Publication IX

Esa Ollila, Visa Koivunen, and Jan Eriksson. 2008. On the Cramér-Rao bound for the constrained and unconstrained complex parameters. In: Proceedings of the 5th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2008). Darmstadt, Germany. 21-23 July 2008, pages 414-418.

ON THE CRAMÉR-RAO BOUND FOR THE CONSTRAINED AND UNCONSTRAINED COMPLEX PARAMETERS

Esa Ollila *

University of Oulu
Department of Mathematical Sciences
P.O. Box 3000, FIN-90014 Oulu, Finland

Visa Koivunen and Jan Eriksson

Helsinki University of Technology
Signal Processing Laboratory
P.O. Box 3000, FIN-02015 HUT, Finland

ABSTRACT

We derive a complex form of the unconstrained and constrained Cramér-Rao lower bound (CRB) of composite real parameters formed by stacking the real and imaginary part of the complex parameters. The derived complex constrained and unconstrained CRB is easy to calculate and possesses similar structure as in the real parameter case but with the real covariance, Jacobian and the Fisher information matrix replaced by complex matrices with analogous interpretations. The advantage of the complex CRB is that it is oftentimes easier to calculate than its real form. It is highlighted that a statistic that attains a bound on the complex covariance matrix alone do not necessarily attain the CRB since complex covariance matrix does not provide a full second-order description of a complex statistic since also the pseudo-covariance matrix is needed. Our derivations also lead to some new insights and theory that are similar to real CRB theory.

1. INTRODUCTION

It is highly useful to have a lower bound for the statistical variability of a statistic. The most well-known lower bound is the Cramér-Rao bound (CRB) which can be used, for example, to show that an unbiased estimator is UMVU (uniformly minimum variance unbiased) estimator. CRB is also related to asymptotic optimality theory and called the information bound, or, the information inequality in single parameter case. See e.g. [1, 2, 3, 4].

The problem of estimating multiple possibly constrained complex parameters from complex-valued data arises frequently in several signal processing applications. Derivations of complex CRB appear in the pioneering works of [5, 1, 6] and some recent derivations appear in [7, 8, 9]. One advantage of our approach is that it shows more transparently the equivalence between the "complex CRB" and the real CRB; see our Theorem 1. Furthermore, our approach leads to easier derivations of the CRB. It is highlighted that a statistic that attains a bound on the complex covariance matrix alone (as in [6, 1]) does not necessarily attain the CRB since complex covariance matrix does not provide a full second-order description of a complex statistic, also pseudo-covariance matrix is needed; c.f. Remark 1 to Corollary 2. It is shown that if and only if the so called pseudo-information matrix vanishes, then the derived bound on the complex covariance matrix of an ubiased estimator is fully equivalent with the real CRB; c.f. Corollary 1[b]. Our derivations also lead to new insights and theory that are analogous to real CRB theory; c.f. Theorem 2.

*The first author performed main part of his research while working as a post-doctoral fellow of Academy of Finland at Signal Processing Laboratory, Helsinki University of Technology.

2. USEFUL MATRIX ALGEBRA

A $k \times k$ complex matrix $\mathbf{C} \in \mathbb{C}^{k\times k}$ is called *symmetric* if $\mathbf{C}^T = \mathbf{C}$ and *Hermitian* if $\mathbf{C}^H = \mathbf{C}$, where $(\cdot)^T$ and $(\cdot)^H$ denotes transpose and Hermitian (complex conjugate) transpose, *i.e.* $(\cdot)^H = [(\cdot)^*]^T$, where $(\cdot)^*$ denotes complex conjugate. Notation $\mathbf{C} \geq \mathbf{B}$ means that the matrix $\mathbf{C} - \mathbf{B}$ is positive semidefinite. Notation $\mathbf{C}^{-*}$ means $(\mathbf{C}^{-1})^*$, $\mathrm{tr}(\mathbf{C})$ denotes the trace of the matrix $\mathbf{C}$ and $|z|$ denotes the modulus of the complex number z. Adopting the notation from [10], we define a complex $2k \times 2k$ matrix $\mathbf{M}_{2k}$ as

$$\mathbf{M}_{2k} \triangleq \frac{1}{2}\begin{pmatrix} \mathbf{I}_k & \mathbf{I}_k \\ -\jmath\mathbf{I}_k & \jmath\mathbf{I}_k \end{pmatrix},$$

where $\mathbf{I}_k$ denotes the $k \times k$ *identity matrix* and $\jmath = \sqrt{-1}$. $\mathbf{M}_{2k}$ is invertible with the inverse $\mathbf{M}_{2k}^{-1} = 2\mathbf{M}_{2k}^H$.

Definition 1 *Define* $\langle\cdot\rangle_{\mathbb{C}} : \mathbb{R}^{2d\times 2k} \mapsto \mathbb{C}^{2d\times 2k}$ *as a mapping*

$$\langle\mathbf{G}\rangle_{\mathbb{C}} = 2\mathbf{M}_{2d}^{-1}\mathbf{G}\mathbf{M}_{2k}$$

that is,

$$\left\langle\begin{pmatrix} \mathrm{Re}[\mathbf{A}] & \mathrm{Re}[\mathbf{B}] \\ \mathrm{Im}[\mathbf{A}] & \mathrm{Im}[\mathbf{B}] \end{pmatrix}\right\rangle_{\mathbb{C}} = \begin{pmatrix} \mathbf{A} - \jmath\mathbf{B} & \mathbf{A} + \jmath\mathbf{B} \\ (\mathbf{A} + \jmath\mathbf{B})^* & (\mathbf{A} - \jmath\mathbf{B})^* \end{pmatrix}$$

for all $\mathbf{A} \in \mathbb{C}^{d\times k}$ *and* $\mathbf{B} \in \mathbb{C}^{d\times k}$.

