Chapter 3

Separable Parameterization

As seen in the topic, the key behind thesis idea is to separate the model parameter in an appropriate way. This task will be described in section 3.1. To investigate the inherent accuracy limitation, we provide a systematic formulation of Cramér-Rao bound in section 3.2. The expression is based on separable parameterization. Since this is a variational result of the lemma 2, it is thus easy to be computed.

Regarding the parameter estimation, an asymptotic maximum likelihood estimator is derived in section 3.3. The purpose of proposed estimator is to outperform previous sub-optimal estimators, especially in non-asymptotic performance. Its asymptotic performances is derived in section 3.4 and then indicated that it holds asymptotic efficiency with respect to the CRB.

As explored the structured covariance matrix, we provide the existence of Toeplitz structure in local scattering model via theorem 2. Next we also extend the idea of AML in section 3.5.3 to the problem of estimating the Toeplitz-Hermitian covariance matrix. In section 3.5, the relationships among the RA, WCM and the proposed AML approaches are provided by two independent lemmas in sections 3.5.2 and 3.5.4. These enable us to a deep insight not only into their methodology connection but also the achievable performance. Finally, we apply the covariance Toeplitzifications into WLS and AML estimators in section 3.6 to reduce the directional estimate errors.

3.1 Separable Parameterization

In this section, we shall separate the parameter into two portions, such as, parameter of interest $\boldsymbol{\vartheta}_\omega(\boldsymbol{\omega}, \boldsymbol{\sigma}_\omega) : \mathbb{R}^{(2N_S)\times 1} \mapsto \mathbb{R}^{(2N_S)\times 1}$ and nuisance parameter $\boldsymbol{\eta}(\boldsymbol{p}, \sigma_n^2) : \mathbb{R}^{(N_S+1)\times 1} \mapsto \mathbb{R}^{(N_S+1)\times 1}$. The reason why we proceed on this is that the unwanted parameters $\boldsymbol{p}$ and σ_n^2 would be eliminated from likelihood function. Let us partition $\boldsymbol{\theta}_\omega$ according to

$$\boldsymbol{\vartheta}_\omega(\boldsymbol{\omega}, \boldsymbol{\sigma}_\omega) \triangleq \begin{bmatrix}\boldsymbol{\omega}^\mathsf{T} & \boldsymbol{\sigma}_\omega^\mathsf{T}\end{bmatrix}^\mathsf{T} \tag{3.1a}$$

$$\boldsymbol{\eta}(\boldsymbol{p}, \sigma_n^2) \triangleq \begin{bmatrix}\boldsymbol{p}^\mathsf{T} & \sigma_n^2\end{bmatrix}^\mathsf{T}. \tag{3.1b}$$

This exhibits a separable form of $\boldsymbol{\theta}_\omega$ as

$$\boldsymbol{\theta}_\omega = \begin{bmatrix}\boldsymbol{\vartheta}_\omega^\mathsf{T} & \boldsymbol{\eta}^\mathsf{T}\end{bmatrix}^\mathsf{T}. \tag{3.2}$$

Next the column-stacking vectorization operator $\boldsymbol{v}_\mathsf{c}(\cdot)$ allows us to represent the vector $\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega) \triangleq \mathcal{E}\left\langle \boldsymbol{x}^*[n_T] \otimes \boldsymbol{x}[n_T]\right\rangle = \boldsymbol{v}_\mathsf{c}\left(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\omega)\right) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{C}^{N_E^2\times 1}$ in such a way that

$$\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega) = \boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\boldsymbol{\eta}(\boldsymbol{p}, \sigma_n^2) \tag{3.3}$$

where $\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega) : \mathbb{R}^{(2N_S)\times 1} \mapsto \mathbb{C}_{\mathbb{F}}^{N_E^2\times(N_S+1)}$ is defined by

$$\begin{aligned}\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega) \triangleq \Big[\ & \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{D}_a(\omega_1)\boldsymbol{B}(\sigma_{\omega_1})\boldsymbol{D}_a^{\mathsf{H}}(\omega_1)) \quad \cdots \\ & \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{D}_a(\omega_{N_S})\boldsymbol{B}(\sigma_{\omega_{N_S}})\boldsymbol{D}_a^{\mathsf{H}}(\omega_{N_S})) \quad \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{I})\ \Big].\end{aligned} \tag{3.4}$$

Note that the covariance vector $\boldsymbol{\xi}_x(\boldsymbol{\theta}_\omega)$ is now linear in nuisance vector $\boldsymbol{\eta}(\boldsymbol{p}, \sigma_n^2)$. This indeed amounts $2N_S$-dimensional parameterization in $\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)$.

3.2 A Formulation of Cramér-Rao Bound

Recall the Slepian-Bangs formula accounting for the zero-mean random vector $\boldsymbol{x}[n_T]$. This leads to the $(n, \acute{n})$-th element of *F*isher *i*nformation *m*atrix (FIM) according to [20]

$$[\boldsymbol{I}_{\mathsf{F}}(\boldsymbol{\theta})]_{[n,\acute{n}]} = N_T \lceil \boldsymbol{\Sigma}_{xx}^{-1}\dot{\boldsymbol{\Sigma}}_{xx}(\theta_n)\boldsymbol{\Sigma}_{xx}^{-1}\dot{\boldsymbol{\Sigma}}_{xx}(\theta_{\acute{n}})\rfloor \tag{3.5}$$

where the scalars θ_n and $\theta_{\acute{n}}$ are the n-th and $\acute{n}$-th elements of the parameter vector $\boldsymbol{\theta}$ for indices $n, \acute{n} \in \{1, 2, \ldots, 3N_S + 1\}$. Let the derivative matrix $\nabla_\theta(\boldsymbol{\xi}_x) \in \mathbb{C}^{N_E^2\times(3N_S+1)}$ of the vector $\boldsymbol{\xi}_x(\boldsymbol{\theta}) \triangleq \boldsymbol{v}_{\mathsf{c}}(\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{C}^{N_E^2\times 1}$ be given by

$$\nabla_\theta(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial\boldsymbol{\theta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\theta}) = \left[\dot{\boldsymbol{\xi}}(\theta_1) \ \cdots \ \dot{\boldsymbol{\xi}}(\theta_{3N_S+1})\right]. \tag{3.6}$$

