Derivation of the importance sampling function 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 importance sampling function

A. Derivation of the Gradient Vectors

The gradient vector

$$\nabla\mathcal{L}(\boldsymbol{\tau}_t) = \nabla\mathcal{L}_z(\boldsymbol{\tau}_t) + \nabla\mathcal{L}_\tau(\boldsymbol{\tau}_t), \tag{1}$$

where

$$\nabla\mathcal{L}_z(\boldsymbol{\tau}_t) = \frac{\partial}{\partial\boldsymbol{\tau}_t}\log\left(p(\boldsymbol{z}_t|\boldsymbol{\alpha}_t^{(i)})\right), \tag{2}$$

$$\nabla\mathcal{L}_\tau(\boldsymbol{\tau}_t) = \frac{\partial}{\partial\boldsymbol{\tau}_t}\log\left(p(\boldsymbol{\tau}_t^{i)}|\boldsymbol{\tau}_{t-1}^{(i)},\boldsymbol{z}_t)\right). \tag{3}$$

We first present the derivation of $\nabla\mathcal{L}_z(\boldsymbol{\tau}_t)$ and then that of $\nabla\mathcal{L}_\tau(\boldsymbol{\tau}_t)$. Some details in the derivation can be found in [1].

A.1 Derivation of $\nabla\mathcal{L}_z(\boldsymbol{\tau}_t)$

$$\nabla\mathcal{L}_z(\boldsymbol{\tau}_t) = \left[\nabla\mathcal{L}_z(\tau_{t,0}),\nabla\mathcal{L}_z(\tau_{t,1}),...,\nabla\mathcal{L}_z(\tau_{t,k_t-1})\right]^T, \tag{4}$$

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

where

$$\begin{aligned}
\nabla\mathcal{L}_z(\tau_{t,k}) &= \frac{\partial}{\partial \tau_{t,k}} \log\left(p(\boldsymbol{z}_t|\boldsymbol{\alpha}_t)\right), \\
&= \frac{\partial}{\partial \tau_{t,k}} \left\{ \frac{-1}{2\sigma_w^2} \left(\boldsymbol{z}_t - \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_t)\boldsymbol{a}_t\right)^T \left(\boldsymbol{z}_t - \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_t)\boldsymbol{a}_t\right) + \kappa_{\sigma_w^2} \right\}, \\
&= \frac{1}{\sigma_w^2} \sum_{l=0}^{L-1} \left(\boldsymbol{z}_t - \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_t)\boldsymbol{a}_t\right)^T \tilde{\boldsymbol{H}}'_l(\tau_{t,k})\boldsymbol{s}_k(t-l), \\
&= \frac{1}{\sigma_w^2} \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}'(\tau_{t,k})\boldsymbol{s}_k(t), \qquad (5)
\end{aligned}$$

where $\boldsymbol{\varepsilon}_t$ is given by

$$\begin{aligned}
\boldsymbol{\varepsilon}_t &= \boldsymbol{z}_t - \tilde{\boldsymbol{H}}_0(\boldsymbol{\tau}_t)\boldsymbol{a}_t, \\
&= \boldsymbol{y}_t - \sum_{l=0}^{L-1} \tilde{\boldsymbol{H}}_l(\boldsymbol{\tau}_t)\boldsymbol{a}_{t-l},
\end{aligned}$$

where $\kappa_{\sigma_w^2}$ is a function of the noise variance σ_w^2, and for $k = 0, ..., k_t - 1$

$$\tilde{\boldsymbol{H}}'_l(\tau_{t,k}) \triangleq \frac{\partial \tilde{\boldsymbol{H}}_l(\tau_{t,k})}{\partial \tau_{t,k}}, \qquad (6)$$

$$\begin{aligned}
\tilde{\boldsymbol{H}}'(\tau_{t,k}) &= \frac{\partial}{\partial \tau_{k,t}} \tilde{\boldsymbol{H}}(\tau_{t,k}), \\
&= \left[\tilde{\boldsymbol{H}}'_0(\tau_{t,k}), \tilde{\boldsymbol{H}}'_1(\tau_{t,k}), \ldots, \tilde{\boldsymbol{H}}'_{L-1}(\tau_{t,k})\right], \qquad (7)
\end{aligned}$$

and $\boldsymbol{s}_k(t)$ is the signal amplitude for the kth source, defined as

$$\boldsymbol{s}_k(t) \quad = \quad [s_k(t), s_k(t-1), \ldots, s_k(t-L+1)]^T. \qquad (8)$$