Mapping $\langle \cdot \rangle_{\mathbb{C}}$ of $\mathbf{G} \in \mathbb{R}^{2d \times 2k}$ produces a complex $2d \times 2k$ matrix of the form

$$\langle \mathbf{G} \rangle_{\mathbb{C}} = \begin{pmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{D}^* & \mathbf{C}^* \end{pmatrix} \quad (1)$$

where $\mathbf{C}$ and $\mathbf{D}$ are complex $d \times k$ matrices. Hence we shall call matrix $\langle \mathbf{G} \rangle_{\mathbb{C}}$ as the *augmented matrix of* $\mathbf{C}$ *and* $\mathbf{D}$. Note that $\mathbf{G} = \mathbf{M}_{2d} \langle \mathbf{G} \rangle_{\mathbb{C}} \mathbf{M}_{2k}^H$

Lemma 1 *The mapping $\langle \cdot \rangle_{\mathbb{C}}$ has the following properties:*

[a] $\mathbf{G} \in \mathbb{R}^{2k \times 2k}$ *is symmetric* $\Leftrightarrow \langle \mathbf{G} \rangle_{\mathbb{C}}$ *is hermitian.*

[b] $\mathbf{G} \in \mathbb{R}^{2k \times 2k}$ *is invertible* $\Leftrightarrow \langle \mathbf{G} \rangle_{\mathbb{C}}$ *is invertible. Moreover,* $\langle \mathbf{G}^{-1} \rangle_{\mathbb{C}} = 4 \langle \mathbf{G} \rangle_{\mathbb{C}}^{-1}$

[c] $\langle \mathbf{G}^T \rangle_{\mathbb{C}} = \langle \mathbf{G} \rangle_{\mathbb{C}}^H$ *for all* $\mathbf{G} \in \mathbb{R}^{2d \times 2k}$.

[d] $\langle \mathbf{G}_1 + \mathbf{G}_2 \rangle_{\mathbb{C}} = \langle \mathbf{G}_1 \rangle_{\mathbb{C}} + \langle \mathbf{G}_2 \rangle_{\mathbb{C}}$ *for all* $\mathbf{G}_1, \mathbf{G}_2 \in \mathbb{R}^{2d \times 2k}$.

[e] $\langle \mathbf{G}_1 \mathbf{G}_2 \rangle_{\mathbb{C}} = \frac{1}{2} \langle \mathbf{G}_1 \rangle_{\mathbb{C}} \langle \mathbf{G}_2 \rangle_{\mathbb{C}}$ *for all* $\mathbf{G}_1 \in \mathbb{R}^{2d \times 2k}$ *and* $\mathbf{G}_2 \in \mathbb{R}^{2k \times 2q}$.

[f] $\mathbf{G} \geq 0 \Leftrightarrow \langle \mathbf{G} \rangle_{\mathbb{C}} \geq 0$ *for all symmetric* $\mathbf{G} \in \mathbb{R}^{2k \times 2k}$.

3. CRAMÉR-RAO LOWER BOUND

The distribution of a complex random vector (r.v) $\mathbf{z} = \mathbf{x} + \jmath \mathbf{y} \in \mathbb{C}^n$ is identified with the real $2n$-variate distribution of the composite real r.v. $\bar{\mathbf{z}} = (\mathbf{x}^T, \mathbf{y}^T)^T$ obtained by stacking the real part and imaginary part of $\mathbf{z}$. The p.d.f. of $\mathbf{z}$ is assumed to depend on the unknown k-variate complex parameter vector $\boldsymbol{\theta} = \boldsymbol{\alpha} + \jmath \boldsymbol{\beta} \in \mathbb{C}^k$. This is equivalent to saying that the p.d.f. of $\bar{\mathbf{z}}$ depends on the unknown $2k$-variate real parameter vector $\bar{\boldsymbol{\theta}} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T \in \mathbb{R}^{2k}$. Hence $f(\mathbf{z}; \boldsymbol{\theta})$ and $f(\bar{\mathbf{z}}; \bar{\boldsymbol{\theta}})$ are two alternative equivalent notations for the p.d.f. of $\mathbf{z}$.

Let

$$\mathbf{t} = \mathbf{t}_R + \jmath \mathbf{t}_I = \mathbf{t}(\mathbf{z}) \in \mathbb{C}^d, \quad d \leq k$$

denote the complex-valued statistic based on $\mathbf{z}$ and write $\bar{\mathbf{t}} = (\mathbf{t}_I^T, \mathbf{t}_R^T)^T$ for its associated composite real form. The expected value of $\mathbf{t}$, $E[\mathbf{t}] = E[\mathbf{t}_R] + \jmath E[\mathbf{t}_I]$, depends naturally on the value of the parameter $\boldsymbol{\theta}$. Hence

$$\mathbf{g}(\boldsymbol{\theta}) = \mathbf{u}(\boldsymbol{\alpha}, \boldsymbol{\beta}) + \jmath \mathbf{v}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = E[\mathbf{t}] : \mathbb{C}^k \to \mathbb{C}^d,$$

that is,

$$\bar{\mathbf{g}}(\bar{\boldsymbol{\theta}}) = \begin{pmatrix} \mathbf{u}(\bar{\boldsymbol{\theta}}) \\ \mathbf{v}(\bar{\boldsymbol{\theta}}) \end{pmatrix} = E[\bar{\mathbf{t}}] : \mathbb{R}^{2k} \to \mathbb{R}^{2d}$$