One can fulfill the FIM (see *e.g.*, [35] and [50]) as

$$\boldsymbol{I}_{\mathsf{F}}(\boldsymbol{\theta}) = N_T \nabla_\theta^{\mathsf{H}}(\boldsymbol{\xi}_x)\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta})\nabla_\theta(\boldsymbol{\xi}_x) \tag{3.7}$$

where $\boldsymbol{\Psi}_{xx} \triangleq \boldsymbol{\Sigma}_{xx}^{\mathsf{T}} \otimes \boldsymbol{\Sigma}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E^2\times N_E^2}$. Straightforward calculating the derivatives and $\nabla_\eta(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial\boldsymbol{\eta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\eta})$, we obtain [19]

$$\dot{\boldsymbol{\xi}}(\omega_{n_S}) = p_{n_S}\boldsymbol{v}_{\mathsf{c}}\left(\dot{\boldsymbol{D}}_a(\omega_{n_S})\boldsymbol{B}(\sigma_{\omega_{n_S}})\boldsymbol{D}_a^{\mathsf{H}}(\omega_{n_S})\right) + p_{n_S}\boldsymbol{v}_{\mathsf{c}}\left(\boldsymbol{D}_a(\omega_{n_S})\boldsymbol{B}(\sigma_{\omega_{n_S}})\dot{\boldsymbol{D}}_a^{\mathsf{H}}(\omega_{n_S})\right) \tag{3.8a}$$

$$\dot{\boldsymbol{\xi}}(\sigma_{\omega_{n_S}}) = p_{n_S}\boldsymbol{v}_{\mathsf{c}}\left(\boldsymbol{D}_a(\omega_{n_S})\dot{\boldsymbol{B}}(\sigma_{\omega_{n_S}})\boldsymbol{D}_a^{\mathsf{H}}(\omega_{n_S})\right) \tag{3.8b}$$

$$\nabla_\eta(\boldsymbol{\xi}_x) = \boldsymbol{\Omega}(\boldsymbol{\omega}, \boldsymbol{\sigma}_\omega) \tag{3.8c}$$

where the derivative matrix $\dot{\boldsymbol{D}}_a(\omega_{n_S})$ is available from

$$\dot{\boldsymbol{D}}_a(\omega_{n_S}) = \imath\boldsymbol{\Lambda}\boldsymbol{D}_a(\omega_{n_S}) \tag{3.9}$$

with

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & N_E - 1 \end{bmatrix}. \tag{3.10}$$

Relying on the PDF encountered, the necessary derivative $\boldsymbol{B}$ can be represented in one's stride as follows [19]

$$[\dot{\boldsymbol{B}}(\sigma_{\omega_\phi})]_{[n_E,\acute{n}_E]} = \begin{cases} \frac{1}{\sigma_{\omega_\phi}}\left(\cos((n_E - \acute{n}_E)\sqrt{3}\sigma_{\omega_\phi}) - [\boldsymbol{B}(\sigma_{\omega_\phi})]_{[n_E,\acute{n}_E]}\right) & ;\text{Uniform} \\ -(n_E - \acute{n}_E)^2\sigma_{\omega_\phi}[\boldsymbol{B}(\sigma_{\omega_\phi})]_{[n_E,\acute{n}_E]} & ;\text{Gaussian} \\ -(n_E - \acute{n}_E)^2\sigma_{\omega_\phi}[\boldsymbol{B}(\sigma_{\omega_\phi})]^2_{[n_E,\acute{n}_E]} & ;\text{Laplacian.} \end{cases} \tag{3.11}$$

By accumulating all derivatives into $\nabla_\omega(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial \boldsymbol{\omega}^\mathsf{T}}\boldsymbol{\xi}_x(\boldsymbol{\omega}) \in \mathbb{C}^{N_E^2 \times N_S}$ and $\nabla_{\sigma_\omega}(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial \boldsymbol{\sigma}_\omega^\mathsf{T}}\boldsymbol{\xi}_x(\boldsymbol{\sigma}_\omega) \in \mathbb{C}^{N_E^2 \times N_S}$ as

$$\nabla_\omega(\boldsymbol{\xi}_x) = \left[\dot{\boldsymbol{\xi}}(\omega_1) \;\; \dot{\boldsymbol{\xi}}(\omega_2) \;\; \cdots \;\; \dot{\boldsymbol{\xi}}(\omega_{N_S})\right] \tag{3.12a}$$

$$\nabla_{\sigma_\omega}(\boldsymbol{\xi}_x) = \left[\dot{\boldsymbol{\xi}}(\sigma_{\omega_1}) \;\; \dot{\boldsymbol{\xi}}(\sigma_{\omega_2}) \;\; \cdots \;\; \dot{\boldsymbol{\xi}}(\sigma_{\omega_{N_S}})\right] \tag{3.12b}$$

the derivative $\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x) \triangleq \frac{\partial}{\partial \boldsymbol{\vartheta}_\omega^\mathsf{T}}\boldsymbol{\xi}_x(\boldsymbol{\vartheta}) \in \mathbb{C}^{N_E^2 \times (2N_S)}$ with respect to $\boldsymbol{\vartheta}_\omega$ can be represented by

$$\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x) = \left[\nabla_\omega(\boldsymbol{\xi}_x) \;\; \nabla_{\sigma_\omega}(\boldsymbol{\xi}_x)\right]. \tag{3.13}$$

Then, it yields

$$\nabla_{\theta_\omega}(\boldsymbol{\xi}_x) = \left[\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x) \;\; \boldsymbol{\Omega}(\boldsymbol{\omega}, \boldsymbol{\sigma}_\omega)\right]. \tag{3.14}$$

Owing to the fact that our parameter estimation algorithms are based on spatial frequency approximation model, the most appropriate benchmark should be the Cramér-Rao bound accounted for the physical model parameter, which is of course derived via the spatial frequency model.