A.2 Derivation of $\nabla\mathcal{L}_\tau(\boldsymbol{\tau}_t)$

$$\nabla\mathcal{L}_\tau(\boldsymbol{\tau}_t) = \left[\nabla\mathcal{L}_\tau(\tau_{t,0}), \nabla\mathcal{L}_\tau(\tau_{t,1}), ..., \nabla\mathcal{L}_\tau(\tau_{t,k_t-1})\right]^T, \qquad (9)$$

where

$$\begin{aligned}
\nabla\mathcal{L}_\tau(\tau_{t,k}) &= \frac{\partial}{\partial \tau_{t,k}} \log\left(p(\boldsymbol{\tau}_t|\boldsymbol{\tau}_{t-1})\right), \\
&= \frac{\partial}{\partial \tau_{t,k}} \left\{ \frac{-1}{2\sigma_v^2} \left(\boldsymbol{\tau}_t - \boldsymbol{\tau}_{\boldsymbol{t-1}}\right)^T \left(\boldsymbol{\tau}_t - \boldsymbol{\tau}_{\boldsymbol{t-1}}\right) + \kappa_{\sigma_v^2} \right\}, \\
&= \frac{-1}{\sigma_v^2}(\tau_{t,k} - \tau_{t-1,k}), \qquad (10)
\end{aligned}$$

where $\kappa_{\sigma_w^2}$ is a function of the noise variance σ_v^2.

As a result, the kth element of the gradient vector $\nabla\mathcal{L}(\boldsymbol{\tau}_t)$ can be expressed as

$$\begin{aligned}
[\nabla\mathcal{L}(\boldsymbol{\tau}_t)]_k &= \nabla\mathcal{L}_z(\tau_{t,k}) + \nabla\mathcal{L}_\tau(\tau_{t,k}), \\
&= \frac{1}{\sigma_w^2} \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}'(\tau_{t,k})\boldsymbol{s}_k(t) - \frac{1}{\sigma_v^2}(\tau_{t,k} - \tau_{t-1,k}). \qquad (11)
\end{aligned}$$

B. Derivation of the Hessian Matrices

The gradient vector

$$\nabla^2 \mathcal{L}(\boldsymbol{\tau}_t) = \nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t) + \nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t), \tag{12}$$

where the k, pth elements of $\nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t)$ and $\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t)$ are given by

$$\left[\nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t)\right]_{k,p} = \frac{\partial^2}{\partial \tau_{t,k} \partial \tau_{t,p}} \log\left(p(\boldsymbol{z}_t | \boldsymbol{\alpha}_t^{(i)})\right), \quad k, p = 0, ..., k_t - 1 \tag{13}$$

$$\left[\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t)\right]_{k,p} = \frac{\partial^2}{\partial \tau_{t,k} \partial \tau_{t,p}} \log\left(p(\boldsymbol{\tau}_t^{i)} | \boldsymbol{\tau}_{t-1}^{(i)}, \boldsymbol{z}_t)\right), \quad k, p = 0, ..., k_t - 1. \tag{14}$$

B.1 Derivation of $\left[\nabla^2 \mathcal{L}_z(\boldsymbol{\tau_t})\right]_{k,p}$

1. If $k \neq p$

$$\begin{aligned}
\left[\nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t)\right]_{k,p} &= \frac{\partial}{\partial \tau_{t,p}} \left\{ \frac{\partial}{\partial \tau_{t,k}} \log\left(p(\boldsymbol{z}_t | \boldsymbol{\alpha}_t^{(i)})\right) \right\}, \\
&= \frac{\partial}{\partial \tau_{t,p}} \nabla \mathcal{L}_z(\tau_{t,k}), \\
&= \frac{\partial}{\partial \tau_{t,p}} \left\{ \frac{1}{\sigma_w^2} \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t) \right\}, \\
&= \frac{-1}{\sigma_w^2} \boldsymbol{s}_p^T(t) \left(\tilde{\boldsymbol{H}}'(\tau_{t,p})\right)^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t),
\end{aligned} \tag{15}$$