Jacobian matrix of $\bar{\mathbf{g}}$ is a $2d \times 2k$ real matrix

$$\mathbf{D}_{\bar{\mathbf{g}}} = \frac{\partial \bar{\mathbf{g}}}{\partial \bar{\boldsymbol{\theta}}} = \begin{pmatrix} \frac{\partial \mathbf{u}}{\partial \boldsymbol{\alpha}} & \frac{\partial \mathbf{u}}{\partial \boldsymbol{\beta}} \\ \frac{\partial \mathbf{v}}{\partial \boldsymbol{\alpha}} & \frac{\partial \mathbf{v}}{\partial \boldsymbol{\beta}} \end{pmatrix}$$

where $\partial / \partial \boldsymbol{\alpha} = (\partial / \alpha_1, \ldots, \partial / \alpha_k)$. Naturally, if the statistic $\mathbf{t}$ is an unbiased estimator of the parameter $\boldsymbol{\theta}$, then $\mathbf{g}(\boldsymbol{\theta}) = \boldsymbol{\theta}$ (*i.e.* $\mathbf{u}(\bar{\boldsymbol{\theta}}) = \boldsymbol{\alpha}$ and $\mathbf{v}(\bar{\boldsymbol{\theta}}) = \boldsymbol{\beta}$) and $\mathbf{D}_{\bar{\mathbf{g}}} = \mathbf{I}_{2k}$.

The *Fisher Information Matrix* (FIM) of $\bar{\boldsymbol{\theta}}$ is defined as

$$\mathbf{J}_{\bar{\boldsymbol{\theta}}} = E[\, \nabla_{\bar{\boldsymbol{\theta}}} \ln f(\mathbf{z}; \boldsymbol{\theta}) \{ \nabla_{\bar{\boldsymbol{\theta}}} \ln f(\mathbf{z}; \boldsymbol{\theta}) \}^T \,],$$

where

$$\nabla_{\bar{\boldsymbol{\theta}}} \ln f(\mathbf{z}; \boldsymbol{\theta}) = \begin{pmatrix} \nabla_{\boldsymbol{\alpha}} \ln f(\mathbf{z}; \boldsymbol{\theta}) \\ \nabla_{\boldsymbol{\beta}} \ln f(\mathbf{z}; \boldsymbol{\theta}) \end{pmatrix}$$

is the (real) gradient of $\ln f$ w.r.t $\bar{\boldsymbol{\theta}}$ and gradient operator is defined as $\nabla_{\boldsymbol{\alpha}} = (\partial / \partial \alpha_1, \ldots, \partial / \partial \alpha_k)^T$. In this paper, it is assumed that FIM $\mathbf{J}_{\bar{\boldsymbol{\theta}}}$ is positive definite so that its inverse exists.

Complete second-order information of complex-valued statistic $\mathbf{t}$ is given by the $2d \times 2d$ real covariance matrix of the composite vector $\bar{\mathbf{t}}$,

$$\begin{aligned} \mathrm{Cov}[\,\bar{\mathbf{t}}\,] &= E[\,(\bar{\mathbf{t}} - E[\,\bar{\mathbf{t}}\,])(\bar{\mathbf{t}} - E[\,\bar{\mathbf{t}}\,])^T\,] \\ &= \begin{pmatrix} \mathrm{Cov}[\,\mathbf{t}_R\,] & \mathrm{Cov}[\mathbf{t}_R, \mathbf{t}_I] \\ \mathrm{Cov}[\mathbf{t}_I, \mathbf{t}_R] & \mathrm{Cov}[\,\mathbf{t}_I\,] \end{pmatrix}. \end{aligned}$$

The *Cramer-Rao bound (CRB)* gives the lower bound on the covariance matrix of the statistic $\bar{\mathbf{t}}$ by stating that under some regularity conditions

$$\mathrm{Cov}[\,\bar{\mathbf{t}}\,] \geq \mathbf{D}_{\bar{\mathbf{g}}} \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \mathbf{D}_{\bar{\mathbf{g}}}^T. \quad (2)$$

The regularity conditions are required for the interchange of certain differentiation and integration operators (see [2] for details). Due to Lemma 1, the CRB (2) is equivalent with the statement

$$\langle \mathrm{Cov}[\,\bar{\mathbf{t}}\,] \rangle_{\mathbb{C}} \geq \langle \mathbf{D}_{\bar{\mathbf{g}}} \, \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \mathbf{D}_{\bar{\mathbf{g}}}^T \rangle_{\mathbb{C}} = \langle \mathbf{D}_{\bar{\mathbf{g}}} \rangle_{\mathbb{C}} \langle \mathbf{J}_{\bar{\boldsymbol{\theta}}} \rangle_{\mathbb{C}}^{-1} \langle \mathbf{D}_{\bar{\mathbf{g}}} \rangle_{\mathbb{C}}^H, \quad (3)$$

which is the *complex form* of the (real) CRB (2). Next we show that complex matrices on the left and right hand side have meaningful interpretations.

4. COMPLEX-VALUED FUNCTIONS AND RANDOM VECTORS: PRELIMINARIES

4.1. Partial derivatives of a complex function

Complex partial derivatives (c.p.d.'s) were first defined upto our best knowledge in the text book by L.V. Ahlfors [11], page 41. Many of the common rules (e.g. the product and the quotient rules) associated with the c.p.d.'s continue to hold in the familiar form known from the real calculus, the notable exception being the chain rule. This makes the c.p.d.'s a useful tool in the analysis of complex functions. See e.g. [12], [1], Chapter 15.6, [13] for a detailed treatment of these partial differential operators.

