Derivation of the PCRB for Wideband Array Signal Processing Using Sequential MC Methods

William Ng, James P. Reilly*, Thia Kirubarajan, and Jean-René Larocque
Department of Electrical and Computer Engineering,
McMaster University, 1280 Main St. W.,
Hamilton, Ontario,
Canada L8S 4K1

I. Derivation of The PCRB

Let $\boldsymbol{y}$ be an observation vector, $\boldsymbol{\theta}$ be an k_0-dimensional parameter vector, $\hat{\boldsymbol{\theta}}(\boldsymbol{y})$ be a function of $\boldsymbol{y}$, which represents an estimate of $\boldsymbol{\theta}$, and $p(\boldsymbol{y},\boldsymbol{\theta})$ be the joint probability density of $(\boldsymbol{y},\boldsymbol{\theta})$. The PCRB on the estimation error on $\boldsymbol{\theta}$ has the form

$$P = E\left\{[\hat{\boldsymbol{\theta}}(\boldsymbol{y})-\boldsymbol{\theta}][\hat{\boldsymbol{\theta}}(\boldsymbol{y})-\boldsymbol{\theta}]^T\right\} \geq \boldsymbol{\mathcal{J}}^{-1}, \tag{1}$$

where $\boldsymbol{\mathcal{J}}$ is the $k_0 \times k_0$ Fisher Information matrix with elements [1]

$$[\boldsymbol{\mathcal{J}}]_{i,j} = E\left[-\frac{\partial^2 \log p(\boldsymbol{y},\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right], \quad i,j = 0,...,k_0-1, \tag{2}$$

provided that the derivatives and expectations in (1) and (2) exist. Let $\nabla_{\boldsymbol{\theta}}$ be the operator of the first-order partial derivative as follows

$$\nabla_{\boldsymbol{\theta}} = \left[\frac{\partial}{\partial\theta_0},...,\frac{\partial}{\partial\theta_{k_0-1}}\right]^T, \tag{3}$$

such that we can express $\boldsymbol{\mathcal{J}}$ as

$$\begin{aligned} \boldsymbol{\mathcal{J}} &= -E\left[\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}^T \log p(\boldsymbol{y},\boldsymbol{\theta})\right], & (4)\\ &= -E\left[\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}^T\left\{\log p(\boldsymbol{y}|\boldsymbol{\theta}) + \log p(\boldsymbol{\theta})\right\}\right], & (5)\\ &= \boldsymbol{\mathcal{J}}_D + \boldsymbol{\mathcal{J}}_P, & (6)\end{aligned}$$

where $\boldsymbol{\mathcal{J}}_D$, also recognized as the standard Fisher Information matrix [2], represents the information obtained from the data, defined as

$$\boldsymbol{\mathcal{J}}_D = -E\left[\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}^T \log p(\boldsymbol{y}|\boldsymbol{\theta})\right], \tag{7}$$

J. Reilly, corresponding author: ph: 905 525 9140 x22895, fax: 905 521 2922, email: reillyj@mcmaster.ca

and $\boldsymbol{\mathcal{J}}_P$ is the information obtained from the *a prior* information, defined as

$$\boldsymbol{\mathcal{J}}_P = -E\left[\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}^T \log p(\boldsymbol{\theta})\right]. \tag{8}$$

Assume that the parameter vector can be partitioned as follows

$$\boldsymbol{\theta} = \left[\boldsymbol{\theta}_a^T, \boldsymbol{\theta}_b^T\right]^T, \quad \boldsymbol{\theta}_a \in \mathcal{R}^{k_a \times 1}, \boldsymbol{\theta}_b \in \mathcal{R}^{k_b \times 1} \tag{9}$$

where $k_a + k_b = k_0$, and the information matrix $\boldsymbol{\mathcal{J}}$ can be partitioned into blocks as follows

$$\boldsymbol{\mathcal{J}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{aa} & \boldsymbol{\mathcal{J}}_{ab} \\ \boldsymbol{\mathcal{J}}_{ba} & \boldsymbol{\mathcal{J}}_{bb} \end{bmatrix}, \tag{10}$$