Since one can write $\boldsymbol{\theta}_\omega(\boldsymbol{\theta}_\phi)$, the CRB derived from spatial frequency model can be transformed into the physical model via the Jacobian matrix [36, pp. 45–46]. In this work, we may express it as $\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega) \triangleq \frac{\partial}{\partial \boldsymbol{\theta}_\phi^\mathsf{T}}\boldsymbol{\theta}_\omega(\boldsymbol{\theta}_\phi) \in \mathbb{R}^{(3N_S+1)\times(3N_S+1)}$. It can be illustrated that

$$\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega) = \begin{bmatrix} \boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega) & \boldsymbol{O} \\ \boldsymbol{O}^\mathsf{T} & \boldsymbol{I} \end{bmatrix} \tag{3.15}$$

whence we have $\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\eta}) \triangleq \frac{\partial}{\partial \boldsymbol{\vartheta}_\phi^\mathsf{T}}\boldsymbol{\eta} = \boldsymbol{O}^\mathsf{T}$, $\boldsymbol{J}_\eta(\boldsymbol{\vartheta}_\phi) \triangleq \frac{\partial}{\partial \boldsymbol{\eta}^\mathsf{T}}\boldsymbol{\vartheta}_\phi = \boldsymbol{O}$ and $\boldsymbol{J}_\eta(\boldsymbol{\vartheta}_\phi) \triangleq \frac{\partial}{\partial \boldsymbol{\eta}^\mathsf{T}}\boldsymbol{\eta} = \boldsymbol{I}_{(N_S+1)}$. Furthermore, the elements of $\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega) \triangleq \frac{\partial}{\partial \boldsymbol{\vartheta}_\phi^\mathsf{T}}\boldsymbol{\vartheta}_\omega(\boldsymbol{\theta}_\phi) \in \mathbb{R}^{(2N_S)\times(2N_S)}$ are given from [19]

$$
\begin{aligned}
[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[n_S,n_S]} &= \dot{\omega}_{n_S}(\phi_{n_S}) \\
&= kd_E \cos(\phi_{n_S}) && (3.16a)\\
[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[N_S+n_S,n_S]} &= \dot{\sigma}_{\omega_{n_S}}(\phi_{n_S}) \\
&= -kd_E \sigma_{\phi_{n_S}} \sin(\phi_{n_S}) && (3.16b)\\
[\boldsymbol{J}_{\vartheta_\phi}(\boldsymbol{\vartheta}_\omega)]_{[N_S+n_S,N_S+n_S]} &= \dot{\sigma}_{\omega_{n_S}}(\sigma_{\phi_{n_S}}) \\
&= kd_E \cos(\phi_{n_S}). && (3.16c)
\end{aligned}
$$

Therefore, the Jacobian matrix $\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)$ might be calculated by

$$\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ) = \nabla_{\theta_\omega}(\boldsymbol{\xi}_\circ)\boldsymbol{J}_{\theta_\phi}(\boldsymbol{\theta}_\omega). \tag{3.17}$$

In conjunction with $\boldsymbol{B}_{\mathrm{CR}(\hat{\boldsymbol{\theta}})}(\boldsymbol{\theta}_\circ) = \boldsymbol{I}_{\mathrm{F}}^{-1}(\boldsymbol{\theta}_\circ)$ [36], the CRB matrix accounted for estimating $\boldsymbol{\theta}_\phi$ is therefore given by

$$\boldsymbol{B}_{\mathrm{CR}(\hat{\boldsymbol{\theta}}_\phi)}(\boldsymbol{\theta}_\circ) = \frac{1}{N_T}\left(\nabla^{\mathsf{H}}_{\theta_\phi}(\boldsymbol{\xi}_\circ)\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\circ)\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)\right)^{-1} \tag{3.18}$$

where $\nabla_{\theta_\phi}(\boldsymbol{\xi}_\circ)$ denotes $\frac{\partial}{\partial\boldsymbol{\theta}_\phi^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\theta}_\phi)|_{\boldsymbol{\theta}_\phi=\boldsymbol{\theta}_\circ}$. Invoking the separable form of (3.2) so that $\nabla_{\theta_\omega}(\boldsymbol{\xi}_x) = [\,\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_x) \quad \boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\,]$, it follows from block-matrix inversion that

$$\boldsymbol{B}_{\mathrm{CR}(\hat{\boldsymbol{\vartheta}}_\omega)}(\boldsymbol{\theta}_\circ) = \frac{1}{N_T}\left(\nabla^{\mathsf{H}}_{\vartheta_\omega}(\boldsymbol{\xi}_\circ)\,\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}(\boldsymbol{\theta}_\circ)\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}\left(\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}(\boldsymbol{\theta}_\circ)\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\circ)\right)}\boldsymbol{\Psi}_{xx}^{-\frac{1}{2}}(\boldsymbol{\theta}_\circ)\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_\circ)\right)^{-1}. \tag{3.19}$$

where $\nabla_{\vartheta_\omega}(\boldsymbol{\xi}_\circ)$ denotes $\frac{\partial}{\partial\boldsymbol{\vartheta}_\omega^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\theta})|_{\boldsymbol{\xi}_x=\boldsymbol{\xi}_\circ}$ and the idempotent and Hermitian matrix $\boldsymbol{\Pi}^{\perp}_{\mathcal{S}_{\mathcal{R}}(\mathbf{A})} = \boldsymbol{I} - \mathbf{A}(\mathbf{A}^{\mathsf{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{H}}$ signifies the orthogonal complementary projection onto the range space of any full-rank matrix $\mathbf{A}$.

3.3 Asymptotic Maximum Likelihood Estimator

Next we shall derive AML estimator. Recall the matrix derivatives with respect to a real-valued scalar x that [51]

$$
\begin{aligned}
\frac{\partial}{\partial \mathrm{x}}\ln|\mathbf{A}(\mathrm{x})| &= \lceil \mathbf{A}^{-1}(\mathrm{x})\dot{\mathbf{A}}(\mathrm{x})\rfloor && (3.20a)\\
\frac{\partial}{\partial \mathrm{x}}\lceil \mathbf{A}(\mathrm{x})\mathbf{B}(\mathrm{x})\rfloor &= \lceil \mathbf{B}(\mathrm{x})\dot{\mathbf{A}}(\mathrm{x})\rfloor + \lceil \mathbf{A}(\mathrm{x})\dot{\mathbf{B}}(\mathrm{x})\rfloor && (3.20b)\\
\frac{\partial}{\partial \mathrm{x}}\mathbf{A}^{-1}(\mathrm{x}) &= -\mathbf{A}^{-1}(\mathrm{x})\dot{\mathbf{A}}(\mathrm{x})\mathbf{A}^{-1}(\mathrm{x}). && (3.20c)
\end{aligned}
$$

