The Cramér-Rao Bound and Its Application to Quantification in MRS

Qiang Ning
Department of Electrical and Computer Engineering
Beckman Institute for Advanced Science and Techonology
University of Illinois at Urbana-Champaign

Abstract

Cramer-Rao Bound (CRB) is an important inequality on the error correlation matrix of an estimator. It describes the theoretical lower bound for an estimator (usually unbiased). Therefore, it is a useful index of how efficiently an estimator is. Here we will summarize the derivation of CRB, list some examples and apply this analysis to the quantification (parameters estimation) of spin-echo signals in MRI.

I. Cramer-Rao Bound

This section and the following examples section are essentially based on Levy's book *Principles of signal detection and parameter estimation* [1].

A. Definition

Let the parameters of an estimator be an m dimensional vector x, and the measurement data be an n dimensional vector Y. The estimator here is denoted by $\hat{X}(Y)$. The Cramer-Rao Bound that we usually used for unbiased estimators is

$$C_E(x) \geq J^{-1}(x), \tag{1}$$

where

$$C_E(x) = E[(x - \hat{X}(Y))(x - \hat{X}(Y))^T] \tag{2}$$

is the error correlation matrix of X, and

$$J(x) = E_Y[\bigtriangledown_x \ln f_Y(Y|x)(\bigtriangledown_x \ln f_Y(Y|x))^T], \tag{3}$$

or equivalently (which is proved in [1]),

$$J(x) = -E_Y[\bigtriangledown_x \bigtriangledown_x^T \ln f_Y(Y|x)], \tag{4}$$

is the so-called Fisher Information in Y about the parameters in X. Note in (3) and (4), $f_Y(Y|X)$ is the conditional probability distribution function of measurement Y on parameters X, and the gradient operator is defined as

$$\bigtriangledown_x = \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \cdots & \frac{\partial}{\partial x_m} \end{bmatrix}^T. \tag{5}$$

In practice, the Fisher Information can be calculated entry-wisely in the following two ways.

$$\begin{aligned} J_{i,j}(x) &= E_Y[\frac{\partial}{\partial x_i}\ln f_Y(Y|x)\frac{\partial}{\partial x_j}\ln f_Y(Y|x)] \\ &= -E_Y[\frac{\partial^2}{\partial x_i \partial x_j}\ln f_Y(Y|x)]. \end{aligned} \tag{6}$$

Given this bound, we have the estimator variance of each parameter,

$$E[(x_i - \hat{x}_i(Y))^2] \geq [J^{-1}]_{ii}(x), \tag{7}$$

which is useful in the analysis of the efficiency of an estimator.

B. Proof

We prove a more generalized form:

$$C_E(x) \geq b(x)b^T(x) + (I_m - \bigtriangledown_X^T b(x))J^{-1}(x)(I_m - \bigtriangledown_X^T b(x))^T, \tag{8}$$

where $b(x)$ is the bias of this estimator,

$$b(x) = x - E(\hat{X}(Y)). \tag{9}$$

It is obvious that when the estimator is unbiased, that is, $b(x) = 0$, (8) degenerates to (1).

Form a $2m$ dimensional vector

$$Z = \begin{bmatrix} x - \hat{X}(Y) - b(x) \\ \bigtriangledown_x \ln(f_Y(Y|x)) \end{bmatrix}. \tag{10}$$

Then let

$$C_Z = E[ZZ^T] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{11}$$

be its correlation matrix. From (2) and (3), we can immediately have

$$C_{11} = C_E(x) - b(x)b^T(x), \tag{12}$$

$$C_{22} = J(x). \tag{13}$$

It remains to evaluate

$$\begin{aligned} C_{21}^T &= C_{12} = E[(x - \hat{X}(Y) - b(x)) \bigtriangledown_X^T \ln(f_Y(Y|x))] \\ &= \int (x - \hat{X}(y) - b(x)) \bigtriangledown_x^T (f_Y(y|x))dy. \end{aligned} \tag{14}$$

Note

$$\begin{aligned} 0 &= \bigtriangledown_x^T \{[x - \hat{X}(y) - b(x)]f_Y(y|x)\} \\ &= f_Y(y|x) \bigtriangledown_x^T [x - \hat{X}(y) - b(x)] + [x - \hat{X}(y) - b(x)] \bigtriangledown_x^T f_Y(y|x) \\ &= f_Y(y|x)[I_m - \bigtriangledown_x^T b(x)] + [x - \hat{X}(y) - b(x)] \bigtriangledown_x^T f_Y(y|x). \end{aligned} \tag{15}$$