where $\boldsymbol{\mathcal{J}}_{ba} = \boldsymbol{\mathcal{J}}_{ab}^T$. The inverses of the submatrices in $\boldsymbol{\mathcal{J}}$ in (10) are the corresponding covariance matrices with other parameters fixed. It can be shown that [3] that the covariance of estimation of $\boldsymbol{\theta}_b$, P_b, is lower bounded by the right-lower block of $\boldsymbol{\mathcal{J}}^{-1}$ as follows

$$\begin{aligned} P_b &= E\left\{\left[\hat{\boldsymbol{\theta}}_b(\boldsymbol{y}) - \boldsymbol{\theta}_b\right]\left[\hat{\boldsymbol{\theta}}_b(\boldsymbol{y}) - \boldsymbol{\theta}_b\right]^T\right\}, && (11) \\ &\geq [\boldsymbol{\mathcal{J}}_{bb} - \boldsymbol{\mathcal{J}}_{ba}\boldsymbol{\mathcal{J}}_{aa}^{-1}\boldsymbol{\mathcal{J}}_{ab}]^{-1}, && (12) \\ &= \boldsymbol{\mathcal{J}}^{-1}(\boldsymbol{\theta}_b), && (13) \end{aligned}$$

provided that $\boldsymbol{\mathcal{J}}_{aa}^{-1}$ exists. The matrix $\boldsymbol{\mathcal{J}}(\boldsymbol{\theta}_b) \in \mathcal{R}^{k_b \times k_b}$, known as the *information submatrix* for parameter $\boldsymbol{\theta}_b$, is given by

$$\boldsymbol{\mathcal{J}}(\boldsymbol{\theta}_b) = \boldsymbol{\mathcal{J}}_{bb} - \boldsymbol{\mathcal{J}}_{ba}\boldsymbol{\mathcal{J}}_{aa}^{-1}\boldsymbol{\mathcal{J}}_{ab}. \tag{14}$$

Given that the state-space model as follows

$$\begin{aligned} \boldsymbol{\tau}(n) &= \boldsymbol{\tau}(n-1) + \sigma_v \boldsymbol{v}(n), && (15) \\ \boldsymbol{y}(n) &= \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}(n))\boldsymbol{a}(n-l) + \sigma_w \boldsymbol{w}(n), && (16) \end{aligned}$$

and that both $\boldsymbol{v}(n)$ and $\boldsymbol{w}(n)$ are *iid* Gaussian random variables with zero mean and unit variance, the total joint probability density function $p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)$, where $\boldsymbol{\mathcal{Y}}_n = \boldsymbol{y}_{1:n}$ and $\boldsymbol{\mathcal{T}}_n = \boldsymbol{\tau}_{1:n}$, can be given as follows

$$p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n) = p(\boldsymbol{\tau}_0)\prod_{j=1}^{n} p(\boldsymbol{y}_j|\boldsymbol{\tau}_j)\prod_{j=1}^{n} p(\boldsymbol{\tau}_j|\boldsymbol{\tau}_{j-1}), \tag{17}$$

where $p(\boldsymbol{\tau}_0)$ is assumed known. According to (4), we can derive an $nk_0 \times nk_0$ information matrix $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{T}}_n)$ from $p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)$. However, instead of computing the information matrix $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{T}}_n)$, we are more interested in computing the $k_0 \times k_0$ instantaneous information submatrix as in (14) for the parameter $\boldsymbol{\tau}_n$.

