldsymbol{B}_{r,j} = \begin{bmatrix}\frac{\partial \tilde{h}_{r,j}(0)}{\partial \beta_{1,j}} & \frac{\partial \tilde{h}_{r,j}(1)}{\partial \beta_{1,j}} & \cdots & \frac{\partial \tilde{h}_{r,j}(\tilde{L}-1)}{\partial \beta_{1,j}} \\ \frac{\partial \tilde{h}_{r,j}(0)}{\partial \beta_{2,j}} & \frac{\partial \tilde{h}_{r,j}(1)}{\partial \beta_{2,j}} & \cdots & \frac{\partial \tilde{h}_{r,j}(\tilde{L}-1)}{\partial \beta_{2,j}} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{h}_{r,j}(0)}{\partial \beta_{M,j}} & \frac{\partial \tilde{h}_{r,j}(1)}{\partial \beta_{M,j}} & \cdots & \frac{\partial \tilde{h}_{r,j}(\tilde{L}-1)}{\partial \beta_{M,j}}\end{bmatrix}, \tag{34}$$

where $\frac{\partial \tilde{h}_{r,j}(l)}{\partial \beta_{m,j}} = \mathrm{sinc}(l - \tau_{m,j}) \exp\left(-i2\pi \frac{d}{\lambda}(r-1)\sin(\alpha_{m,j})\right)$. The next two derivatives $\partial\tilde{\boldsymbol{h}}/\partial\boldsymbol{\tau}$ and $\partial\tilde{\boldsymbol{h}}/\partial\boldsymbol{\alpha}$ share the same form to $\partial\tilde{\boldsymbol{h}}/\partial\boldsymbol{\beta}$, but respectively with the following entries

$$\begin{aligned}\frac{\partial \tilde{h}_{r,j}(l)}{\partial \tau_{m,j}} &= \beta_{m,j}\left(\frac{\sin(l-\tau_{m,j})}{(\ell-\tau_{m,j})^2} - \frac{\cos(l-\tau_{m,j})}{\ell-\tau_{m,j}}\right) \\ &\qquad \times \exp\left(-i2\pi\frac{d}{\lambda}(r-1)\sin(\alpha_{m,j})\right),\end{aligned} \tag{35}$$

$$\begin{aligned}\frac{\partial \tilde{h}_{r,j}(l)}{\partial \alpha_{m,j}} &= \beta_{m,j}\,\mathrm{sinc}(\ell-\tau_{m,j})\frac{-i2\pi dq\cos(\alpha_{m,j})}{\lambda} \\ &\qquad \times \exp\left(-i2\pi\frac{d}{\lambda}(r-1)\sin(\alpha_{m,j})\right).\end{aligned} \tag{36}$$

Now, let us denote the error mean $\boldsymbol{e}_n^{(p)} = \tilde{\boldsymbol{m}}_n^{(p)} - \boldsymbol{m}_n^{(p)}$ and $\boldsymbol{E}_n^{(p)} = \sigma_v^2 \boldsymbol{I}_{N_r K} + \boldsymbol{e}_n^{(p)} \boldsymbol{e}_n^{(p)H}$, we have

$$\begin{aligned}&\mathbb{E}_f\big\{\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big\} = \boldsymbol{e}_n^{(p)}, \\ &\mathbb{E}_f\big\{\big(\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big)\big(\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big)^\top\big\} = \boldsymbol{e}_n^{(p)} \boldsymbol{e}_n^{(p)\top}, \\ &\mathbb{E}_f\big\{\big(\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big)\big(\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\big)^H\big\} = \boldsymbol{E}_n^{(p)}, \\ &\mathbb{E}_f\big\{\|\boldsymbol{y}_n^{(p)} - \tilde{\boldsymbol{m}}_n^{(p)}\|_2^2\big\} = \sigma_v^2 PK + \|\boldsymbol{e}_n^{(p)}\|_2^2.\end{aligned}$$

