*Progress In Electromagnetics Research, PIER 102, 31–48, 2010*

# ACCURATE PARAMETER ESTIMATION FOR WAVE EQUATION

**F. K. W. Chan, H. C. So, S.-C. Chan, W. H. Lau and C. F. Chan**

Department of Electronic Engineering
City University of Hong Kong
83 Tat Chee Avenue, Kowloon Tong, Hong Kong, China

**Abstract**—Waves arise in many physical phenomena which have applications such as describing the voltage along a transmission line and medical imaging modality of elastography. In this paper, estimating the parameters for two forms of lossy wave equations, which correspond to multi-mode and multi-dimensional waves, are tackled. By exploiting the linear prediction property of the noise-free signals, an iterative quadratic maximum likelihood (IQML) approach is devised for accurate parameter estimation. Simulation results show that the estimation performance of the proposed IQML algorithms can attain the optimal benchmark, namely, Cramér-Rao lower bound, at sufficiently high signal-to-noise ratio and/or large data size conditions.

## 1. INTRODUCTION

Consider the electromagnetic wave equation for a homogeneous and isotropic medium at a fixed frequency:

$$\bigtriangledown^2 U(x,y,z) + k^2 U(x,y,z) = 0 \tag{1}$$

where $U(x,y,z)$ is the complex field amplitude, $k$ is the propagation constant, $\bigtriangledown^2$ is the Laplacian operator of the form

$$\bigtriangledown^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2}$$

The solution can be expressed as a linear combination of plane waves, that is [1]

$$\begin{aligned} U(x,y,z) \;=\; \sum_{m=1}^{\infty} &[A_m \exp(j(k_{x,m}x + k_{y,m}y + k_{z,m}z)) \\ &+ B_m \exp(-j(k_{x,m}x + k_{y,m}y + k_{z,m}z))] \end{aligned} \tag{3}$$

Corresponding author: F. K. W. Chan (ckwf@hkexperts.com).

where $[k_{x,m} \;\; k_{y,m} \;\; k_{z,m}]^T$ is the wave number vector of the $m$th wave and $A_m$ and $B_m$ are the associated forward and backward complex amplitudes, respectively.

When the wave equation is applied to a channel waveguide, the propagation direction is confined to one-dimension (1-D), say, $z$, the guided field measured along the propagation direction $z$ can be described by reducing (3) to

$$U(z) = \sum_{m=1}^{M} [A_m \exp(jk_{z,m}z) + B_m \exp(-jk_{z,m}z)] \tag{4}$$

where $M$ is the number of nondegenerate modes supported. Examples of channel waveguide include optical fibers, ridge waveguides, and coaxial cables. As (4) is a realistic model of the physical phenomenon in many scenarios [2] such as describing the voltage along a transmission line and medical imaging modality of elastography, it is of interest to find the complex-valued parameters, namely, $\{A_m\}$, $\{B_m\}$ and $\{k_{z,m}\}$.

To the best of our knowledge, Oliphant [2] was the first to tackle this parameter estimation problem with the use of a nonlinear least squares (NLS) algorithm. Although the NLS estimator can attain the maximum likelihood (ML) performance under white Gaussian noise environment, it is hard to implement in practice as its objective function is multi-modal. Recently, So *et al.* [3] have devised an iterative quadratic maximum likelihood (IQML) [4–6] algorithm for the 1-D wave equation in white Gaussian noise by making use of the linear prediction (LP) property [7, 8] in $U(z)$. Apart from more computationally attractive, it is demonstrated in [3] that the IQML scheme can provide optimum accuracy even with a smaller threshold signal-to-noise ratio (SNR) than that of the NLS estimator. However, both [2, 3] only address the simplest scenario of $M = 1$ in (4) which corresponds to the 1-D single-mode waveguide. For $M > 1$, the waveguide is multi-mode and a generalized treatment is required, which will be discussed in detail in this work. Moreover, when the wave equation is applied to planar waveguides or free-space, the propagation direction is no longer confined to 1-D. That is, the field is expressed by (3) without further simplification and its parameter estimation will also be investigated.