Differentiating the ML function with respect to the unstructured nuisance parameter $\boldsymbol{\eta}$, it yields

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\eta}^{\mathsf{T}}}\ell_{\mathrm{ML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega,\boldsymbol{\eta}) &= -\hat{\boldsymbol{\xi}}_x^{\mathsf{H}}\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\omega)\frac{\partial}{\partial\boldsymbol{\eta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\omega,\boldsymbol{\eta}) + \boldsymbol{\xi}_x^{\mathsf{H}}(\boldsymbol{\theta}_\omega)\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\omega)\frac{\partial}{\partial\boldsymbol{\eta}^{\mathsf{T}}}\boldsymbol{\xi}_x(\boldsymbol{\vartheta}_\omega,\boldsymbol{\eta})\\ &= (\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\boldsymbol{\eta}-\hat{\boldsymbol{\xi}}_x)^{\mathsf{H}}\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\omega)\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\end{aligned} \tag{3.21}$$

where the vector $\hat{\boldsymbol{\xi}}_x \in \mathbb{C}^{N_E^2\times 1}$ is derived from $\hat{\boldsymbol{\xi}}_x = \boldsymbol{v}_{\mathsf{c}}(\hat{\boldsymbol{\Sigma}}_{xx})$. Forcing $\frac{\partial}{\partial\boldsymbol{\eta}}\ell_{\mathrm{ML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega,\boldsymbol{\eta}) \trianglerighteq \boldsymbol{0}$, we obtain a critical relationship

$$\boldsymbol{\eta} = \left(\boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\vartheta}_\omega)\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\omega)\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\right)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\vartheta}_\omega)\,\boldsymbol{\Psi}_{xx}^{-1}(\boldsymbol{\theta}_\omega)\hat{\boldsymbol{\xi}}_x. \tag{3.22}$$

It is worthwhile to note that in (3.49), the nuisance parameter $\boldsymbol{\eta}$ depends itself on model parameter $\boldsymbol{\theta}$. Therefore, the existence of the above solution are quite hard to compute. Since $\hat{\boldsymbol{\Sigma}}_{xx} = \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ) + O_p(1/\sqrt{N_T})$ [1], or equivalently [52, p. 194]

$$\operatorname*{plim}_{N_T\to\infty}\frac{1}{N_T}\sum_{n_T=1}^{N_T}\boldsymbol{x}[n_T]\boldsymbol{x}^{\mathsf{H}}[n_T] = \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}_\circ) \tag{3.23}$$

we may replace $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\vartheta}_\omega,\boldsymbol{\eta})$ with the nonparametric estimate $\hat{\boldsymbol{\Sigma}}_{xx}$ without loss of asymptotic efficiency. If we designate the nonparametric estimate $\hat{\boldsymbol{\Psi}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E^2\times N_E^2}$ as $\hat{\boldsymbol{\Psi}}_{xx} \triangleq \hat{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \hat{\boldsymbol{\Sigma}}_{xx}$, it follows from (3.49) that the AML nuisance estimate becomes

$$\hat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega) = \left(\boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\vartheta}_\omega)\,\hat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\right)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}(\boldsymbol{\vartheta}_\omega)\,\hat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x. \tag{3.24}$$

Plugging the incomplete $\hat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega)$ into (3.3), we obtian

$$\boldsymbol{v}_{\mathsf{c}}(\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\vartheta}_\omega)) = \boldsymbol{\Omega}(\boldsymbol{\vartheta}_\omega)\hat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega) \tag{3.25}$$

where $\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\vartheta}_\omega) \triangleq \boldsymbol{\Sigma}_{xx}(\boldsymbol{\vartheta}_\omega,\hat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega)) : \mathbb{R}^{(2N_S)\times 1} \mapsto \mathbb{C}_{\mathbb{H}}^{N_E\times N_E}$ is the concentrated covariance for AML estimate.

Proposition 1. *If we define* $\ell_{\mathrm{AML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega) \triangleq \ell_{\mathrm{ML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega,\hat{\boldsymbol{\eta}}_{\mathrm{AML}}(\boldsymbol{\vartheta}_\omega))$, *the exact ML estimator in* (2.20) *can be approximated as*

$$\hat{\boldsymbol{\vartheta}}_{\mathrm{AML}} = \arg\min_{\boldsymbol{\vartheta}_\omega}\ell_{\mathrm{AML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega) \tag{3.26a}$$

$$\ell_{\mathrm{AML}}^{[N_T]}(\boldsymbol{\vartheta}_\omega) = \lceil\widehat{\boldsymbol{\Sigma}}_{xx}^{-1}(\boldsymbol{\vartheta}_\omega)\hat{\boldsymbol{\Sigma}}_{xx}\rfloor + \ln|\widehat{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\vartheta}_\omega)|. \tag{3.26b}$$

1

Definition 3. $\mathrm{x}_n = O_p(\mathrm{b}_n)$ *if* $\frac{\mathrm{x}_n}{\mathrm{b}_n} = O_p(1)$.

Definition 4. $\mathrm{x}_n = O_p(1)$ *if for every* $\epsilon > 0$, $\exists$ *a number* $k(\epsilon)$ *and an integer* $n(\epsilon)$ *such that if* $n > n(\epsilon)$, *then* $Pr(|\mathrm{x}_n| \le k(\epsilon)) \ge 1-\epsilon$.

Searching the minimum solution for $\hat{\boldsymbol{\vartheta}}_{\text{AML}}$ *as* $2N_s$ *dimensions in* (3.26), *we immediately obtain physical angle estimates via* (2.8).

Substituting it into $\hat{\boldsymbol{\eta}}_{\text{AML}}(\boldsymbol{\vartheta}_\omega)$, the AML estimate of model parameter $\boldsymbol{\theta}$ is thus

$$\hat{\boldsymbol{\theta}}_{\text{AML}} \triangleq \left[\hat{\boldsymbol{\vartheta}}_{\text{AML}} \quad \hat{\boldsymbol{\eta}}^{\mathsf{T}}_{\text{AML}}(\hat{\boldsymbol{\vartheta}}_{\text{AML}})\right]^{\mathsf{T}}. \tag{3.27}$$

As described above, it is noteworthy that the estimation in this way invokes the reparamterization, the concentration of likelihood function, and the asymptotic approximation. Therefore, it automatically originates an alternative way for lowering computational complexity withe respect to the exact ML method. Next we shall provide an asymptotic performance assessment of the proposed AML estimator.