Integrating this identity with respect to y gives

$$C_{12} = -I_m + \bigtriangledown_x^T b(x). \tag{16}$$

Assuming the information matrix $C_{22} = J(x)$ is positive definite (i.e., invertible, which is usually true) at point X, we can perform Schur decomposition to $C_Z(x)$:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} I_m & C_{12}C_{22}^{-1} \\ 0 & I_m \end{pmatrix} \begin{pmatrix} C_{11} - C_{12}C_{22}^{-1}C_{21} & 0 \\ 0 & C_{22} \end{pmatrix} \begin{pmatrix} I_m & 0 \\ C_{22}^{-1}C_{21} & I_m \end{pmatrix}.$$

Because the correlation matrix $C_Z(x) \geq 0$, $C_{11} - C_{12}C_{22}^{-1}C_{21}$ is also non-negative definite, i.e.,

$$C_E(x) - b(x)b^T(x) - (I_m - \bigtriangledown_x^T b(x))J^{-1}(x)(I_m - \bigtriangledown_x^T b(x))^T \geq 0, \tag{17}$$

which is exactly (8).

□

C. *Example*

1) CRB Calculation: Consider the estimation of the mean and variance of a sequence $\{Y_k, 1 \leq k \leq N\}$ of i.i.d. $N(m, v)$ gaussian random variables.

$$f_Y(y_k|m,v) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{(y_k-m)^2}{2v}}, k = 1, 2, \ldots, N,$$

$$f_Y(\mathbf{y}|m,v) = \frac{1}{(\sqrt{2\pi v})^N} e^{-\frac{\sum_{k=1}^N (y_k-m)^2}{2v}},$$

$$L(\mathbf{y}|m,v) = \ln(f_Y(\mathbf{y}|m,v)) = -\frac{N}{2}\ln(2\pi v) - \frac{1}{2v}\sum_{k=1}^N (y_k - m)^2.$$

First order derivatives are

$$\frac{\partial L}{\partial m} = \frac{1}{v}\sum_{k=1}^N (y_k - m), \tag{18}$$

and

$$\frac{\partial L}{\partial v} = -\frac{N}{2v} + \frac{1}{2v^2}\sum_{k=1}^N (y_k - m)^2. \tag{19}$$

Then second derivatives are

$$\frac{\partial^2 L}{\partial m^2} = -\frac{N}{v},$$

$$\frac{\partial^2 L}{\partial v^2} = \frac{N}{2v^2} - \frac{1}{v^3}\sum_{k=1}^N (y_k - m)^2,$$

and

$$\frac{\partial^2 L}{\partial m \partial v} = -\frac{1}{v^2}\sum_{k=1}^N (y_k - m).$$

Therefore the Fisher Information matrix is

$$J(m,v) = N \begin{pmatrix} v^{-1} & 0 \\ 0 & (2v^2)^{-1} \end{pmatrix}, \tag{20}$$

$$J^{-1}(m,v) = N^{-1} \begin{pmatrix} v & 0 \\ 0 & 2v^2 \end{pmatrix}. \tag{21}$$

Then the error variances of the entries of any unbiased estimator must satisfy

$$E[(m - \hat{m}(Y))^2] \geq \frac{v}{N}, \tag{22}$$

$$E[(v - \hat{v}(Y))^2] \geq \frac{2v^2}{N}. \tag{23}$$

2) Efficiency: An unbiased estimator is efficient if it reaches its CRB, which means $C_E(X) - J^{-1}(X) = 0$. Further if only a partial of the eigenvalues of $C_E(X) - J^{-1}(X)$ is zero, it is called partially efficient.

By setting (18) and (19) to be zero, we have

$$\hat{m}_{ML}(Y) = \frac{1}{N}\sum_{k=1}^{N} y_k, \tag{24}$$

and

$$\hat{v}_{ML}(Y) = \frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{m})^2. \tag{25}$$

It is proved in [1] that

$$E[\hat{m}_{ML}(Y)] = m, \tag{26}$$

$$E[\hat{v}_{ML}(Y)] = \frac{N-1}{N}v, \tag{27}$$

$$E[(m - \hat{m}_{ML}(Y))^2] = \frac{v}{N}, \tag{28}$$

$$E[(v - \hat{v}_{ML}(Y))^2] = \frac{2N-1}{N^2}v^2. \tag{29}$$

From the results we see that the unbiased estimator $\hat{m}_{ML}(Y)$ is efficient because it reaches its CRB as defined in (21). Even though the estimator $\hat{v}_{ML}(Y)$ is biased, its error variance, however, is smaller than the CRB for unbiased estimators. This implies that it is not necessary that biased estimators have poor error variance performances, instead allowing a small bias can sometimes be benificial.