Define the partial derivative of a complex function $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{u}(\boldsymbol{\alpha}, \boldsymbol{\beta}) + \jmath \mathbf{v}(\boldsymbol{\alpha}, \boldsymbol{\beta}) : \mathbb{C}^k \to \mathbb{C}^d$ w.r.t $\boldsymbol{\alpha} = \mathrm{Re}[\boldsymbol{\theta}]$ as $\partial \mathbf{g} / \partial \boldsymbol{\alpha} =$

$\partial \mathbf{u}/\partial \boldsymbol{\alpha} + \jmath \partial \mathbf{v}/\partial \boldsymbol{\alpha}$ and define $\partial \mathbf{g}/\partial \boldsymbol{\beta}$ analogously. Then, the complex partial differential operators $\partial/\partial \boldsymbol{\theta}$ and $\partial/\partial \boldsymbol{\theta}^*$ are defined as

$$\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} = \frac{1}{2}\left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\alpha}} - \jmath \frac{\partial \mathbf{g}}{\partial \boldsymbol{\beta}}\right), \quad \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} = \frac{1}{2}\left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\alpha}} + \jmath \frac{\partial \mathbf{g}}{\partial \boldsymbol{\beta}}\right).$$

Definition 2 *Complex Jacobian matrix of a complex function* $\mathbf{g} : \mathbb{C}^k \to \mathbb{C}^d$ *is defined as the complex* $2d \times 2k$ *matrix*

$$\boldsymbol{\mathcal{D}}_{\mathbf{g}} = \begin{pmatrix} \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} & \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} \\ \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*}\right]^* & \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right]^* \end{pmatrix},$$

i.e. it is the augmented matrix of $\partial \mathbf{g}/\partial \boldsymbol{\theta}$ *and* $\partial \mathbf{g}/\partial \boldsymbol{\theta}^*$.

It is easy to verify that

$$\langle \mathbf{D}_{\bar{\mathbf{g}}} \rangle_{\mathbb{C}} = 2 \cdot \boldsymbol{\mathcal{D}}_{\mathbf{g}}. \tag{4}$$

4.2. Complex covariances and the complex FIM

Covariance matrix of a complex r.v. $\mathbf{t} = \mathbf{t}_R + \jmath \mathbf{t}_I \in \mathbb{C}^d$ is

$$\begin{aligned} \mathrm{Cov}[\,\mathbf{t}\,] &= E[\,(\mathbf{t} - E[\mathbf{t}])(\mathbf{t} - E[\mathbf{t}])^H\,] \\ &= \mathrm{Cov}[\,\mathbf{t}_R\,] + \mathrm{Cov}[\,\mathbf{t}_I\,] + \jmath\{\mathrm{Cov}[\mathbf{t}_I, \mathbf{t}_R] - \mathrm{Cov}[\mathbf{t}_R, \mathbf{t}_I]\} \end{aligned}$$

and the *pseudo-covariance matrix* [14] of $\mathbf{t}$ is

$$\begin{aligned} \mathrm{Pcov}[\,\mathbf{t}\,] &= E[(\mathbf{t} - E[\mathbf{t}])(\mathbf{t} - E[\mathbf{t}])^T] \\ &= \mathrm{Cov}[\,\mathbf{t}_R\,] - \mathrm{Cov}[\,\mathbf{t}_I\,] + \jmath\{\mathrm{Cov}[\mathbf{t}_I, \mathbf{t}_R] + \mathrm{Cov}[\mathbf{t}_R, \mathbf{t}_I]\}. \end{aligned}$$

Note that $\mathrm{Cov}[\,\mathbf{t}\,]$ is hermitian and positive semidefinite complex $d \times d$ matrix and $\mathrm{Pcov}[\,\mathbf{t}\,]$ is symmetric complex $d \times d$ matrix. Write

$$\widehat{\mathbf{t}} = (\mathbf{t}^T, \mathbf{t}^H)^T$$

for the *augmented vector* of $\mathbf{t}$ formed by stacking $\mathbf{t}$ and its complex conjugate $\mathbf{t}^*$. Note that $\bar{\mathbf{t}} = (\mathbf{t}_R^T, \mathbf{t}_I^T)^T = \mathbf{M}_{2d}\widehat{\mathbf{t}}$. It is well-known (e.g. [10]) that

$$\langle \mathrm{Cov}[\,\bar{\mathbf{t}}\,] \rangle_{\mathbb{C}} = \begin{pmatrix} \mathrm{Cov}[\,\mathbf{t}\,] & \mathrm{Pcov}[\,\mathbf{t}\,] \\ \mathrm{Pcov}[\,\mathbf{t}\,]^* & \mathrm{Cov}[\,\mathbf{t}\,]^* \end{pmatrix} = \mathrm{Cov}[\widehat{\mathbf{t}}], \tag{5}$$

i.e. operator $\langle \cdot \rangle_{\mathbb{C}}$ maps the covariance matrix of the composite real r.v. $\bar{\mathbf{t}}$ to the covariance matrix of the augmented r.v. $\widehat{\mathbf{t}}$.

As in [12], we define the complex gradient operator as $\nabla_{\boldsymbol{\theta}^*} = (\partial/\partial \boldsymbol{\theta}^*)^T = (\partial/\partial \theta_1^*, \ldots, \partial/\partial \theta_k^*)^T$. Thus

$$\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z}; \boldsymbol{\theta}) = \frac{1}{2}\Big\{\nabla_{\boldsymbol{\alpha}} \ln f(\mathbf{z}; \boldsymbol{\theta}) + \jmath \nabla_{\boldsymbol{\beta}} \ln f(\mathbf{z}; \boldsymbol{\theta})\Big\}.$$

Then, we call the $k \times k$ complex matrices,

$$\begin{aligned} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} &= E[\,\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z}; \boldsymbol{\theta}) \{\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z}; \boldsymbol{\theta})\}^H\,], \\ \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} &= E[\,\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z}; \boldsymbol{\theta}) \{\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z}; \boldsymbol{\theta})\}^T\,] \end{aligned}$$

as the (complex) *information matrix* and the *pseudo-information matrix*, respectively.