Let $\boldsymbol{\mathcal{T}}_n$ be partitioned as $\left[\boldsymbol{\mathcal{T}}_{n-1}^T, \boldsymbol{\tau}_n^T\right]^T$. Following (9)-(10), we can express $\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{T}}_n) \in \mathcal{R}^{nk_0 \times nk_0}$ as

$$\begin{aligned}
\boldsymbol{\mathcal{J}}(\boldsymbol{\mathcal{T}}_n) &= \begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{B}_n \\ \boldsymbol{B}_n^T & \boldsymbol{C}_n \end{bmatrix}, \\
&= \begin{bmatrix} -E\left[\nabla_{\boldsymbol{\mathcal{T}}_{n-1}} \nabla_{\boldsymbol{\mathcal{T}}_{n-1}}^T \log p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)\right] & -E\left[\nabla_{\boldsymbol{\tau}_n} \nabla_{\boldsymbol{\mathcal{T}}_{n-1}}^T \log p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)\right] \\ -E\left[\nabla_{\boldsymbol{\mathcal{T}}_{n-1}} \nabla_{\boldsymbol{\tau}_n}^T \log p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)\right] & -E\left[\nabla_{\boldsymbol{\tau}_n} \nabla_{\boldsymbol{\tau}_n}^T \log p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n)\right] \end{bmatrix},
\end{aligned} \tag{18}$$

provided that the derivatives and the expectations exist. As a result, according to (14), we obtain an expression of $\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_n) \in \mathcal{R}^{k_0 \times k_0}$ as follows

$$\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_n) = \boldsymbol{C}_n - \boldsymbol{B}_n^T \boldsymbol{A}_n^{-1} \boldsymbol{B}_n. \tag{19}$$

In order to get a recursive update equation of $\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_{n+1})$, given $\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_n)$ and $\boldsymbol{y}_{n+1}$, we need to first consider the joint probability function $p(\boldsymbol{\mathcal{Y}}_{n+1}, \boldsymbol{\mathcal{T}}_{n+1})$ as follows

$$\begin{aligned}
p(\boldsymbol{\mathcal{Y}}_{n+1}, \boldsymbol{\mathcal{T}}_{n+1}) &= p(\boldsymbol{y}_{n+1} | \boldsymbol{\mathcal{T}}_{n+1}, \boldsymbol{\mathcal{Y}}_n) p(\boldsymbol{\tau}_{n+1} | \boldsymbol{\mathcal{T}}_n, \boldsymbol{\mathcal{Y}}_n) p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n), \\
&= p(\boldsymbol{y}_{n+1} | \boldsymbol{\tau}_{n+1}) p(\boldsymbol{\tau}_{n+1} | \boldsymbol{\tau}_n) p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n),
\end{aligned} \tag{20}$$

where we use the fact that $\boldsymbol{\mathcal{T}}_n$ is independent of $\boldsymbol{y}_{n+1}$ and that innovations of $\boldsymbol{\tau}_n$ are independent.

Accordingly, the information matrix $\boldsymbol{\mathcal{J}}(\mathcal{T}_{n+1})$ with $\boldsymbol{\mathcal{T}}_{n+1}$ partitioned as $\left[\boldsymbol{\mathcal{T}}_n^T, \boldsymbol{\tau}_{n+1}^T\right]^T$ can be shown to be

$$\begin{aligned}
\boldsymbol{\mathcal{J}}(\mathcal{T}_{n+1}) &= -E\left[\nabla_{\boldsymbol{\mathcal{T}}_{n+1}} \nabla_{\boldsymbol{\mathcal{T}}_{n+1}}^T \log p(\boldsymbol{\mathcal{Y}}_{n+1}, \boldsymbol{\mathcal{T}}_{n+1})\right], \\
&= -E\left[\nabla_{\boldsymbol{\mathcal{T}}_{n+1}} \nabla_{\boldsymbol{\mathcal{T}}_{n+1}}^T \left\{ \log p(\boldsymbol{y}_{n+1} | \boldsymbol{\tau}_{n+1}) + \log p(\boldsymbol{\tau}_{n+1} | \boldsymbol{\tau}_n) + \log p(\boldsymbol{\mathcal{Y}}_n, \boldsymbol{\mathcal{T}}_n) \right\}\right], \\
&= \begin{bmatrix} \boldsymbol{A}_{n+1} & \boldsymbol{B}_{n+1} \\ \boldsymbol{B}_{n+1}^T & \boldsymbol{C}_{n+1} \end{bmatrix},
\end{aligned} \tag{21}$$