Another fact mentioned in [1] on pp. 145 is that all ML estimators have these two properties:

1) Asymptotically unbiased, i.e., $\lim_{N\to\infty} b(X) = 0$;
2) Asymptotically efficient, i.e., $\lim_{N\to\infty} E[(x_i - \hat{x}_i(Y))^2] = [J^{-1}]_{ii}(X)$;

which can be seen from the example above.

II. Quantification of Single Spin-Echo Signals

A. Formulation

We use the model that introduced in Chap. 5.2 *Proposed formulation* [2], ignoring baseline signals.

$$s[m] = s_{metab}[m] + \xi[m], \tag{30}$$

$$\begin{aligned} s_{metab}[m] &= e^{i\phi_0}\sum_{n=1}^{N} a_n(TE)\varphi_{n,TE}[m]\psi_{n,d_n}[m], \\ m &= 0,1,\ldots,M-1, \end{aligned} \tag{31}$$

where $\xi[m] \sim N(0,\sigma^2)$ is a complex gaussian noise, ϕ_0 is a zero-order term phase, $a_n(TE)$ is a real positive amplitude assumed to exponentially decay with respect to TE

$$a_n(TE) = c_n e^{-TE/T_{2,n}}, \tag{32}$$

and $\varphi_{n,TE}[m]$ and $\psi_{n,d_n}[m]$ are metabolite basis function and signal decay, respectively, defined as

$$\varphi_{n,TE}[m] = \sum_{l=1}^{L_n} \alpha_{l,n}(TE)e^{-i\beta_{l,n}(TE)}e^{-i2\pi f_{l,n}(TE)m\Delta t}, \tag{33}$$

$$\psi_{n,d_n}[m] = e^{-m\Delta t/d_n}. \tag{34}$$

In the above formulations, $T_{2,n}$ is a metabolite-dependent relaxation constant, d_n is a real lineshape parameter and Δt is the sampling time. $\alpha_{l,n}(TE)$, $\beta_{l,n}(TE)$ and $f_{l,n}(TE)$ are relative amplitude, phase and frequency of the l-th resonance of the n-th metabolite which can all be determined from quantum mechanical simulations.

The parameter vector that we are going to estimate is

$$\theta = [a_1,\ldots,a_N,d_1,\ldots,d_N,\phi_0]^T. \tag{35}$$

B. Entrywise Derivation of CRB

The probability distribution function of an n-channel i.i.d. real gaussian variable $X \sim N(0,\sigma^2 I_n)$, $X \in \mathbb{C}^n$ is

$$f_X(x) = \frac{1}{(2\pi\sigma^2)^{n/2}}e^{-\frac{\|x\|^2}{2\sigma^2}}. \tag{36}$$

The likelihood function of $s[m], m = 0,\ldots,M-1$ is therefore

$$\begin{aligned} L(s[m]) &= \frac{1}{(\pi\sigma^2)^M}e^{-\frac{|\xi[m]|^2}{\sigma^2}}, \\ \ln L(s[m]) &= \text{const} - \frac{1}{\sigma^2}\sum_{m=0}^{M-1}|\xi[m]|^2, \end{aligned} \tag{37}$$

where $\xi[m] \sim N(0,\sigma^2)$. Using the result that $\frac{\partial\|z\|^2}{\partial\alpha} = z(\frac{\partial z^*}{\partial\alpha}) + z^*(\frac{\partial z}{\partial\alpha}) = 2Re\{z(\frac{\partial z^*}{\partial\alpha})\}$ where $z \in \mathbb{C}$, $\alpha \in \mathbb{R}$, we have

$$\frac{\partial \ln L}{\partial\theta_k} = -\frac{1}{\sigma^2}\sum_{m=0}^{M-1}\left\{\xi[m]\frac{\partial\xi^*[m]}{\partial\theta_k} + \xi^*[m]\frac{\partial\xi[m]}{\partial\theta_k}\right\}, \tag{38}$$