Definition 3 *Complex FIM of complex parameter* $\boldsymbol{\theta}$ *is defined as*

$$\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} & \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} \\ \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}^* & \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^* \end{pmatrix},$$

i.e. it is the augmented matrix of information matrix $\boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}$ *and pseudo-information matrix* $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}$.

It is easy to verify that $\langle \mathbf{J}_{\bar{\boldsymbol{\theta}}} \rangle_{\mathbb{C}} = 4\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}$, which together with Lemma 1[b] shows that

$$\langle \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \rangle_{\mathbb{C}} = 4\langle \mathbf{J}_{\bar{\boldsymbol{\theta}}} \rangle_{\mathbb{C}}^{-1} = \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}. \tag{6}$$

i.e. the inverse of the FIM of real parameter $\bar{\boldsymbol{\theta}}$ is mapped to inverse of the complex FIM of $\boldsymbol{\theta}$.

5. CRAMÉR-RAO LOWER BOUND FOR UNCONSTRAINED PARAMETERS

Based on equations (4) and (6), the complex form (3) of the CRB (2) can be written in a neat form as is illustrated below.

Theorem 1 *Let* $\mathbf{t} = \mathbf{t}_R + \jmath \mathbf{t}_I \in \mathbb{C}^d$ *be the complex valued statistic. The CRB on the covariance matrix of* $\bar{\mathbf{t}} = (\mathbf{t}_R^T, \mathbf{t}_I^T)^T$ *is equivalent with the following bound on* $\widehat{\mathbf{t}} = (\mathbf{t}^T, \mathbf{t}^H)^T$ *:*

$$\mathrm{Cov}[\,\bar{\mathbf{t}}\,] \geq \mathbf{D}_{\bar{\mathbf{g}}}\, \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1}\, \mathbf{D}_{\bar{\mathbf{g}}}^T \;\Leftrightarrow\; \mathrm{Cov}[\,\hat{\mathbf{t}}\,] \geq \boldsymbol{\mathcal{D}}_{\mathbf{g}}\, \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\, \boldsymbol{\mathcal{D}}_{\mathbf{g}}^H. \tag{7}$$

Statistic $\mathbf{t}$ *attains the CRB in that* $\mathrm{Cov}[\,\bar{\mathbf{t}}\,] = \mathbf{D}_{\bar{\mathbf{g}}} \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \mathbf{D}_{\bar{\mathbf{g}}}^T$ *if and only if* $\mathrm{Cov}[\,\widehat{\mathbf{t}}\,] = \boldsymbol{\mathcal{D}}_{\mathbf{g}} \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} \boldsymbol{\mathcal{D}}_{\mathbf{g}}^H$.

Using the well-known result for the inverse of a partitioned matrix, we may write

$$\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} & \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} \\ \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}^* & \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^* \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1} & -\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}} \\ -\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^H \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1} & \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-*} \end{pmatrix}, \tag{8}$$

where $\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}} = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} - \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-*} \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}^*$ and $\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}} = \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-*}$. The following corollary to Theorem 1 is worth pointing out.

Corollary 1 *Let* $\mathbf{t}$ *be unbiased estimator of* $\boldsymbol{\theta}$*, i.e.* $\boldsymbol{\theta} = E[\mathbf{t}]$.
[a] *Then*

$$\mathrm{Cov}[\,\bar{\mathbf{t}}\,] \geq \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \;\Leftrightarrow\; \mathrm{Cov}[\,\hat{\mathbf{t}}\,] \geq \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} \tag{9}$$

and the CRB is attained in that $\mathrm{Cov}[\,\bar{\mathbf{t}}\,] = \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \Leftrightarrow \mathrm{Cov}[\,\widehat{\mathbf{t}}\,] = \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} \Leftrightarrow \mathrm{Cov}[\,\mathbf{t}\,] = \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}$ *and* $\mathrm{Pcov}[\,\mathbf{t}\,] = -\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}$.
[b] *If pseudo-information matrix vanishes, i.e.* $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} = \mathbf{0}$*, then*

$$\mathrm{Cov}[\,\bar{\mathbf{t}}\,] \geq \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \;\Leftrightarrow\; \mathrm{Cov}[\,\mathbf{t}\,] \geq \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-1}$$

and the CRB is attained in that $\mathrm{Cov}[\,\bar{\mathbf{t}}\,] = \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} \Leftrightarrow \mathrm{Cov}[\,\mathbf{t}\,] = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-1}$.

For example, in the scalar parameter case, a statistic $t = t(\mathbf{z}) \in \mathbb{C}$ that is an unbiased estimator of $\theta \in \mathbb{C}$ satisfies

$$\begin{pmatrix} \mathrm{Var}[\,t\,] & \mathrm{Pvar}[\,t\,] \\ \mathrm{Pvar}[\,t\,]^* & \mathrm{Var}[\,t\,] \end{pmatrix} \geq \frac{1}{i_\theta^2 - |p_\theta|^2} \begin{pmatrix} i_\theta & -p_\theta \\ -p_\theta{}^* & i_\theta \end{pmatrix}$$

where $i_\theta = E[|\frac{\partial}{\partial \theta^*} \ln f(\mathbf{z};\theta)|^2]$, $p_\theta = E[(\frac{\partial}{\partial \theta^*} \ln f(\mathbf{z};\theta))^2]$ and $\mathrm{Pvar}[\,t\,] = E[(t - E[t])^2]$. Thus, $\mathrm{Var}[\,t\,] = E[|t - E(t)|^2] \geq i_\theta/(i_\theta^2 - |p_\theta|^2)$.