The rest of the paper is organized as follows. In Section 2, we develop an IQML-based parameter estimator for the wave equation of (4) in the presence of white Gaussian noise, which can be considered as an extension of [3]. The basic idea is to solve for $\{k_{z,m}\}$ in an iterative manner first and then estimate $\{A_m\}$ and $\{B_m\}$ according to linear least squares (LLS). In Section 3, we generalize (3) to the $P$-dimensional ($P$-D) model with $P \geq 2$. Prior to employing the

IQML approach, we exploit the subspace technique [9] to use the principal singular vectors as the inputs to the algorithm. Cramér-Rao lower bound (CRLB) [10] computation for the investigated signal models, which gives a lower bound on the variance attainable by any unbiased estimator using the same input data, is provided in Section 4. Simulation results are included in Section 5 to evaluate the estimation performance of the proposed approach by comparing with the CRLB. Finally, conclusions are drawn in Section 6. A list of mathematical symbols that are used in the paper is given in Table 1.

**Table 1.** List of symbols.

| Symbol | Meaning |
|---|---|
| $\mathbb{C}^{M \times N}$ | set of $M \times N$ complex matrices |
| $^T$ | transpose |
| $^H$ | conjugate transpose |
| $^*$ | complex conjugate |
| $^{-1}$ | inverse |
| vec | vectorization |
| $\otimes$ | Kronecker product |
| $E$ | expectation |
| $\mathbf{I}_i$ | $i \times i$ identity matrix |
| $\mathbf{0}_{i \times j}$ | $i \times j$ zero matrix |
| $\Re(a)$ | real part of $a$ |
| $\Im(a)$ | imaginary part of $a$ |
| $\hat{a}$ | estimate of $a$ |
| $[\mathbf{a}]_i$ | $i$th element of $\mathbf{a}$ |
| $[\mathbf{A}]_{i,j}$ | $(i,j)$ entry of $\mathbf{A}$ |
| $\mathbf{A}'$ | derivative of $\mathbf{A}$ |
| diag $(\cdot)$ | diagonal matrix |
| Toeplitz $(\mathbf{a}, \mathbf{b}^T)$ | Toeplitz matrix with first column $\mathbf{a}$ and first row $\mathbf{b}^T$ |

## 2. ESTIMATION FOR MULTI-MODE SCENARIOS

In this Section, we develop the parameter estimation algorithm for lossy wave equation corresponding to 1-D multi-mode waveguide. Based on uniformly sampling a noisy version of (4), the observed signal model is

$$R(n) = U(n) + Q(n) \tag{5}$$

where

$$U(n) = \sum_{m=1}^{M} \left[A_m \exp(jk_m n) + B_m \exp(-jk_m n)\right], \quad n = 1, 2, \ldots, N \quad (6)$$

The $k_m$ represents the wave number where we drop the subscript $_z$ for notation simplicity, while $A_m$ and $B_m$ are the associated amplitudes determined by the auxiliary or boundary conditions, of the $m$th wave. All $k_m$, $A_m$ and $B_m$, $m = 1, 2, \ldots, M$, are unknown complex-valued constants to be estimated, while the number of modes, $M$, is assumed known. The additive noises $\{Q(n)\}$ are complex uncorrelated white Gaussian processes with variances $\sigma^2$. The task is to find all unknown parameters given the $N$ measurements of $\{R(n)\}$. Our methodology is first to find the nonlinear parameters of wave numbers. The linear $\{A_m\}$ and $\{B_m\}$ are then estimated using an LLS [10] fit.