where

$$\begin{aligned}
\xi^*[m] &= s^*[m] - e^{-i\phi_0}\sum_{n=1}^{N} a_n(TE)\varphi^*_{n,TE}[m]\psi^*_{n,d_n}[m], \\
(\frac{\partial \xi[m]}{\partial a_k})^* &= \frac{\partial \xi^*[m]}{\partial a_k} = -e^{-i\phi_0}\varphi^*_{k,TE}[m]\psi^*_{k,d_k}[m], && (39)\\
(\frac{\partial \xi[m]}{\partial d_k})^* &= \frac{\partial \xi^*[m]}{\partial d_k} = -e^{-i\phi_0}a_k(TE)\varphi^*_{k,TE}[m]\frac{\partial \psi^*_{k,d_k}[m]}{\partial d_k}, && (40)\\
(\frac{\partial \xi[m]}{\partial \phi_0})^* &= \frac{\partial \xi^*[m]}{\partial \phi_0} = ie^{-i\phi_0}\sum_{n=1}^{N} a_n(TE)\varphi^*_{n,TE}[m]\psi^*_{n,d_n}[m], && (41)\\
k &= 1,2,\ldots,N.
\end{aligned}$$

By (6), we have

$$F_{p,q}(\theta) = E_\xi\left[\left(\frac{\partial \mathrm{ln}L}{\partial \theta_p}\right)\left(\frac{\partial \mathrm{ln}L}{\partial \theta_q}\right)\right] \tag{42}$$

In calculating that, we will need the expectation of $\xi^*[m_1]\xi[m_2]$ and $\xi[m_1]\xi[m_2]$, $\forall m_1, m_2 = 1,\ldots,M$. Because the gaussian noise channels are i.i.d., it is easy to get

$$\begin{aligned}
E_\xi\{\xi^*[m_1]\xi[m_2]\} &= \begin{cases}\sigma^2, & \text{for } m_1 = m_2\\ 0, & \text{for } m_1 \neq m_2\end{cases} && (43)\\
E_\xi\{\xi[m_1]\xi[m_2]\} &= 0, \text{for } m_1 \neq m_2.
\end{aligned}$$

To handle $\xi[m]\xi[m]$, we firstly decompose it into real variables:

$$\xi[m] = \xi_r[m] + i\xi_i[m], \tag{44}$$

where $\xi_r[m], \xi_i[m] \sim N(0, \frac{\sigma^2}{2})$ and are independent with each other. And then

$$E_\xi\{\xi[m]\xi[m]\} = E_\xi\{\xi_r^2 + 2i\xi_r\xi_i - \xi_i^2\} = E_\xi\{\xi_r^2 - \xi_i^2\} = 0. \tag{45}$$

Therefore

$$E_\xi\{\xi[m_1]\xi[m_2]\} = 0, \forall m_1, m_2 = 1,\ldots,M. \tag{46}$$

Taking (43) and (46) into consideration, the entries of the Fisher Information matrix is

$$\begin{aligned}
F_{p,q}(\theta) &= \frac{1}{\sigma^4}E_\xi\left\{\sum_{m_1=0}^{M-1}\left[\xi[m_1]\frac{\partial\xi^*[m_1]}{\partial\theta_p}+\xi^*[m_1]\frac{\partial\xi[m_1]}{\partial\theta_p}\right]\right. \\
&\left.\sum_{m_2=0}^{M-1}\left[\xi[m_2]\frac{\partial\xi^*[m_2]}{\partial\theta_q}+\xi^*[m_2]\frac{\partial\xi[m_2]}{\partial\theta_q}\right]\right\} \\
&= \frac{1}{\sigma^4}E_\xi\left\{\sum_{m=0}^{M-1}\left[\xi[m]\xi^*[m]\frac{\partial\xi^*[m]}{\partial\theta_p}\frac{\partial\xi[m]}{\partial\theta_q}+\xi^*[m]\xi[m]\frac{\partial\xi[m]}{\partial\theta_p}\frac{\partial\xi^*[m]}{\partial\theta_q}\right]\right\} \\
&= \frac{1}{\sigma^2}\sum_{m=0}^{M-1}\left[\frac{\partial\xi^*[m]}{\partial\theta_p}\frac{\partial\xi[m]}{\partial\theta_q}+\frac{\partial\xi[m]}{\partial\theta_p}\frac{\partial\xi^*[m]}{\partial\theta_q}\right] \\
&= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\frac{\partial\xi^*[m]}{\partial\theta_p}\frac{\partial\xi[m]}{\partial\theta_q}\right\}. \qquad (47)
\end{aligned}$$

(48)