Accordingly, we obtain

$$\boldsymbol{J}_n(\sigma_v^2,\sigma_v^2) = -\boldsymbol{A}_n(\sigma_v^2,\sigma_v^2) = -\frac{N_r K}{\sigma_v^4}, \tag{37}$$

$$\boldsymbol{J}_n(\boldsymbol{\psi},\sigma_v^2) = (\boldsymbol{J}_{\boldsymbol{\psi},\sigma_v^2})^H = \boldsymbol{0}, \tag{38}$$

$$\boldsymbol{A}_n(\boldsymbol{\psi},\sigma_v^2) = (\boldsymbol{A}_{\boldsymbol{\psi},\sigma_v^2})^H = \frac{1}{\sigma_v^4}\boldsymbol{G}^H{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{e}, \tag{39}$$

$$\boldsymbol{J}_n(\boldsymbol{\psi},\boldsymbol{\psi}) = -\frac{1}{\sigma_v^4}\boldsymbol{G}^H{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{E}_n^{(p)}\boldsymbol{\mathcal{X}}_n^{(p)}\boldsymbol{G}, \tag{40}$$

$$\boldsymbol{A}_n(\boldsymbol{\psi},\boldsymbol{\psi}) = \frac{1}{\sigma_v^2}\boldsymbol{G}^H{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{\mathcal{X}}_n^{(p)}\boldsymbol{G}. \tag{41}$$

In the case of training-based estimation, the $\mathrm{MCRB}_{\mathrm{OP}}(\boldsymbol{\theta})$ ("OP" standards for "only pilot") is given by

$$\mathrm{MCRB}_{\mathrm{OP}}(\boldsymbol{\theta}) = \boldsymbol{A}_{\mathrm{OP}}^{\#}(\boldsymbol{\theta})\boldsymbol{J}_{\mathrm{OP}}(\boldsymbol{\theta})\boldsymbol{A}_{\mathrm{OP}}^{\#}(\boldsymbol{\theta}), \tag{42}$$

where

$$\boldsymbol{J}_{\mathrm{OP}}(\boldsymbol{\theta}) = \frac{-1}{\sigma_v^4}\begin{bmatrix}\boldsymbol{G}^H\left(\sum_{n=1}^{N_p}{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{E}\boldsymbol{\mathcal{X}}_n^{(p)}\right)\boldsymbol{G} & \boldsymbol{0}\\ \boldsymbol{0} & N_pN_rK\end{bmatrix}, \tag{43}$$

$$\boldsymbol{A}_{\mathrm{OP}}(\boldsymbol{\theta}) = \frac{1}{\sigma_v^2}\begin{bmatrix}\boldsymbol{G}^H\left(\sum_{n=1}^{N_p}{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{\mathcal{X}}_n^{(p)}\right)\boldsymbol{G} & \boldsymbol{G}^H{\boldsymbol{\mathcal{X}}_n^{(p)}}^H\boldsymbol{e}\\ \boldsymbol{e}^H\boldsymbol{\mathcal{X}}_n^{(p)}\boldsymbol{G} & \dfrac{N_pN_rK}{\sigma_v^2}\end{bmatrix}. \tag{44}$$

When dealing with semi-blind channel estimation, the two matrices $\boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta})$ and $\boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta})$ are derived from

$$\boldsymbol{J}_{\mathrm{P}}(\boldsymbol{\theta}) = \begin{bmatrix}\boldsymbol{J}_{\mathrm{OP}}(\boldsymbol{\theta}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}_{\boldsymbol{\sigma}_x^2}\end{bmatrix}, \boldsymbol{A}_{\mathrm{P}}(\boldsymbol{\theta}) = \begin{bmatrix}\boldsymbol{A}_{\mathrm{OP}}(\boldsymbol{\theta}) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}_{\boldsymbol{\sigma}_x^2}\end{bmatrix}. \tag{45}$$