where the terms $\boldsymbol{A}_{n+1}$, $\boldsymbol{B}_{n+1}$, and $\boldsymbol{C}_{n+1}$ are given by

$$\boldsymbol{A}_{n+1} = \begin{bmatrix} \boldsymbol{A}_n & \boldsymbol{B}_n \\ \boldsymbol{B}_n^T & \boldsymbol{C}_n + \boldsymbol{D}_n^{11} \end{bmatrix}, \tag{22}$$

$$\boldsymbol{B}_{n+1} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{D}_n^{12} \end{bmatrix}, \tag{23}$$

$$\boldsymbol{C}_{n+1} = \boldsymbol{D}_n^{22}, \tag{24}$$

and the terms $\boldsymbol{D}_n^{11} \in \mathcal{R}^{k_0 \times k_0}$, $\boldsymbol{D}_n^{12} \in \mathcal{R}^{k_0 \times k_0}$, $\boldsymbol{D}_n^{21} \in \mathcal{R}^{k_0 \times k_0}$, and $\boldsymbol{D}_n^{22} \in \mathcal{R}^{k_0 \times k_0}$ are defined as follows

$$\boldsymbol{D}_n^{11} = E\left[-\nabla_{\boldsymbol{\tau}_n}\nabla_{\boldsymbol{\tau}_n}^T \log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right], \tag{25}$$

$$\boldsymbol{D}_n^{12} = E\left[-\nabla_{\boldsymbol{\tau}_{n+1}}\nabla_{\boldsymbol{\tau}_n}^T \log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right], \tag{26}$$

$$\boldsymbol{D}_n^{21} = \left[\boldsymbol{D}_n^{12}\right]^T, \tag{27}$$

$$\begin{aligned}\boldsymbol{D}_n^{22} =& E\left[-\nabla_{\boldsymbol{\tau}_{n+1}}\nabla_{\boldsymbol{\tau}_{n+1}}^T \log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right] + \\ & E\left[-\nabla_{\boldsymbol{\tau}_{n+1}}\nabla_{\boldsymbol{\tau}_{n+1}}^T \log p\left(\boldsymbol{y}_{n+1}|\boldsymbol{\tau}_{n+1}\right)\right].\end{aligned} \tag{28}$$

Thus the information submatrix $\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_{n+1})$ can be given by the inverse of the right-lower submatrix of $\boldsymbol{\mathcal{J}}^{-1}(\boldsymbol{\mathcal{T}}_{n+1})$ as in (14) by

$$\begin{aligned}\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_{n+1}) &= \boldsymbol{C}_{n+1} - \boldsymbol{B}_{n+1}^{n}\boldsymbol{A}_{n+1}^{-1}\boldsymbol{B}_{n+1}, \\ &= \boldsymbol{D}_n^{22} - \left[\boldsymbol{0}, \boldsymbol{D}_n^{21}\right]\left[\begin{array}{cc}\boldsymbol{A}_n & \boldsymbol{B}_n \\ \boldsymbol{B}_n^T & \boldsymbol{C}_n + \boldsymbol{D}_n^{11}\end{array}\right]^{-1}\left[\begin{array}{c}\boldsymbol{0} \\ \boldsymbol{D}_n^{12}\end{array}\right], \\ &= \boldsymbol{D}_n^{22} - \boldsymbol{D}_n^{21}\left[\boldsymbol{D}_n^{11} + \boldsymbol{C}_n - \boldsymbol{B}_n^T\boldsymbol{A}_n^{-1}\boldsymbol{B}_n\right]^{-1}\boldsymbol{D}_n^{12}, \\ &= \boldsymbol{D}_n^{22} - \boldsymbol{D}_n^{21}\left[\boldsymbol{D}_n^{11} + \boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_n)\right]^{-1}\boldsymbol{D}_n^{12},\end{aligned} \tag{29}$$

which is the desired recursive update equation of the information submatrix for $\boldsymbol{\tau}_{n+1}$. The initial information submatrix $\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_0)$ can be computed from the *a priori* probability function $p(\boldsymbol{\tau}_0)$ as follows

$$\boldsymbol{\mathcal{J}}(\boldsymbol{\tau}_0) = -E\left[\nabla_{\boldsymbol{\tau}_0}\nabla_{\boldsymbol{\tau}_0}^T \log p(\boldsymbol{\tau}_0)\right]. \tag{30}$$