Substituting (39)(40)(41) into (47), we get for $\forall p,q=1,\ldots,N$

$$\begin{aligned}
F_{a_p,a_q}(\theta) &= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\frac{\partial\xi^*[m]}{\partial a_p}\frac{\partial\xi[m]}{\partial a_q}\right\} \\
&= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{-e^{-i\phi_0}\varphi^*_{p,TE}[m]\psi^*_{p,d_p}[m](-e^{i\phi_0}\varphi_{q,TE}[m]\psi_{q,d_q}[m])\right\} \\
&= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\varphi^*_{p,TE}[m]\psi^*_{p,d_p}[m]\varphi_{q,TE}[m]\psi_{q,d_q}[m]\right\}, \qquad (49)
\end{aligned}$$

$$\begin{aligned}
F_{d_p,d_q}(\theta) &= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\frac{\partial\xi^*[m]}{\partial d_p}\frac{\partial\xi[m]}{\partial d_q}\right\} \\
&= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{a_pa_q\varphi^*_{p,TE}[m]\frac{\partial\psi^*_{p,d_p}[m]}{\partial d_p}\varphi_{q,TE}[m]\frac{\partial\psi_{q,d_q}[m]}{\partial d_q}\right\}, \qquad (50)
\end{aligned}$$

$$\begin{aligned}
F_{\phi_0,\phi_0}(\theta) &= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\frac{\partial\xi^*[m]}{\partial\phi_0}\frac{\partial\xi[m]}{\partial\phi_0}\right\} \\
&= \frac{2}{\sigma^2}\sum_{m=0}^{M-1}Re\left\{\sum_{n_1=1}^{N}a_{n_1}\varphi^*_{n_1,TE}[m]\psi^*_{n_1,d_{n_1}}[m]\right. \\
&\left.\sum_{n_2=1}^{N}a_{n_2}\varphi_{n_2,TE}[m]\psi_{n_2,d_{n_2}}[m]\right\}
\end{aligned}$$

$$= \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{\sum_{n_1=1}^{N}\sum_{n_2=1}^{N} a_{n_1}a_{n_2}\varphi^*_{n_1,TE}[m]\psi^*_{n_1,d_{n_1}}[m]\varphi_{n_2,TE}[m]\psi_{n_2,d_{n_2}}[m]\right\}, \quad (51)$$

$$F_{a_p,d_q}(\theta) = \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{\frac{\partial\xi^*[m]}{\partial a_p}\frac{\partial\xi[m]}{\partial d_q}\right\}$$

$$= \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{a_q\varphi^*_{p,TE}[m]\psi^*_{p,d_p}[m]\varphi_{q,TE}[m]\frac{\partial\psi_{q,d_q}[m]}{\partial d_q}\right\}, \quad (52)$$

$$F_{a_p,\phi_0}(\theta) = \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{\frac{\partial\xi^*[m]}{\partial a_p}\frac{\partial\xi[m]}{\partial \phi_0}\right\}$$

$$= \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{i\varphi^*_{p,TE}[m]\psi^*_{p,d_p}[m]\sum_{n=1}^{N} a_n\varphi_{n,TE}[m]\psi_{n,d_n}[m]\right\}$$

$$= -\frac{2}{\sigma^2}\sum_{m=0}^{M-1} Im\left\{\varphi^*_{p,TE}[m]\psi^*_{p,d_p}[m]\sum_{n=1}^{N} a_n\varphi_{n,TE}[m]\psi_{n,d_n}[m]\right\}, \quad (53)$$

$$F_{d_p,\phi_0}(\theta) = \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{\frac{\partial\xi^*[m]}{\partial d_p}\frac{\partial\xi[m]}{\partial \phi_0}\right\}$$

$$= \frac{2}{\sigma^2}\sum_{m=0}^{M-1} Re\left\{ia_p\varphi^*_{p,TE}[m]\frac{\partial\psi^*_{p,d_p}[m]}{\partial d_p}\sum_{n=1}^{N} a_n\varphi_{n,TE}[m]\psi_{n,d_n}[m]\right\}$$

$$= -\frac{2}{\sigma^2}\sum_{m=0}^{M-1} Im\left\{a_p\varphi^*_{p,TE}[m]\frac{\partial\psi^*_{p,d_p}[m]}{\partial d_p}\sum_{n=1}^{N} a_n\varphi_{n,TE}[m]\psi_{n,d_n}[m]\right\}. \quad (54)$$

Here we have got the entries of $F_{p,q}(\theta)$ above its diagonal, and the rest of the entries are the conjugate transpose of the upper triangular part.

$$F(\theta) = \begin{pmatrix} F_{a,a} & F_{a,d} & F_{a,\phi_0} \\ F^H_{a,d} & F_{d,d} & F_{d,\phi_0} \\ F^H_{a,\phi_0} & F^H_{d,\phi_0} & F_{\phi_0,\phi_0} \end{pmatrix}. \quad (55)$$

C. Matrix Derivation of CRB

In deriving the CRB entrywisely, we see the formulations are much complicated due to a large number of summations. If we utilize matrix derivatives, the labor can be greatly reduced.