3.4 Asymptotic Distribution of AML Estimate

In sensor array signal processing, the asymptotic behavior of ML estimates has been reported in [53] that the stochastic ML estimate based on Gaussian signal assumption is asymptotically attainable to the stochastic CRB even in the situation missing a priori distribution assumed for signal.

In this section, we shall first carry out what appropriate sense that governs the convergence of the estimated $\hat{\boldsymbol{\vartheta}}_{\text{AML}}$ sequence. Later, the implied convergence in distribution will be inspected by deriving the asymptotic efficiency relative to CRB. Without loss of generality, the parameter $\boldsymbol{\vartheta}$ might be no longer specified as physical or spatial frequency model.

3.4.1 Consistency

Let the limitation of AML criterion $\bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta})$ be a deterministic function designated by

$$\bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta}) \triangleq \lim_{N_T\to\infty} \mathcal{E}\left\langle \ell^{[N_T]}_{\text{AML}}(\boldsymbol{\vartheta})\right\rangle. \tag{3.28}$$

As a consequence, suppose that there exists a root of such a constant function [54, p. 281] in such a way that

$$\boldsymbol{\vartheta}_\circ = \arg\min_{\boldsymbol{\vartheta}} \bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta}). \tag{3.29}$$

In domain $\mathcal{D}_{\mathcal{M}}$ of a uniformly stable model [54, p. 252], the AML criterion uniformly converges to the limiting function $\bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta})$ with probability one, *i.e.*, [54, p. 254] and [21, p. 90]

$$\lim_{N_T\to\infty} \Pr\left(\sup_{\boldsymbol{\vartheta}\in\mathcal{D}_{\mathcal{M}}} \left|\ell^{[N_T]}_{\text{AML}}(\boldsymbol{\vartheta}) - \bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta})\right| > \varepsilon\right) = 0 \quad ;\forall\varepsilon > 0. \tag{3.30}$$

This follows, under the analogous argument, that for any model without ambiguities in array manifold the AML estimate $\hat{\boldsymbol{\vartheta}}_{\text{AML}}$ almost surely converges to the true quantity $\boldsymbol{\vartheta}_{\circ}$ as[2]

$$\lim_{N_T\to\infty} \Pr\left(\sup_{\boldsymbol{\vartheta}\in\mathcal{D}_{\mathcal{M}}} |\hat{\boldsymbol{\vartheta}}_{\text{AML}} - \boldsymbol{\vartheta}_{\circ}| > \varepsilon\right) = 0 \quad ;\forall\varepsilon > \boldsymbol{0}. \tag{3.31}$$

One may translate (3.30) and (3.31) into [21, pp. 89–90]

$$\Pr\left(\lim_{N_T\to\infty} \ell_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) = \bar{\ell}_{\text{AML}}(\boldsymbol{\vartheta})\right) = 1 \tag{3.32a}$$

$$\Pr\left(\lim_{N_T\to\infty} \hat{\boldsymbol{\vartheta}}_{\text{AML}} = \boldsymbol{\vartheta}_{\circ}\right) = 1. \tag{3.32b}$$

The above convergence enables the AML estimator to be a *consistent estimator* [55, p. 384] and in the sense of *strong consistency* [56, p. 48].

3.4.2 Asymptotic Efficiency

Let the gradient $\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) : \mathbb{R}^{(2N_S)\times 1} \mapsto \mathbb{R}^{(2N_S)\times 1}$ and the Hessian matrix $\boldsymbol{H}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) : \mathbb{R}^{(2N_S)\times 1} \mapsto \mathbb{R}_{\mathbb{S}}^{(2N_S)\times(2N_S)}$ be

$$\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) \triangleq \frac{\partial}{\partial\boldsymbol{\vartheta}} \ell_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) \tag{3.33a}$$

$$\boldsymbol{H}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}) \triangleq \frac{\partial^2}{\partial\boldsymbol{\vartheta}\otimes\partial\boldsymbol{\vartheta}^{\mathsf{T}}} \ell_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}). \tag{3.33b}$$

It is obvious that a necessary condition of the extremal point existence in the AML function must be exactly satisfied by $\boldsymbol{g}_{\text{AML}}^{[N_T]}(\hat{\boldsymbol{\vartheta}}_{\text{AML}}) = \boldsymbol{0}$. Assume that there is a parameter vector $\boldsymbol{\vartheta}_{\epsilon}$, aligned in a neighborhood region between the true value $\boldsymbol{\vartheta}_{\circ}$ and the AML estimate $\hat{\boldsymbol{\vartheta}}_{\text{AML}}$. This might be assumed in such a way that the Taylor series can be expanded around $\hat{\boldsymbol{\vartheta}}_{\text{AML}}$ as

$$\boldsymbol{0} = \boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}_{\circ}) + \boldsymbol{H}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}_{\epsilon})(\hat{\boldsymbol{\vartheta}}_{\text{AML}} - \boldsymbol{\vartheta}_{\circ}). \tag{3.34}$$

Under the same reason in (3.30), the AML Hessian *strongly converges* to its asymptotic amount according to

$$\Pr\left(\lim_{N_T\to\infty} \boldsymbol{H}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}_{\epsilon}) = \bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta}_{\circ})\right) = 1 \tag{3.35}$$

where the limiting Hessian matrix $\bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta}) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{R}_{\mathbb{S}}^{(2N_S)\times(2N_S)}$ can be defined as

$$\bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta}) \triangleq \lim_{N_T\to\infty} \mathcal{E}\left\langle \boldsymbol{H}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta})\right\rangle. \tag{3.36}$$

[2]The notation $|\cdot|$ in this expression denotes the absolute-valued operator.