B. Data-based MCRB

We know that the true distribution of $\boldsymbol{y}_n^{(d)}$ is $\mathcal{CN}(\boldsymbol{0},\boldsymbol{C})$ while the users assume $\boldsymbol{y}_n^{(d)} \sim \mathcal{CN}(\boldsymbol{0},\tilde{\boldsymbol{C}})$, where $\tilde{\boldsymbol{C}}$ is defined as in (17). The vector of unknown parameters is now $\boldsymbol{\theta} = \left[\boldsymbol{\psi}^{\top},\sigma_v^2,\boldsymbol{\sigma}_x^2\right]^{\top}$.
The derivative of $\ell_d(\boldsymbol{y}_d|\boldsymbol{\theta})$ is given by

$$\frac{\partial\ell_d}{\partial\theta_i} = -N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\right\} + \sum_{n=1}^{N_d}{\boldsymbol{y}_n^{(d)}}^H\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\tilde{\boldsymbol{C}}^{-1}\boldsymbol{y}_n^{(d)}. \tag{46}$$

where $\frac{\partial\tilde{\boldsymbol{C}}}{\partial\beta_i} = \sigma_x^2\tilde{\boldsymbol{\lambda}}\frac{\partial\tilde{\boldsymbol{\lambda}}^H}{\partial\tilde{\boldsymbol{h}}}\frac{\partial\tilde{\boldsymbol{h}}}{\partial\beta_i}$, $\frac{\partial\tilde{\boldsymbol{C}}}{\partial\tau_i} = \sigma_x^2\tilde{\boldsymbol{\lambda}}\frac{\partial\tilde{\boldsymbol{\lambda}}^H}{\partial\tilde{\boldsymbol{h}}}\frac{\partial\tilde{\boldsymbol{h}}}{\partial\tau_i}$, $\frac{\partial\tilde{\boldsymbol{C}}}{\partial\alpha_i} = \sigma_x^2\tilde{\boldsymbol{\lambda}}\frac{\partial\tilde{\boldsymbol{\lambda}}^H}{\partial\tilde{\boldsymbol{h}}}\frac{\partial\tilde{\boldsymbol{h}}}{\partial\tau_{\alpha_i}}$, $\frac{\partial\tilde{\boldsymbol{C}}}{\sigma_{x_j}^2} = \tilde{\boldsymbol{\lambda}}_j\tilde{\boldsymbol{\lambda}}_j^H$, and $\frac{\partial\tilde{\boldsymbol{C}}}{\sigma_v^2} = \boldsymbol{I}_{N_rK}$.
At $\boldsymbol{\theta} = \boldsymbol{\theta}_{pt}$, we obtain $\mathbb{E}_f\{\frac{\partial\ell_d}{\partial\theta_i}\} = 0$ and hence

$$N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\right\} = \sum_{n=1}^{N_d}\mathbb{E}_f\left\{{\boldsymbol{y}_n^{(d)}}^H\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\tilde{\boldsymbol{C}}^{-1}\boldsymbol{y}_n^{(d)}\right\}. \tag{47}$$

Accordingly, taking the expectation of $\{\frac{\partial\ell_d}{\partial\theta_i}\frac{\partial\ell_d}{\partial\theta_j}\}$ and $\{\frac{\partial^2\ell_d}{\partial\theta_i\partial\theta_j}\}$ over the true $f(\boldsymbol{y})$ result in

$$\begin{aligned}[\boldsymbol{J}_{\mathrm{D}}(\boldsymbol{\theta})]_{ij} &= N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_j}\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}\right\} \\ &+N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\left(\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}-\boldsymbol{I}\right)\right\}\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_j}\left(\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}-\boldsymbol{I}\right)\right\},\end{aligned} \tag{48}$$

$$\begin{aligned}[\boldsymbol{A}_{\mathrm{D}}(\boldsymbol{\theta})]_{ij} &= -N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_j}\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\left(\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}-\boldsymbol{I}\right)\right\} \\ &+N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial^2\tilde{\boldsymbol{C}}}{\partial\theta_j\partial\theta_i}\left(\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}-\boldsymbol{I}\right)\right\} - N_d\,\mathrm{tr}\left\{\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_i}\tilde{\boldsymbol{C}}^{-1}\frac{\partial\tilde{\boldsymbol{C}}}{\partial\theta_j}\tilde{\boldsymbol{C}}^{-1}\boldsymbol{C}\right\}.\end{aligned} \tag{49}$$