Rewrite the signal model in matrix form.

$$\mathbf{s} = e^{i\phi_0}\mathbf{Z}\mathbf{a} + \xi, \quad (56)$$

where $\mathbf{Z}_{m,n} = \varphi_{n,TE}[m]\psi_{d,d_n}[m],\ m = 0, \ldots, M-1,\ n = 1, \ldots, N$ and

$$\mathbf{s} = \begin{pmatrix} s[0] \\ s[1] \\ \vdots \\ s[M-1] \end{pmatrix}, \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}, \mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{pmatrix}, \xi = \begin{pmatrix} \xi[0] \\ \xi[1] \\ \vdots \\ \xi[M-1] \end{pmatrix}. \tag{57}$$

I will not keep these variables bold face hereafter for simplicity. Then the likelihood function is

$$\ln L(s) = \text{const} - \frac{1}{\sigma^2}\|\xi\|^2. \tag{58}$$

According to my previous weekly summary (week 12) of the gradient calculation of least squares problems, we have

$$\triangledown_\theta \ln L = -\frac{1}{\sigma^2}(J^H\xi + (J^H\xi)^*), \tag{59}$$

where J is the jacobian matrix of ξ over θ (denote the Fisher Information matrix by another notation F to avoid conflict).

$$\begin{aligned} J &= \frac{\partial \xi}{\partial \theta} \\ &= \left[\ \frac{\partial \xi}{\partial a} \quad \frac{\partial \xi}{\partial d} \quad \frac{\partial \xi}{\partial \phi_0}\ \right] \\ &= \left[\ -e^{i\phi_0}Z \quad -e^{i\phi_0}DA \quad -ie^{i\phi_0}Za\ \right], \end{aligned} \tag{60}$$

where

$$D = \left[\ \frac{\partial Z_1}{\partial d_1} \quad \ldots \quad \frac{\partial Z_N}{\partial d_N}\ \right], \tag{61}$$

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{pmatrix}. \tag{62}$$

By (3) and (59), we have

$$\begin{aligned} F &= E\left\{(\triangledown_\theta \ln L)(\triangledown_\theta \ln L)^H\right\} \\ &= \frac{1}{\sigma^4}E\left\{(J^H\xi + (J^H\xi)^*)(\xi^H J + (\xi^H J)^*)\right\} \\ &= \frac{2}{\sigma^2}Re\{J^H J\}. \end{aligned} \tag{63}$$

Using (60), we have

$$F_{a,a}(\theta) = \frac{2}{\sigma^2} Re\{Z^H Z\}, \tag{64}$$

$$F_{d,d}(\theta) = \frac{2}{\sigma^2} Re\{A^H D^H D A\}, \tag{65}$$

$$F_{\phi_0,\phi_0}(\theta) = \frac{2}{\sigma^2} Re\{a^H Z^H Z a\}, \tag{66}$$

$$F_{a,d}(\theta) = \frac{2}{\sigma^2} Re\{Z^H D A\}, \tag{67}$$

$$F_{a,\phi_0}(\theta) = -\frac{2}{\sigma^2} Im\{Z^H Z a\}, \tag{68}$$

$$F_{d,\phi_0}(\theta) = -\frac{2}{\sigma^2} Im\{A^H D^H Z a\}. \tag{69}$$

Finally

$$F(\theta) = \begin{pmatrix} F_{a,a} & F_{a,d} & F_{a,\phi_0} \\ F_{a,d}^H & F_{d,d} & F_{d,\phi_0} \\ F_{a,\phi_0}^H & F_{d,\phi_0}^H & F_{\phi_0,\phi_0} \end{pmatrix}. \tag{70}$$

REFERENCES

[1] B. C. Levy, *Principles of signal detection and parameter estimation*. Springer, 2008, ch. 4.4, pp. 131–150.

[2] H. M. Nguyen, "Towards high-resolution magnetic resonance spectroscopic imaging: Spatiotemporal denoising and echo-time selection," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2011.