For large enough N_T, the AML residual $\tilde{\boldsymbol{\vartheta}}_{\text{AML}} \triangleq \hat{\boldsymbol{\vartheta}}_{\text{AML}} - \boldsymbol{\vartheta}_\circ$ might be expressed as[3] (see *e.g.*, [57, p. 11])

$$\tilde{\boldsymbol{\vartheta}}_{\text{AML}} = -\bar{\boldsymbol{H}}_{\text{AML}}^{-1}(\boldsymbol{\theta}_\circ)\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}_\circ) + o_p\left(\frac{1}{\sqrt{N_T}}\right) \tag{3.37}$$

whence the limiting Hessian matrix is assumed to be invertible. Based on the Lyapunov central limit theorem [53], the gradient vector with true parameter will *converge in distribution* to normal distribution, *i.e.*, asymptotically distribute as

$$\sqrt{N_T}\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta}_\circ) \overset{\infty}{\approx} \mathcal{N}_{\mathcal{R}}(\boldsymbol{0}; \overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta}_\circ)) \tag{3.38}$$

where the asymptotic covariance matrix due to gradient vector $\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta})$, $\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta}) : \mathbb{R}^{(3N_S+1)\times 1} \mapsto \mathbb{R}_{\mathbb{S}}^{(2N_S)\times(2N_S)}$, is defined as

$$\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta}) \triangleq \lim_{N_T\to\infty} N_T \mathcal{E}\left\langle \boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta})\left(\boldsymbol{g}_{\text{AML}}^{[N_T]}(\boldsymbol{\vartheta})\right)^{\mathsf{T}}\right\rangle. \tag{3.39}$$

The transformation in (3.37) implies that the error of AML estimate also converges in distribution to normal distribution with

$$\sqrt{N_T}(\hat{\boldsymbol{\vartheta}}_{\text{AML}} - \boldsymbol{\vartheta}_\circ) \overset{\infty}{\approx} \mathcal{N}_{\mathcal{R}}(\boldsymbol{0}; \overline{\boldsymbol{B}}_{\text{AML}}(\boldsymbol{\theta}_\circ)) \tag{3.40}$$

where the limiting AML residual covariance is given by

$$\overline{\boldsymbol{B}}_{\text{AML}}(\boldsymbol{\theta}_\circ) = \bar{\boldsymbol{H}}_{\text{AML}}^{-1}(\boldsymbol{\theta}_\circ)\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta}_\circ)\bar{\boldsymbol{H}}_{\text{AML}}^{-1}(\boldsymbol{\theta}_\circ). \tag{3.41}$$

In appendix 5.2, it was derived that

$$\overline{\boldsymbol{Q}}_{\text{AML}}(\boldsymbol{\theta}_\circ) = \bar{\boldsymbol{H}}_{\text{AML}}(\boldsymbol{\theta}_\circ) = \frac{1}{N_T}\boldsymbol{B}_{\text{CR}(\hat{\boldsymbol{\vartheta}})}^{-1}(\boldsymbol{\theta}_\circ). \tag{3.42}$$

Together with (3.41), now we can find the AML error covariance theoretically from $\bar{\boldsymbol{B}}_{\text{AML}}(\boldsymbol{\theta}_\circ) \triangleq \frac{1}{N_T}\overline{\boldsymbol{B}}_{\text{AML}}(\boldsymbol{\theta}_\circ)$, *i.e.*,

$$\bar{\boldsymbol{B}}_{\text{AML}}(\boldsymbol{\theta}_\circ) = \boldsymbol{B}_{\text{CR}(\hat{\boldsymbol{\vartheta}})}(\boldsymbol{\theta}_\circ). \tag{3.43}$$

Indeed the AML standard deviation coincides in large sample with the square root of Cramér-Rao bound (3.18).

3

Definition 5. $\mathrm{x}_n = o_p(\mathrm{b}_n)$ *if* $\frac{\mathrm{x}_n}{\mathrm{b}_n} = o_p(1)$.

Definition 6. $\mathrm{x}_n = o_p(1)$ *if* x_n converges in probability *to zero* ($\mathrm{x}_n \xrightarrow{Pr} 0$), *or mathematically speaking in the sense of* $\underset{n\to\infty}{plim}$, $\lim_{n\to\infty} Pr(|\mathrm{x}_n| > \varepsilon) = 0\,; \forall \varepsilon > 0$.

3.5 Toeplitzification Relationships

Before proceeding on the methodologies of Toeplitzification, it is noteworthy to see how the array covariance matrix $\boldsymbol{\Sigma}_{xx}$ can be Toeplitz.

3.5.1 How the array covariance matrix can be Toeplitz

For more clarification, we develop a theorem on such.

Theorem 2 (How can the array covariance matrix $\boldsymbol{\Sigma}_{xx}$ be Toeplitz ?). *If we employ the ULA indicated in* (2.1), *then it follows that the array covariance matrix $\boldsymbol{\Sigma}_{xx}$ is also Toeplitz, i.e.,*

$$[\boldsymbol{\Sigma}_{xx}]_{[n_L+1,1]} = [\boldsymbol{\Sigma}_{xx}]_{[n_L+n_l,n_l]} \; ; \forall n_l \in \{1, 2, \ldots, N_E - n_L\} \tag{3.44}$$

where $n_L \triangleq n_E - \acute{n}_E \in \{0, 1, N_E - 1\}$ is the index of sub-diagonal or lag.

Proof. See appendix. □

Remark 1. *The reason why Toeplitz structure is held itself belongs to the assumption that sensor elements are aligned as the uniform linear array with phase reference at the first element.*

Remark 2. *It is worthwhile to point out that theorem 2 does not make any assumption on the symmetry of the angle deviation which is, in general, modelled as random distribution symmetric around nominal direction. Therefore, another virtute derived from the latter theorem is that the parameter estimation with this additional information is as general as that recently reported in [58].*

Theorem 2 enables us to a priori knowledged of array covariance structure. Towards this end, all elements in array covariance should be forced to make the estimated covariance matrix Hermitian and Toeplitz. Due to approaching the exact value more rapid than the ordinary sample covariance, the estimated covariance matrix with Toeplitz constraint seems advantageous when having to replace the parametric array covariance matrix with a non-parametric estimate.

3.5.2 Relationship between RA and WCM Toeplitzifications

It is however fruitful to see the relation between the WCM and RA estimate. Furthermore, if it needs to deploy one of both estimates, *i.e.*, the WCM and RA fundamental vector estimates, their statistical properties should be concerned.