C. Transformation

It is desired to estimate the channel coefficients in time domain: $\tilde{\boldsymbol{h}} = g(\boldsymbol{\theta})$ where $g(\boldsymbol{\theta})$ is defined in (7). More concretely, the MCRB for channel taps can be expressed as follows

$$\mathrm{MCRB}_{\mathrm{SB}}\left(\tilde{\boldsymbol{h}}\right) = \boldsymbol{A}_{\mathrm{SB}}^{\#}\big(g(\boldsymbol{\theta})\big)\boldsymbol{J}_{\mathrm{SB}}\big(g(\boldsymbol{\theta})\big)\boldsymbol{A}_{\mathrm{SB}}^{\#}\big(g(\boldsymbol{\theta})\big). \tag{50}$$

Thanks to the derivative chain rule, we have

$$\frac{\partial\ell(\boldsymbol{y}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \boldsymbol{T}^H(\boldsymbol{\theta})\frac{\partial\ell\big(\boldsymbol{y}|g(\boldsymbol{\theta})\big)}{\partial g(\boldsymbol{\theta})}, \tag{51}$$

where $\boldsymbol{T}(\boldsymbol{\theta})$ is the Jacobian matrix of the transformation

$$\boldsymbol{T}(\boldsymbol{\theta}) \triangleq \frac{\partial g(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \begin{bmatrix}\boldsymbol{G} & \boldsymbol{0}_{\sigma_v^2,\boldsymbol{\sigma}_x^2}\end{bmatrix}, \tag{52}$$

where $\partial g(\boldsymbol{\theta})/\partial\sigma_v^2 = \boldsymbol{0}$, $\partial g(\boldsymbol{\theta})/\partial\boldsymbol{\sigma}_x^2 = \boldsymbol{0}$, and $\boldsymbol{G} = \partial g(\boldsymbol{\psi})/\partial\boldsymbol{\psi}$ is already provided in (31).
Accordingly, the matrix $\boldsymbol{J}_{\mathrm{SB}}(\boldsymbol{\theta})$ can be represented by

$$\begin{aligned}\boldsymbol{J}_{\mathrm{SB}}(\boldsymbol{\theta}) &= \mathbb{E}_f\left\{\boldsymbol{T}^H(\boldsymbol{\theta})\frac{\partial\ell\big(\boldsymbol{y}|g(\boldsymbol{\theta})\big)}{\partial g(\boldsymbol{\theta})}\left(\frac{\partial\ell\big(\boldsymbol{y}|g(\boldsymbol{\theta})\big)}{\partial g(\boldsymbol{\theta})}\right)^H\boldsymbol{T}(\boldsymbol{\theta})\right\} \\ &= \boldsymbol{T}^H(\boldsymbol{\theta})\boldsymbol{J}_{\mathrm{SB}}\big(g(\boldsymbol{\theta})\big)\boldsymbol{T}(\boldsymbol{\theta}).\end{aligned} \tag{53}$$

Similarly, we also derive

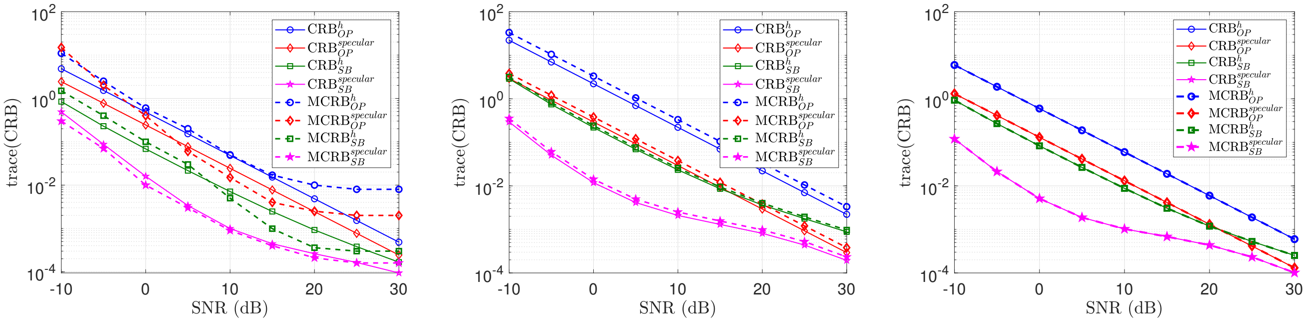
$$\boldsymbol{A}_{\mathrm{SB}}(\boldsymbol{\theta}) = \boldsymbol{T}^H(\boldsymbol{\theta})\boldsymbol{A}_{\mathrm{SB}}\big(g(\boldsymbol{\theta})\big)\boldsymbol{T}(\boldsymbol{\theta}). \tag{54}$$