II. Derivation of $\boldsymbol{D}_n^{11}$, $\boldsymbol{D}_n^{12}$, $\boldsymbol{D}_n^{21}$, and $\boldsymbol{D}_n^{22}$

Given the state-space model in (16) and that both $\boldsymbol{v}(n)$ and $\boldsymbol{w}(n)$ are *iid* Gaussian random variables with zero mean and unit variance, the functions $\log p(\boldsymbol{y}_{n+1}|\boldsymbol{\tau}_{n+1})$ and $\log p(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n)$ are given as follows

$$\begin{aligned}&\log p\left(\boldsymbol{y}_{n+1}|\boldsymbol{\tau}_{n+1}\right) = \kappa_{\sigma_w} - \\ &\frac{1}{2\sigma_w^2}\left(\boldsymbol{y}_{n+1} - \sum_{l=0}^{L-1}\tilde{\boldsymbol{H}}(\boldsymbol{\tau}_{n+1})\boldsymbol{a}_{n-l+1}\right)^T\left(\boldsymbol{y}_{n+1} - \sum_{l=0}^{L-1}\tilde{\boldsymbol{H}}(\boldsymbol{\tau}_{n+1})\boldsymbol{a}_{n-l+1}\right),\end{aligned} \tag{31}$$

$$\log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right) = \kappa_{\sigma_v} - \frac{1}{2\sigma_v^2}\left(\boldsymbol{\tau}_{n+1} - \boldsymbol{\tau}_n\right)^T\left(\boldsymbol{\tau}_{n+1} - \boldsymbol{\tau}_n\right), \tag{32}$$

where κ_{σ_w} and κ_{σ_v} are a function of σ_w^2 and σ_v^2, respectively. Next we will present the derivations of $\boldsymbol{D}_n^{11}$, $\boldsymbol{D}_n^{12}$, $\boldsymbol{D}_n^{21}$, and $\boldsymbol{D}_n^{22}$ in sequel.

A. Derivation of $\boldsymbol{D}_n^{11}$

The term $\boldsymbol{D}_n^{11}$ is defined as

$$E\left\{-\nabla_{\boldsymbol{\tau}_n}\nabla_{\boldsymbol{\tau}_n}^T \log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right\}, \tag{33}$$

whose i,jth element of $\boldsymbol{D}_n^{11}$ is defined as follows

$$\begin{aligned}[\boldsymbol{D}_n^{11}]_{i,j} &= \frac{\partial}{\partial \tau_{t,j}}\left\{\frac{-\partial}{\partial \tau_{t,i}}\log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right\}, \quad i,j=0,...,k_0-1\\ &= \frac{-1}{\sigma_v^2}\frac{\partial}{\partial \tau_{t,j}}\left\{\left(\tau_{n+1,i}-\tau_{t,i}\right)\right\},\\ &= \begin{cases}0, & \text{if } i\neq j\\ \frac{1}{\sigma_v^2}, & \text{if } i=j\end{cases}.\end{aligned} \tag{34}$$

In other words, the matrix $\boldsymbol{D}_n^{11}$ is a diagonal matrix defined as

$$\boldsymbol{D}_n^{11} = \frac{1}{\sigma_v^2}\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \dots & 0 & 1\end{bmatrix}. \tag{35}$$