Corollary 2 *Eq.* (7) *implies the following bound on the covariance matrix* $\mathrm{Cov}[\,\mathbf{t}\,]$ *of* $\mathbf{t}$*:*

$$\mathrm{Cov}[\,\mathbf{t}\,] \geq \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1} \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right]^H - \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-*} \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*}\right]^H + \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^H \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1} \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right]^H + \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1} \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}} \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*}\right]^H.$$

If $\mathbf{t}$ *is an ubiased estimator of* $\boldsymbol{\theta} \in \mathbb{C}^d$*, i.e.* $\boldsymbol{\theta} = E[\,\mathbf{t}\,]$*, then eq.* (9) *implies that*

$$\mathrm{Cov}[\,\mathbf{t}\,] \geq \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}$$

and that $\sum_{i=1}^d \mathrm{Var}[\,t_i\,] \geq \mathrm{tr}\{\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}\}$.

Remark 1. Corollary 2 gives a bound *solely* on the covariance matrix $\mathrm{Cov}[\,\mathbf{t}\,]$ of the statistic $\mathbf{t}$. If an unbiased estimator $\mathbf{t}$ attains the bound on the covariance matrix alone in that $\mathrm{Cov}[\,\mathbf{t}\,] = \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}$, it does not imply that $\mathbf{t}$ attains the CRB (9) since also $\mathrm{Pcov}[\,\mathbf{t}\,] = -\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}$ needs to hold (c.f. Corollary 1[a]). Only if the pseudo-information vanishes, *i.e.* $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} = \mathbf{0}$, then $\mathrm{Cov}[\,\mathbf{t}\,] = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-1}$ implies that $\mathbf{t}$ attains the CRB (c.f. Corollary 1[b]).

Note that the bound on the covariance matrix in Corollary 2 is essentially the same as the bound in [6] derived by imitating the proof in the real case. Our derivation avoids technical difficulties associated with such approach such as the legitimity of changing operations of complex interegration and differentiation.

Theorem below characterizes the case when an unbiased estimator that attains the CRB may be found in terms of the complex score function.

Theorem 2 *Assume statistic* $\mathbf{t} = \mathbf{t}_R + \jmath \mathbf{t}_I$ *is an unbiased estimator of* $\boldsymbol{\theta}$*. Then,* $\mathbf{t}$ *attains the* CRB *if and only if*

$$\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z};\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}(\mathbf{t} - \boldsymbol{\theta}) + \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}(\mathbf{t} - \boldsymbol{\theta})^*. \tag{10}$$

If pseudo-information vanishes, i.e. $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} = \mathbf{0}$*, then* $\mathbf{t}$ *attains the* CRB *if and only if* $\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z};\boldsymbol{\theta}) = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}(\mathbf{t} - \boldsymbol{\theta})$.

As an example, consider least squares (LS) estimation of the parameter $\boldsymbol{\theta} \in \mathbb{C}^d$ in the regression model $\mathbf{z} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \in \mathbb{C}^n$, where $\mathbf{X} \in \mathbb{C}^{n \times d}$ is fixed (known) full column rank design matrix and the error term $\boldsymbol{\varepsilon} \in \mathbb{C}^n$ has generaralized complex Gaussian distribution [10] with $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $\mathrm{Cov}[\,\boldsymbol{\varepsilon}\,] = \sigma^2 \mathbf{I}_n$ and $\mathrm{Pcov}[\,\boldsymbol{\varepsilon}\,] = \tau \mathbf{I}_n$, where $\tau \in \mathbb{C}$ and $\sigma^2 > 0$ are known scalars. If $\tau = 0$, then $\boldsymbol{\varepsilon}$ has conventional circular Gaussian distribution. Then p.d.f. of $\mathbf{z}$ then is $f(\mathbf{z};\boldsymbol{\theta}) = \pi^{-n}|\boldsymbol{\Gamma}|^{-1/2}\exp\{-\frac{1}{2}\hat{\varepsilon}^H \boldsymbol{\Gamma}^{-1}\hat{\varepsilon}\}$, where $\boldsymbol{\Gamma}$ is the augmented matrix of $\mathrm{Cov}[\,\boldsymbol{\varepsilon}\,]$ and $\mathrm{Pcov}[\,\boldsymbol{\varepsilon}\,]$ and $\hat{\varepsilon}$ is the augmented vector of $\boldsymbol{\varepsilon} = \mathbf{z} - \mathbf{X}\boldsymbol{\theta}$. The LS-estimate (LSE) of $\boldsymbol{\theta}$ is $\mathbf{t} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{z}$ and it is unbiased; see [1], Chapter 15 for details. The log-likelihood (ignoring the additive contant terms not depending on the unknown parameter) is

$$\ln f(\mathbf{z};\boldsymbol{\theta}) = -b\boldsymbol{\varepsilon}^H\boldsymbol{\varepsilon} + \frac{b}{2}\varrho^*\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} + \frac{b}{2}\varrho\boldsymbol{\varepsilon}^H\boldsymbol{\varepsilon}^*,$$

where $b = \sigma^2/(\sigma^4 - |\tau|^2)$ and $\varrho = \tau/\sigma^2$. The complex gradient of $\ln f$ becomes

$$\nabla_{\boldsymbol{\theta}^*} \ln f(\mathbf{z};\boldsymbol{\theta}) = b\{\mathbf{X}^H\boldsymbol{\varepsilon} - \varrho\mathbf{X}^H\boldsymbol{\varepsilon}^*\}$$

and $\boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} = b\mathbf{X}^H\mathbf{X}$ and $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} = -b\varrho\mathbf{X}^H\mathbf{X}^*$ are obtained for information matrix and pseudo-information matrix, respectively. If $\tau = 0$, then $\boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} = \mathbf{0}$, and it is easy to see that LSE $\mathbf{t}$ attains the bound on the covariance matrix (since $\mathrm{Cov}[\,\mathbf{t}\,] = \sigma^2(\mathbf{X}^H\mathbf{X})^{-1} = \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^{-1}$) and hence due to Corollary 1[b] it attains the CRB. However, if $\tau \neq 0$ (*i.e.* the error terms have non-circular Gaussian distribution), the CRB is not attained. Indeed $\boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}$ is equal to $\mathrm{Cov}[\,\mathbf{t}\,]$ only if $\tau = 0$.