Finally, the expression (50) becomes

$$\begin{aligned}\mathrm{MCRB}_{\mathrm{SB}}\left(\tilde{\boldsymbol{h}}\right) &= \boldsymbol{T}(\boldsymbol{\theta})\boldsymbol{A}_{\mathrm{SB}}^{\#}(\boldsymbol{\theta})\boldsymbol{T}^H(\boldsymbol{\theta})(\boldsymbol{T}^H(\boldsymbol{\theta}))^{\#} \\ &\quad\times\boldsymbol{J}_{\mathrm{SB}}(\boldsymbol{\theta})\boldsymbol{T}^{\#}(\boldsymbol{\theta})\boldsymbol{T}(\boldsymbol{\theta})\boldsymbol{A}_{\mathrm{SB}}^{\#}(\boldsymbol{\theta})\boldsymbol{T}^H(\boldsymbol{\theta}) \\ &= \boldsymbol{T}(\boldsymbol{\theta})\,\mathrm{MCRB}_{\mathrm{SB}}(\boldsymbol{\theta})\boldsymbol{T}^H(\boldsymbol{\theta}).\end{aligned} \tag{55}$$

V. RESULTS & DISCUSSIONS

In this section, we illustrate the behavior of the proposed MCRBs in two scenarios: underspecification and overspecification. We compare the MCRBs with the classical CRBs which are derived from the assumed model without being aware of the misspecification to demonstrate the validity of the proposed bounds. We follow the experiment setup of Rekik *et al.* in [10] and simulation parameters are specified in Tab. I.
Fig. 1(a) plots the trace of (M)CRBs w.r.t. the channel parameters versus SNR $= -10\log_{10}(\sigma_v^2)$ in the presence of channel order underestimation. First, we can see that the proposed MCRBs tend to converge towards error levels as SNR increases, probably because there exist a nonzero error mean between the true and the assumed one. It suggests that the performance of channel estimators can not exceed a deterministic threshold even when SNR goes to infinity. By contrast, the classical CRBs are approximately proportional to the noise variance which do not reflect the true performance bounds of channel estimators under the channel order underspecification.



(a) Underestimation $\tilde{L} = 3 < L_{tr} = 4$ (b) Overestimation $\tilde{L} = 5 > L_{tr} = 4$ (c) Perfect Specification $\tilde{L} = L_{tr} = 4$

Fig. 1: Proposed MCRB bounds for semi-blind channel estimation.

TABLE I: Simulation parameters

Parameters	Specifications
MIMO	$N_t = 2,\ N_r = 3$
Antenna	$d/\lambda = 1/2$
OFDM subcarriers	$K = 64$
Multi-paths	$M = 4$
Channel order	$L = 4$
Number of pilot symbols	$N_p = 1$
Number of data symbols	$N_d = 40$
Channel fading	$\boldsymbol{\beta} = [0.8, 0.6; 0.4, 0.2; 0.9, 0.7; 0.5, 0.3]$;
DOA	$\boldsymbol{\alpha} = [\pi/2, \pi/4; \pi/6, \pi/8; \pi/3, \pi/5; \pi/7, \pi/9]$
Time delay	$\boldsymbol{\tau} = [0.1, 0.2; 0.3, 0.4; 0.2, 0.3; 0.4, 0.5]$
Signal power	$\boldsymbol{\sigma}_x^2 = [1, 1]$

Second, we can notice that introducing the prior specular parameters (DOA, time delay and fading) will gain the theoretical performance limit of channel estimators despite the channel order is underspecified. Indeed, such information can be seen as constraints imposed on the channel taps, hence the resulting $\mathrm{MCRB}^{\text{specular}}$ can be referred to as a constrained MCRB which is often lower than the unconstrained bound. Third, the MCRBs for semi-blind estimation are lower than ones for training-based estimation to demonstrate the effectiveness of semi-blind approaches.