Lemma 3 (Reformable weight from WCM to RA). *Let the optimal weight* $\boldsymbol{W}$ *in* (2.52) *be the identity matrix. Then, a consistent fundamental vector estimate* $\hat{\boldsymbol{\tau}}_{\mathsf{CM}} \in \mathbb{R}^{(2N_E-1)\times 1}$ *which is given by*

$$\begin{aligned}\hat{\boldsymbol{\tau}}_{\mathsf{CM}} &\triangleq \arg\min_{\boldsymbol{\tau}} \|\bar{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx}\|_I^2 \\ &= (\boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \breve{\boldsymbol{\Xi}}\, \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\xi}}_x \\ &= \boldsymbol{\Upsilon}^{-1} (\breve{\boldsymbol{\Xi}}^{\mathsf{T}} \breve{\boldsymbol{\Xi}})^{-1} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\xi}}_x\end{aligned} \tag{3.45}$$

is equivalent to that available from the RA method i.e.,

$$\hat{\boldsymbol{\tau}}_{\mathsf{CM}} = \hat{\boldsymbol{\tau}}_{\mathsf{RA}}. \tag{3.46}$$

Proof. See appendix. □

Notice that (2.51) might be rewritten as

$$\begin{aligned}\boldsymbol{v}_{\mathsf{c}}(\hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathsf{RA}})) &= \boldsymbol{\Pi}_{\mathcal{S}_{\mathcal{R}}(\breve{\boldsymbol{\Xi}}\boldsymbol{\Upsilon})} \hat{\boldsymbol{\xi}}_x \\ &= \boldsymbol{\Pi}_{\mathcal{S}_{\mathcal{R}}(\breve{\boldsymbol{\Xi}})} \hat{\boldsymbol{\xi}}_x\end{aligned} \tag{3.47}$$

where $\boldsymbol{\Pi}_{\mathcal{S}_{\mathcal{R}}(\mathbf{A})} = \mathbf{A}(\mathbf{A}^{\mathsf{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{H}}$ signifies the orthogonal projection onto the range space of any full rank matrix $\mathbf{A}$. Such a lemma is aims at illustrating the similarity between unweighed WCM approach and RA method.

Remark 3. *Based on best linear unbiased estimator (BLUE) viewpoint [36], the performance of WCM must be, in principle, superior than that available from the RA method because the first one invoked the optimal weight to reduce the residual.*

3.5.3 AML Toeplitzification

In this section, we derive an array covariance Toeplitizification and then consider its relationship to other loss function. Note that $\frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} \bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) = \breve{\boldsymbol{\Xi}}\, \boldsymbol{\Upsilon}$. Differentiating (2.20b) with respect to $\boldsymbol{\tau}$, it yields

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} \ell_{\mathrm{ML}}^{[N_T]}(\boldsymbol{\tau}) &= -\hat{\boldsymbol{\xi}}_x^{\mathsf{H}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} \bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) + \bar{\boldsymbol{\xi}}_x^{\mathsf{H}}(\boldsymbol{\tau}) \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \frac{\partial}{\partial \boldsymbol{\tau}^{\mathsf{T}}} \bar{\boldsymbol{\xi}}_x(\boldsymbol{\tau}) \\ &= (\breve{\boldsymbol{\Xi}}\, \boldsymbol{\Upsilon} \boldsymbol{\tau} - \hat{\boldsymbol{\xi}}_x)^{\mathsf{H}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \breve{\boldsymbol{\Xi}}\, \boldsymbol{\Upsilon}\end{aligned} \tag{3.48}$$

where $\bar{\boldsymbol{\Psi}}_{xx} \triangleq \bar{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \bar{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}^{N_E^2 \times N_E^2}$, and $\hat{\boldsymbol{\xi}}_x = \boldsymbol{v}_{\mathsf{c}}(\hat{\boldsymbol{\Sigma}}_{xx})$. Forcing $\frac{\partial}{\partial \boldsymbol{\tau}} \ell_{\mathrm{ML}}^{[N_T]}(\boldsymbol{\tau}) \trianglerighteq \boldsymbol{0}$, we obtain

$$\begin{aligned}\boldsymbol{\tau} &= \left(\boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \breve{\boldsymbol{\Xi}}\, \boldsymbol{\Upsilon}\right)^{-1} \boldsymbol{\Upsilon}^{\mathsf{H}} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \hat{\boldsymbol{\xi}}_x \\ &= \boldsymbol{\Upsilon}^{-1} \left(\breve{\boldsymbol{\Xi}}^{\mathsf{T}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \breve{\boldsymbol{\Xi}}\right)^{-1} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \bar{\boldsymbol{\Psi}}_{xx}^{-1}(\boldsymbol{\tau}) \hat{\boldsymbol{\xi}}_x.\end{aligned} \tag{3.49}$$

In what follows, we shall designate the Toeplitz-constrained covariance estimate as $\hat{\bar{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathsf{ML}})$.

Proposition 2. *Since* $\hat{\boldsymbol{\Psi}}_{xx} \triangleq \hat{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \hat{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}^{N_E^2 \times N_E^2}$ *converges in probability to* $\bar{\boldsymbol{\Psi}}_{xx}(\boldsymbol{\tau})$, *we can replace the parametric* $\bar{\boldsymbol{\Psi}}_{xx}(\boldsymbol{\tau})$ *with the non-parametric* $\hat{\boldsymbol{\Psi}}_{xx}$, *without loss of asymptotic performance. Based on* (3.49), *the Toeplitz-constrained covariance estimate can be approximated as*

$$\boldsymbol{v}_{\mathsf{c}}(\hat{\hat{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathrm{AML}})) = \breve{\boldsymbol{\Xi}}\big(\breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \breve{\boldsymbol{\Xi}}\big)^{-1} \breve{\boldsymbol{\Xi}}^{\mathsf{T}} \hat{\boldsymbol{\Psi}}_{xx}^{-1} \hat{\boldsymbol{\xi}}_x \tag{3.50}$$

where AML stands for the attribution of asymptotic maximum likelihood criterion.

