

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

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

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

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

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

$$\nabla_{\boldsymbol{\theta}} = \left[\frac{\partial}{\partial \theta_0}, \dots, \frac{\partial}{\partial \theta_{k_0-1}} \right]^T, \quad (3)$$

such that we can express \mathcal{J} as

$$\mathcal{J} = -E \left[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \log p(\mathbf{y}, \boldsymbol{\theta}) \right], \quad (4)$$

$$= -E \left[\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \left\{ \log p(\mathbf{y}|\boldsymbol{\theta}) + \log p(\boldsymbol{\theta}) \right\} \right], \quad (5)$$

$$= \mathcal{J}_D + \mathcal{J}_P, \quad (6)$$

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

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

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

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

Assume that the parameter vector can be partitioned as follows

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

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

$$\mathcal{J} = \begin{bmatrix} \mathcal{J}_{aa} & \mathcal{J}_{ab} \\ \mathcal{J}_{ba} & \mathcal{J}_{bb} \end{bmatrix}, \quad (10)$$

where $\mathcal{J}_{ba} = \mathcal{J}_{ab}^T$. The inverses of the submatrices in \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 \mathcal{J}^{-1} as follows

$$P_b = E \left\{ \left[\hat{\boldsymbol{\theta}}_b(\mathbf{y}) - \boldsymbol{\theta}_b \right] \left[\hat{\boldsymbol{\theta}}_b(\mathbf{y}) - \boldsymbol{\theta}_b \right]^T \right\}, \quad (11)$$

$$\geq [\mathcal{J}_{bb} - \mathcal{J}_{ba} \mathcal{J}_{aa}^{-1} \mathcal{J}_{ab}]^{-1}, \quad (12)$$

$$= \mathcal{J}^{-1}(\boldsymbol{\theta}_b), \quad (13)$$

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

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

Given that the state-space model as follows

$$\boldsymbol{\tau}(n) = \boldsymbol{\tau}(n-1) + \sigma_v \mathbf{v}(n), \quad (15)$$

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} \tilde{\mathbf{H}}_l(\boldsymbol{\tau}(n)) \mathbf{a}(n-l) + \sigma_w \mathbf{w}(n), \quad (16)$$

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

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

where $p(\boldsymbol{\tau}_0)$ is assumed known. According to (4), we can derive an $nk_0 \times nk_0$ information matrix $\mathcal{J}(\mathcal{T}_n)$ from $p(\mathbf{Y}_n, \mathcal{T}_n)$. However, instead of computing the information matrix $\mathcal{J}(\mathcal{T}_n)$, 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 \mathcal{T}_n be partitioned as $[\mathcal{T}_{n-1}^T, \boldsymbol{\tau}_n^T]^T$. Following (9)-(10), we can express $\mathcal{J}(\mathcal{T}_n) \in \mathcal{R}^{nk_0 \times nk_0}$ as

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

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

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

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

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

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

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

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

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

$$\mathbf{A}_{n+1} = \begin{bmatrix} \mathbf{A}_n & \mathbf{B}_n \\ \mathbf{B}_n^T & \mathbf{C}_n + \mathbf{D}_n^{11} \end{bmatrix}, \quad (22)$$

$$\mathbf{B}_{n+1} = \begin{bmatrix} \mathbf{0} \\ \mathbf{D}_n^{12} \end{bmatrix}, \quad (23)$$

$$\mathbf{C}_{n+1} = \mathbf{D}_n^{22}, \quad (24)$$

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

follows

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

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

$$\mathbf{D}_n^{21} = [\mathbf{D}_n^{12}]^T, \quad (27)$$

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

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

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

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

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

II. DERIVATION OF \mathbf{D}_n^{11} , \mathbf{D}_n^{12} , \mathbf{D}_n^{21} , AND \mathbf{D}_n^{22}

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

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

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

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 \mathbf{D}_n^{11} , \mathbf{D}_n^{12} , \mathbf{D}_n^{21} , and \mathbf{D}_n^{22} in sequel.

A. Derivation of \mathbf{D}_n^{11}

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

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

whose i, j th element of \mathbf{D}_n^{11} is defined as follows

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

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

$$\mathbf{D}_n^{11} = \frac{1}{\sigma_v^2} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{bmatrix}. \quad (35)$$

B. Derivation of \mathbf{D}_n^{12}

The i, j th element of \mathbf{D}_n^{12} is defined as follows

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

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

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

C. Derivation of \mathbf{D}_n^{22}

The i, j th element of \mathbf{D}_n^{22} is defined as follows

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

where

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

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

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

and

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

REFERENCES

- [1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications 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.