B. Derivation of $\boldsymbol{D}_n^{12}$

The i,jth element of $\boldsymbol{D}_n^{12}$ is defined as follows

$$\begin{aligned}[\boldsymbol{D}_n^{12}]_{i,j} &= \frac{\partial}{\partial \tau_{n+1,j}}\left\{\frac{-\partial}{\partial \tau_{t,i}}\log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right)\right\}, \quad i,j=0,...,k_0-1\\ &= \frac{-1}{\sigma_v^2}\frac{\partial}{\partial \tau_{n+1,j}}\left\{\left(\tau_{n+1,i}-\tau_{t,i}\right)\right\},\\ &= \begin{cases}0, & \text{if } i\neq j\\ \frac{-1}{\sigma_v^2}, & \text{if } i=j\end{cases}.\end{aligned} \tag{36}$$

In other words, the matrix $\boldsymbol{D}_n^{12}$ and $\boldsymbol{D}_n^{21} = \left[\boldsymbol{D}_n^{12}\right]^T$ are a diagonal matrix defined as

$$\boldsymbol{D}_n^{12} = \boldsymbol{D}_n^{21} = \frac{-1}{\sigma_v^2}\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \dots & 0 & 1\end{bmatrix}. \tag{37}$$

C. Derivation of $\boldsymbol{D}_n^{22}$

The i,jth element of $\boldsymbol{D}_n^{22}$ is defined as follows

$$\begin{aligned}[\boldsymbol{D}_n^{22}]_{i,j} &= \frac{\partial}{\partial \tau_{n+1,j}}\left\{\frac{-\partial}{\partial \tau_{n+1,i}}\log p\left(\boldsymbol{\tau}_{n+1}|\boldsymbol{\tau}_n\right) + \frac{-\partial}{\partial \tau_{n+1,i}}\log p\left(\boldsymbol{y}_{n+1}|\boldsymbol{\tau}_{n+1}\right)\right\}, \quad i,j=0,...,k_0-1,\\ &= \frac{\partial}{\partial \tau_{n+1,j}}\left\{\frac{1}{\sigma_v^2}(\tau_{n+1,i}-\tau_{t,i}) - \frac{1}{\sigma_w^2}\boldsymbol{\epsilon}_{n+1}^T\tilde{\boldsymbol{H}}'(\tau_{n+1,i})\boldsymbol{s}_i(n+1)\right\},\\ &= \begin{cases}\frac{1}{\sigma_w^2}f_{i,j}(n+1), & \text{if } i\neq j\\ \frac{1}{\sigma_v^2}+\frac{1}{\sigma_w^2}\left(\boldsymbol{\epsilon}_{n+1}^T\tilde{\boldsymbol{H}}''(\tau_{n+1,i})\boldsymbol{s}_i(n+1)+f_{i,i}(n+1)\right), & \text{if } i=j\end{cases},\end{aligned} \tag{38}$$

where

$$\boldsymbol{\epsilon}_{n+1} = \boldsymbol{y}_{n+1} - \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}_{n+1})\boldsymbol{a}_{n-l+1}, \tag{39}$$

$$\tilde{\boldsymbol{H}}'(\tau_{n+1,i}) = \frac{\partial}{\partial \tau_{n+1,i}} \tilde{\boldsymbol{H}}(\tau_{n+1,i}), \tag{40}$$

$$\tilde{\boldsymbol{H}}''(\tau_{n+1,i}) = \frac{\partial}{\partial \tau_{n+1,i}} \tilde{\boldsymbol{H}}'(\tau_{n+1,i}), \tag{41}$$

and

$$f_{i,j}(n+1) = -\boldsymbol{s}_j^T(n+1)\left(\tilde{\boldsymbol{H}}'(\tau_{n+1,j})\right)^T \tilde{\boldsymbol{H}}'(\tau_{n+1,i})\boldsymbol{s}_i(n+1). \tag{42}$$

References

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Appliations to Tracking and Navigation*. New York: John Wiley & Sons, 2001.

[2] X.-R. Li, Y. Bar-Shalom, and T. Kirubarajan, *Estimation, Tracking and Navigation: Theory, Algorithms and Software*. New York: John Wiley & Sons, June, 2001.

[3] P. Tichavsky, H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1386–1396, May 1998.