3.5.4 Relationship between WCM and AML Toeplitzifications

We then recall the *w*eighted *c*ovariance-*m*atching (WCM) criterion [35]. Let a *w*eighted *l*east *s*quares (WLS) criterion be [48]

$$f_{\mathrm{WLS}}(\boldsymbol{\tau}|\boldsymbol{W}) \triangleq \|\bar{\boldsymbol{\Sigma}}_{xx}(\boldsymbol{\tau}) - \hat{\boldsymbol{\Sigma}}_{xx}\|_{\boldsymbol{W}}^2 \tag{3.51}$$

where $\|\mathbf{A}\|_{\mathbf{W}}^2 \triangleq \boldsymbol{v}_{\mathsf{c}}^{\mathsf{H}}(\mathbf{A})\,\mathbf{W}^{-1}\boldsymbol{v}_{\mathsf{c}}(\mathbf{A})$ designates a weighted version of the Euclidean norm. By replacing the consistent weight matrix $\hat{\boldsymbol{W}}$ as $\hat{\boldsymbol{\Psi}}_{xx}$, the WCM estimate of Toepltiz lag is given from

$$\hat{\boldsymbol{\tau}}_{\mathrm{WCM}} = \arg\min_{\boldsymbol{\tau}} f_{\mathrm{WLS}}(\boldsymbol{\tau}|\hat{\boldsymbol{W}}). \tag{3.52}$$

Lemma 4 (Equivalence of WLS and AML). *Based on the weighted covariance-matching criterion* (3.51), *the WCM estimate yields the same solution as provided by the AML estimate, i.e.,*

$$\hat{\boldsymbol{\tau}}_{\mathrm{WCM}} = \hat{\boldsymbol{\tau}}_{\mathrm{AML}}. \tag{3.53}$$

Proof. See appendix. □

Remark 4. *The lemma 4 allows us to omit the numerical verification since the WCM and AML are available from the same expression (see* (2.54) *and* (3.50)*).*

3.6 Application of Toeplitz-Constrained Covariances

Since the array covariance matrix is itself Toeplitz, one has to be aware of such a matrix structure during parameter estimation.

3.6.1 Improved WLS Estimators

For terminology concise, one may recall $\hat{\boldsymbol{\vartheta}}_{\mathrm{WLS}}$ as *ordinary WLS* estimate when replacing $\boldsymbol{W}$ in (3.51) with $\hat{\boldsymbol{W}}$ in (2.53) and (2.21).

Proposition 3. *By substituting* $\boldsymbol{W}$ *in* (3.51) *with* $\hat{\hat{\boldsymbol{W}}}(\hat{\boldsymbol{\tau}}_{\mathrm{RA}}) \triangleq \hat{\hat{\boldsymbol{\Sigma}}}_{xx}^{\mathsf{T}}(\hat{\boldsymbol{\tau}}_{\mathrm{RA}}) \otimes \hat{\hat{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathrm{RA}})$ *and* $\hat{\hat{\boldsymbol{W}}}(\hat{\boldsymbol{\tau}}_{\mathrm{WCM}}) \triangleq \hat{\hat{\boldsymbol{\Sigma}}}_{xx}^{\mathsf{T}}(\hat{\boldsymbol{\tau}}_{\mathrm{WCM}}) \otimes \hat{\hat{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathrm{WCM}})$, *we shall designate the improved solution as "RA-WLS" and "WCM-WLS" estimates, respectively.*

3.6.2 Improved AML Estimator

To estimate nominal direction in the presence of spatially distributed source, the comparative study of asymptotically efficient estimators in what follows of the thesis is shown that the large-sample approximated maximum likelihood estimator outperforms the WLS-based estimator [35] in non-asymptotic region. Here we try to incorporate the Toeplitz-constrained covariance estimated from (3.50) into the asymptotic maximum likelihood estimator shown earlier.

Let $\check{\boldsymbol{\Sigma}}_{xx}(\omega) : [-kd_E, kd_E] \mapsto \mathbb{C}_{\mathbb{H}}^{N_E \times N_E}$ be a concentrated covariance matrix which is calculated from

$$\boldsymbol{v}_{\mathsf{c}}(\check{\boldsymbol{\Sigma}}_{xx}(\omega)) = \boldsymbol{\Omega}(\omega)\big(\boldsymbol{\Omega}^{\mathsf{H}}(\omega)\,\widehat{\boldsymbol{\Psi}}_{xx}^{-1}\boldsymbol{\Omega}(\omega)\big)^{-1}\boldsymbol{\Omega}^{\mathsf{H}}(\omega)\,\widehat{\boldsymbol{\Psi}}_{xx}^{-1}\hat{\boldsymbol{\xi}}_x \tag{3.54}$$

where $\widehat{\boldsymbol{\Psi}}_{xx} \triangleq \widehat{\boldsymbol{\Sigma}}_{xx}^{\mathsf{T}} \otimes \widehat{\boldsymbol{\Sigma}}_{xx} \in \mathbb{C}_{\mathbb{H}}^{N_E^2 \times N_E^2}$ is arbitrary non-parametric estimate of $\boldsymbol{\Psi}_{xx}(\boldsymbol{\theta})$. Then, the nominal direction estimate in this way can be given by

$$\hat{\omega}_{\mathrm{AML}} = \arg\min_{\omega} \ell_{\mathrm{AML}}^{[N_T]}(\omega) \tag{3.55a}$$

$$\ell_{\mathrm{AML}}^{[N_T]}(\omega) = \lceil \check{\boldsymbol{\Sigma}}_{xx}^{-1}(\omega)\hat{\boldsymbol{\Sigma}}_{xx} \rfloor + \ln|\check{\boldsymbol{\Sigma}}_{xx}(\omega)|. \tag{3.55b}$$

Plugging $\widehat{\boldsymbol{\Sigma}}_{xx}$ in (3.54) as $\hat{\boldsymbol{\Sigma}}_{xx}$ derived from (2.21), the above $\hat{\omega}_{\mathrm{AML}}$ owing to (3.55) will be called the "*ordinary AML*" estimate.

Proposition 4. *By invoking $\widehat{\boldsymbol{\Sigma}}_{xx} = \hat{\hat{\boldsymbol{\Sigma}}}_{xx}(\hat{\boldsymbol{\tau}}_{\mathrm{AML}})$ according to (3.50) and (3.54), we shall designate the solution $\hat{\omega}_{\mathrm{AML}}$ in (3.55) as the estimate of "AML with Toeplitz constraint".*