6. CRAMER-RAO LOWER BOUND FOR CONSTRAINED PARAMETERS

In the constrained parameter problem, $l < k$ complex constaints are imposed on $\boldsymbol{\theta}$ as follows

$$\mathbf{c}(\boldsymbol{\theta}) = \mathbf{a}(\boldsymbol{\alpha},\boldsymbol{\beta}) + \jmath\mathbf{b}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathbf{0}.$$

Assume that the Jacobian matrix of $\bar{\mathbf{c}}(\bar{\boldsymbol{\theta}}) = (\mathbf{a}(\bar{\boldsymbol{\theta}})^T, \mathbf{b}(\bar{\boldsymbol{\theta}})^T)^T$, denoted by $\mathbf{D}_{\bar{\mathbf{c}}} \in \mathbb{R}^{2l \times 2k}$, exists and has full rank $2l$. In the constrained case, the *constrained CRB* [3, 4] gives the lower bound on the covariance matrix of $\bar{\mathbf{t}}$ by stating that under some regularity conditions

$$\mathrm{Cov}[\,\bar{\mathbf{t}}\,] \geq \mathbf{D}_{\bar{\mathbf{g}}}\,\mathrm{CRB}(\bar{\boldsymbol{\theta}})\,\mathbf{D}_{\bar{\mathbf{g}}}^T, \tag{11}$$

where

$$\mathrm{CRB}(\bar{\boldsymbol{\theta}}) = \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1} - \mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1}\mathbf{D}_{\bar{\mathbf{c}}}^T\left(\mathbf{D}_{\bar{\mathbf{c}}}\,\mathbf{J}_{\bar{\boldsymbol{\theta}}}^{-1}\mathbf{D}_{\bar{\mathbf{c}}}^T\right)^{-1}\mathbf{D}_{\bar{\mathbf{c}}}\,\mathbf{J}_{\boldsymbol{\theta}}^{-1}.$$

It follows by Lemma 1 and (4) that (11) is equivalent with

$$\langle\mathrm{Cov}[\,\bar{\mathbf{t}}\,]\rangle_{\mathbb{C}} = \mathrm{Cov}[\,\widehat{\mathbf{t}}\,] \geq \boldsymbol{\mathcal{D}}_{\mathbf{g}}\,\langle\mathrm{CRB}(\bar{\boldsymbol{\theta}})\rangle_{\mathbb{C}}\,\boldsymbol{\mathcal{D}}_{\mathbf{g}}^H.$$

Similarly, using Lemma 1 and eq.'s (4) and (6) we may write $\mathrm{CRB}(\bar{\boldsymbol{\theta}})$ into a following complex form

$$\langle\mathrm{CRB}(\bar{\boldsymbol{\theta}})\rangle_{\mathbb{C}} = \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} - \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{D}}_{\mathbf{c}}^T\left(\boldsymbol{\mathcal{D}}_{\mathbf{c}}\,\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\mathcal{D}}_{\mathbf{c}}^H\right)^{-1}\boldsymbol{\mathcal{D}}_{\mathbf{c}}\,\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1},$$

where $\boldsymbol{\mathcal{D}}_{\mathbf{c}}$ is the $2l \times 2k$ complex Jacobian of $\mathbf{c}$, *i.e.* the augmented matrix of $\partial\mathbf{c}/\partial\boldsymbol{\theta}$ and $\partial\mathbf{c}/\partial\boldsymbol{\theta}^*$. See also [8] who consider the case that FIM is singular, *i.e.* not invertible.

7. CONCLUSIONS

Complex form of the unconstrained and constrained CRB was derived. Derived complex form of the CRB is based on complex FIM and complex Jacobian matrix. The calculation of these quantities is essentially similar to the real case. Hence the complex CRB is oftentimes much easier to calculate than its real form which looses the simple description of the statistic/parameters offered by complex number notations.

8. PROOFS OF THE RESULTS

Proof of Lemma 1. Properties [a]-[e] are immediate or almost immediate.

[f] It is easy to verify that the eigenvalues of $\langle \mathbf{G} \rangle_{\mathbb{C}}$ are twice the eigenvalues of $\mathbf{G}$. Next recall that any complex hermitian (or real symmetric) matrix is positive semidefinite $\Leftrightarrow$ all its eigenvalues are non-negative (e.g. Th. 7.2.1. in [15]). This gives the claim.