Fig. 1(b) illustrates the performance lower bounds when the channel order is overestimated. In fact, the overspecification is not really a case of misspecification, as indicated in our companion work [12]. We can see that the proposed MCRBs follow the classical CRBs in this oversepcification. The MCRBs become the CRBs when the channel order is perfectly specified, as shown in Fig. 1(c).

VI. Conclusions

In this paper, we considered the problem of analyzing the theoretical performance limit of semi-blind channel estimation techniques with side information under the channel order misspecification. Two closed-form expressions of the MCRB were derived for unbiased training-based estimators and semi-blind estimators. Numerical experiments were provided to illustrate the validity of the proposed bounds.

VII. Acknowledgment

This work was funded by by the National Foundation for Science and Technology Development of Vietnam (Grant No. 01/2019/TN).

References

[1] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Commun. Surv. Tutor.*, vol. 9, no. 2, pp. 18–48, 2007.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice Hall PTR, 1993.

[3] E. de Carvalho and D. T. M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *IEEE Works. Signal Process. Adv. Wirel. Commun.*, 1997, pp. 129–132.

[4] L. Berriche, K. Abed-Meraim, and J. C. Belfiore, "Cramer-Rao bounds for MIMO channel estimation," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, 2004, pp. 397–400.

[5] S. Abdallah and I. N. Psaromiligkos, "Exact Cramer–Rao bounds for semiblind channel estimation in amplify-and-forward two-way relay networks employing square QAM," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 12, pp. 6955–6967, 2014.

[6] A. Ladaycia, A. Mokraoui, K. Abed-Meraim, and A. Belouchrani, "Performance bounds analysis for semi-blind channel estimation in MIMO-OFDM communications systems," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 9, pp. 5925–5938, 2017.

[7] E. Nayebi and B. D. Rao, "Semi-blind channel estimation for multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 540–553, 2018.

[8] O. Rekik, A. Ladaycia, A. Mokraoui, and K. Abed-Meraim, "CRB-based performance analysis of semi-blind channel estimation for massive MIMO-OFDM systems with pilot contamination," *IET Commun.*, vol. 13, no. 20, pp. 3479–3488, 2019.

[9] Q. Mayyala, K. Abed-Meraim, and A. Zerguine, "On the performance evaluation of blind system identification in presence of side information," in *Int. Wirel. Comm. Mobile Comput. Conf.*, 2017, pp. 218–222.

[10] O. Rekik, A. Mokraoui, T. Quynh, and K. Abed-Meraim, "Performance bound analysis of side information effect on MIMO-OFDM channels semi-blind identification," in *European Signal Process. Conf.*, 2021 (submitted).

[11] L. T. Thanh, K. Abed-Meraim, and N. Linh-Trung, "Performance lower bounds of blind system identification techniques in the presence of channel order estimation error," in *European Signal Process. Conf.*, 2021 (submitted).

[12] ——, "Misspecified cramer-rao bounds for blind channelestimation under channel order misspecification," *IEEE Trans. Signal Process.*, 2021 (under review).

[13] Q. H. Vuong, "Cramer-Rao bounds for misspecified models," *Caltech*, 1986.

[14] C. D. Richmond and L. L. Horowitz, "Parameter bounds on estimation accuracy under model misspecification," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2263–2278, 2015.

[15] S. Fortunati, F. Gini, M. S. Greco, and C. D. Richmond, "Performance bounds for parameter estimation under misspecified models: Fundamental findings and applications," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 142–157, 2017.

[16] H. White, "Maximum likelihood estimation of misspecified models," *Econometrica*, pp. 1–25, 1982.