2. If $k = p$

$$\begin{aligned}
\left[\nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t)\right]_{k,k} &= \frac{\partial^2}{\partial \tau_{t,k}^2} \log\left(p(\boldsymbol{z}_t | \boldsymbol{\alpha}_t^{(i)})\right), \\
&= \frac{\partial}{\partial \tau_{t,k}} \nabla \mathcal{L}_z(\tau_{t,k}), \\
&= \frac{\partial}{\partial \tau_{t,k}} \left\{ \frac{1}{\sigma_w^2} \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t) \right\}, \\
&= \frac{1}{\sigma_w^2} \left\{ \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}''(\tau_{t,k}) \boldsymbol{s}_k(t) - \boldsymbol{s}_p^T(t) \left(\tilde{\boldsymbol{H}}'(\tau_{t,p})\right)^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t) \right\},
\end{aligned} \tag{16}$$

where

$$\tilde{\boldsymbol{H}}''(\tau_{t,k}) = \frac{\partial}{\partial \tau_{t,k}} \tilde{\boldsymbol{H}}'(\tau_{t,k}). \tag{17}$$

B.2 Derivation of $\left[\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau_t})\right]_{k,p}$

1. If $k \neq p$

$$
\begin{aligned}
\left[\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t)\right]_{k,p} &= \frac{\partial}{\partial \tau_{t,p}} \left\{ \frac{\partial}{\partial \tau_{t,k}} \log\left(p(\boldsymbol{\tau}_t | \boldsymbol{\tau}_{t-1})\right) \right\}, \\
&= \frac{\partial}{\partial \tau_{t,p}} \nabla \mathcal{L}_\tau(\tau_{t,k}), \\
&= \frac{\partial}{\partial \tau_{t,p}} \left\{ \frac{1}{\sigma_v^2} (\tau_{t,k} - \tau_{t-1,k}) \right\} \\
&= 0. \qquad (18)
\end{aligned}
$$

2. If $k = p$

$$
\begin{aligned}
\left[\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t)\right]_{k,k} &= \frac{\partial^2}{\partial \tau_{t,k}^2} \log\left(p(\boldsymbol{\tau}_t | \boldsymbol{\tau}_{t-1})\right), \qquad (19) \\
&= \frac{\partial}{\partial \tau_{t,k}} \nabla \mathcal{L}_\tau(\tau_{t,k}), \qquad (20) \\
&= \frac{\partial}{\partial \tau_{t,k}} \left\{ \frac{1}{\sigma_v^2} (\tau_{t,k} - \tau_{t-1,k}) \right\}, \qquad (21) \\
&= \frac{1}{\sigma_v^2}. \qquad (22)
\end{aligned}
$$

Therefore, the k,pth element of the Hessian matrix $\nabla^2 \mathcal{L}(\boldsymbol{\tau}_t)$ can be expressed as follows

$$
\begin{aligned}
\left[\nabla^2 \mathcal{L}(\boldsymbol{\tau}_t)\right]_{k,p} &= \left[\nabla^2 \mathcal{L}_z(\boldsymbol{\tau}_t)\right]_{k,p} + \left[\nabla^2 \mathcal{L}_\tau(\boldsymbol{\tau}_t)\right]_{k,p}, \qquad (23) \\
&= \begin{cases} -\sigma_w^{-2} \boldsymbol{s}_k^T(t) \left(\tilde{\boldsymbol{H}}'(\tau_{t,k})\right)^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t), & \text{if } k = p, \\ \sigma_w^{-2} \left\{ \boldsymbol{\varepsilon}_t^T \tilde{\boldsymbol{H}}''(\tau_{t,k}) \boldsymbol{s}_k(t) - \boldsymbol{s}_p^T(t) \left(\tilde{\boldsymbol{H}}'(\tau_{t,p})\right)^T \tilde{\boldsymbol{H}}'(\tau_{t,k}) \boldsymbol{s}_k(t) \right\} + \sigma_v^{-2}, & \text{if } k \neq p. \end{cases} \qquad (24)
\end{aligned}
$$

References

[1] W. Ng, J. P. Reilly, and T. Kirubarajan, "The derivation of the theoretical CRLB for wide-band array signal processing: Static scenario," May 2002. Please download the document at http://www.ece.mcmaster.ca/∼reilly.