Proof of Corollary 2. If a matrix $\mathbf{A}$ is positive semidefinite, then any principal sub-matrix of $\mathbf{A}$ is positive semidefinite as well (c.f. [15]). Therefore,

$$[\mathrm{Cov}[\,\widehat{\mathbf{t}}\,]]_{1:d} \geq [\boldsymbol{\mathcal{D}}_{\mathbf{g}} \boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} \boldsymbol{\mathcal{D}}_{\mathbf{g}}^{H}]_{1:d}, \tag{12}$$

where $[\mathbf{A}]_{1:d}$ denotes the principal $d \times d$ sub-matrix of matrix $\mathbf{A}$ obtained by including the rows and columns $1, \ldots, d$. Now, $[\mathrm{Cov}[\,\widehat{\mathbf{t}}\,]]_{1:d} = \mathrm{Cov}[\,\mathbf{t}\,]$ and the matrix on the RHS of (12) can be found to be of the stated form after substituting the expression (8) for $\boldsymbol{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}$ in (12). If $\mathbf{t}$ is unbiased estimator, *i.e.* $\mathbf{g}(\boldsymbol{\theta}) = \boldsymbol{\theta}$, then $\partial \mathbf{g}/\partial \boldsymbol{\theta} = \partial \boldsymbol{\theta}/\partial \boldsymbol{\theta} = \mathbf{I}_d$ and $\partial \mathbf{g}/\partial \boldsymbol{\theta}^* = \partial \boldsymbol{\theta}/\partial \boldsymbol{\theta}^* = \mathbf{0}$ and thus the expression simplifies to $\mathrm{Cov}[\,\mathbf{t}\,] \geq \boldsymbol{\mathcal{R}}_{\boldsymbol{\theta}}^{-1}$. □

Proof of Theorem 2. Recall that a statistic $\bar{\mathbf{t}} = (\mathbf{t}_R^T, \mathbf{t}_I^T)^T \in \mathbb{R}^{2k}$ that is an unbiased estimator of $\bar{\boldsymbol{\theta}} = (\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T)^T \in \mathbb{R}^{2k}$ attains the CRB if and only if (see e.g. [1]) $\nabla_{\bar{\boldsymbol{\theta}}} \ln f = \mathbf{J}_{\bar{\boldsymbol{\theta}}}(\bar{\mathbf{t}} - \bar{\boldsymbol{\theta}})$. Thus,

$$\nabla_{\bar{\boldsymbol{\theta}}} \ln f = (1/2)\mathbf{M}_{2k} \cdot [2\mathbf{M}_{2k}^{-1}\mathbf{J}_{\bar{\boldsymbol{\theta}}}\mathbf{M}_{2k}] \cdot [\mathbf{M}_{2k}^{-1}(\bar{\mathbf{t}} - \bar{\boldsymbol{\theta}})], \ i.e.$$

$$\begin{pmatrix} \nabla_{\boldsymbol{\alpha}} \ln f \\ \nabla_{\boldsymbol{\beta}} \ln f \end{pmatrix} = \begin{pmatrix} \mathbf{I}_k & \mathbf{I}_k \\ -\jmath \mathbf{I}_k & \jmath \mathbf{I}_k \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}} & \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}} \\ \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}^* & \boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}^* \end{pmatrix} \cdot \begin{pmatrix} \mathbf{t} - \boldsymbol{\theta} \\ (\mathbf{t} - \boldsymbol{\theta})^* \end{pmatrix}.$$

Hence

$$\nabla_{\boldsymbol{\theta}^*} \ln f = \frac{1}{2} \begin{pmatrix} \mathbf{I}_k & \jmath \mathbf{I}_k \end{pmatrix} \begin{pmatrix} \nabla_{\boldsymbol{\alpha}} \ln f \\ \nabla_{\boldsymbol{\beta}} \ln f \end{pmatrix}$$

is equal to $\boldsymbol{\mathcal{I}}_{\boldsymbol{\theta}}(\mathbf{t} - \boldsymbol{\theta}) + \boldsymbol{\mathcal{P}}_{\boldsymbol{\theta}}(\mathbf{t} - \boldsymbol{\theta})^*$. □

9. REFERENCES

[1] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice-Hall, New Jersey, 1993.

[2] P.J. Bickel and K. A. Doksum, *Mathematical statistics - basic ideas and selected topics VOL. I*, Prentice-Hall, NJ: Upper Saddle river, 2001.

[3] J. D. Gorman and A. O. Hero, “Lower bound for parametric estimation with constraints,” *IEEE Trans. Inform. Theory*, vol. 26, no. 6, pp. 1285–1301, 1990.

[4] T. L. Marzetta, “A simple derivation of the constrained multiple parameter cramer-rao bound,” *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2247 – 2249, 1993.

[5] S. F. Yau and Y. Bresler, “A compact Cramér-Rao bound expression for parametric estimation of superimposed signals,” *IEEE Trans. Signal Processing*, vol. 40, pp. 1226–1230, 1992.

[6] A. van den Bos, “A Cramér-Rao lower bound for complex parameters,” *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2859, 1994.

[7] E. de Carvalho, J. Cioffi, and D. T. M. Slock, “Cramer-Rao bounds for blind multichannel estimation,” in *Proc. IEEE Global telecommunications conference*, San Francisco, USA, Nov. 27 – Dec. 1, 2000.

[8] A. K. Jagannatham and B. D. Rao, “Cramer-rao lower bound for constrained complex parameters,” *IEEE Singal Proc. Lett.*, vol. 11, no. 11, pp. 875–878, 2004.

[9] S. T. Smith, “Statistical resolution limits and the complexified cram’er-rao bound,” *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1597 – 1609, 2005.

[10] B. Picinbono, “Second order complex random vectors and normal distributions,” *IEEE Trans. Signal Processing*, vol. 44, no. 10, pp. 2637–2640, 1996.

[11] L.V. Ahlfors, *Complex analysis*, McGraw-Hill, New York, 1953.

[12] D. H. Brandwood, “A complex gradient operator and its applications in adaptive array theory,” *IEE Proc. F and H*, vol. 1, pp. 11–16, 1983.

[13] A. Hjorungnes and D. Gesbert, “Complex-valued matrix differentiation: techniques and key results,” *IEEE Trans. Signal Processing*, vol. 55, pp. 2740–2746, 2007.

[14] F. D Neeser and J. L. Massey, “Proper complex random processes with applications to information theory,” *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.

[15] R. A. Horn and C. A. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.