We notice that each mode of $U(n)$ can be expressed as

$$\begin{aligned} & A_m \exp(jk_m n) + B_m \exp(-jk_m n) \\ & = 0.5(A_m + B_m + A_m - B_m) \exp(jk_m n) \\ & +0.5(B_m + A_m - A_m + B_m) \exp(-jk_m n) \\ & = (A_m + B_m) \cos(k_m n) + j(A_m - B_m) \sin(k_m n) = \alpha_m \cos(k_m n + \phi_m) \end{aligned} \quad (7)$$

where

$$\alpha_m = 2\sqrt{A_m B_m} \quad (8)$$

$$\phi_m = \tan^{-1}\left(\frac{j(B_m - A_m)}{A_m + B_m}\right) \quad (9)$$

From (7), it is clear that $U(n)$ is a sum of $M$ sinusoids with complex-valued frequencies. Making use of the sinusoidal LP relationship that $U(n)$ can be expressed as a linear combination of its past $2M$ samples, we have [8]:

$$\sum_{n=0}^{2M} \lambda_n U(n - 2M) = 0, \ \lambda_n = \lambda_{2M-n}, \ \lambda_0 = 1, \ n = 2M + 1, \ldots, N \quad (10)$$

where $\{\lambda_n\}$ are called the LP coefficients. The wave numbers $\{k_m\}$ are related to the following polynomial constructed from $\{\lambda_n\}$:

$$\sum_{n=0}^{2M} \lambda_n z^{2M-n} = 0 \quad (11)$$

whose roots are $z = \exp(\pm jk_m)$, $m = 1, 2, \ldots, M$ [8]. With the use of $\lambda_n = \lambda_{2M-n}$, (10) is expressed in matrix form as

$$\mathbf{u} = \mathbf{U}\boldsymbol{\lambda} \quad (12)$$

where

$$\mathbf{u} = -\boldsymbol{\mu}_1 - \boldsymbol{\mu}_{2M+1} \tag{13}$$

$$\mathbf{U} = [\boldsymbol{\mu}_2 + \boldsymbol{\mu}_{2M} \quad \boldsymbol{\mu}_3 + \boldsymbol{\mu}_{2M-1} \quad \dots \quad \boldsymbol{\mu}_{M+2} + \boldsymbol{\mu}_M \quad \boldsymbol{\mu}_{M+1}] \tag{14}$$

$$\boldsymbol{\lambda} = [\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_M]^T \tag{15}$$

$$\boldsymbol{\mu}_n = [U(n) \quad U(n+1) \quad \dots \quad U(N-2M-1+n)]^T \tag{16}$$

As $\{U(n)\}$ are not available, we have to perform estimation from $\{R(n)\}$. Let $\mathbf{R}$ and $\mathbf{r}$ be the noisy versions of $\mathbf{U}$ and $\mathbf{u}$ constructed from $\{R(n)\}$. According to (12), we now have:

$$\mathbf{r} \approx \mathbf{R}\boldsymbol{\lambda} \tag{17}$$

Employing the weighted least squares (WLS) technique, the estimate of $\boldsymbol{\lambda}$, denoted by $\hat{\boldsymbol{\lambda}}$, is [10]

$$\hat{\boldsymbol{\lambda}} = \arg\min_{\breve{\boldsymbol{\lambda}}} \left(\mathbf{R}\breve{\boldsymbol{\lambda}} - \mathbf{r}\right)^H \mathbf{W} \left(\mathbf{R}\breve{\boldsymbol{\lambda}} - \mathbf{r}\right) = \left(\mathbf{R}^H\mathbf{W}\mathbf{R}\right)^{-1} \mathbf{R}^H\mathbf{W}\mathbf{r} \tag{18}$$

where $\mathbf{W}$ is the weighting matrix. To achieve the best accuracy, we employ the Markov estimate [10] of $\mathbf{W}$ and its inverse is derived as:

$$\begin{aligned}\mathbf{W}^{-1} &= E\left\{(\mathbf{R}\boldsymbol{\lambda} - \mathbf{r})(\mathbf{R}\boldsymbol{\lambda} - \mathbf{r})^H\right\} = E\{\boldsymbol{\Lambda}\boldsymbol{r}\boldsymbol{r}^H\boldsymbol{\Lambda}^H\} \\ &= \boldsymbol{\Lambda}E\left\{(\boldsymbol{u}+\boldsymbol{q})(\boldsymbol{u}+\boldsymbol{q})^H\right\}\boldsymbol{\Lambda}^H = \boldsymbol{\Lambda}E\left\{\boldsymbol{q}\boldsymbol{q}^H\right\}\boldsymbol{\Lambda}^H = \sigma^2\boldsymbol{\Lambda}\boldsymbol{\Lambda}^H \end{aligned} \tag{19}$$

where

$$\begin{aligned}\boldsymbol{\Lambda} \;=\; \mathrm{Toeplitz}\Big(&\left[1 \quad \mathbf{0}_{1\times(N-2M-1)}\right]^T, \\ &\left[1 \quad \lambda_1 \quad \dots \quad \lambda_M \quad \lambda_{M-1} \quad \dots \quad \lambda_1 \quad 1 \quad \mathbf{0}_{1\times(N-2M-1)}\right]\Big)\end{aligned} \tag{20}$$

$$\boldsymbol{r} = [R(1) \quad R(2) \quad \dots \quad R(N)]^T \tag{21}$$

$$\boldsymbol{u} = [U(1) \quad U(2) \quad \dots \quad U(N)] \tag{22}$$

$$\boldsymbol{q} = [Q(1) \quad Q(2) \quad \dots \quad Q(N)]^T \tag{23}$$

Note that $\boldsymbol{\Lambda}\boldsymbol{u} = \mathbf{0}_{(N-2M-1)\times 1}$ and it is unnecessary to know $\sigma^2$ as it will be canceled out in (18), that is, we only need (19) up to a multiplying scalar. As $\mathbf{W}$ is parameterized by the unknown parameter $\boldsymbol{\lambda}$, we estimate $\boldsymbol{\lambda}$ in an iterative manner and the estimation procedure is summarized as follows.

**(i)** Set $\mathbf{W} = \mathbf{I}_{N-2M}$

**(ii)** Calculate $\hat{\boldsymbol{\lambda}}$ using (18)

**(iii)** Compute an updated version of $\mathbf{W}$ using (19) with $\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}$

**(iv)** Repeat Steps (ii)–(iii) until a stopping criterion is reached

**(v)** Solve the roots of (11) with $\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}$ to obtain the estimate of $k_m$, namely, $\hat{k}_m$, $m = 1, 2, \ldots, M$

Upon convergence of $\hat{\boldsymbol{\lambda}}$, the bias of (18) can be investigated by premultiplying both sides by $\mathbf{R}^H\mathbf{W}\mathbf{R}$ to construct:

$$f(\hat{\boldsymbol{\lambda}}) = \mathbf{R}^H\mathbf{W}(\mathbf{R}\hat{\boldsymbol{\lambda}} - \mathbf{r}) = \mathbf{0}_{(N-2M)\times 1} \tag{24}$$

where $\mathbf{W}$ is a function of $\hat{\boldsymbol{\lambda}}$. For sufficiently large SNR and/or data size, $\hat{\boldsymbol{\lambda}}$ will be located at a reasonable proximity of $\boldsymbol{\lambda}$. Denoting $\boldsymbol{\Delta\lambda} = \hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}$ and using Taylor's series to expand $f(\hat{\boldsymbol{\lambda}})$ around $\boldsymbol{\lambda}$ yields:

$$\begin{aligned}\mathbf{0}_{(N-2M)\times 1} &\approx f(\boldsymbol{\lambda}) + \left(\mathbf{R}^H\mathbf{W}'\left(\mathbf{I}_M \otimes (\mathbf{R}\boldsymbol{\lambda} - \mathbf{r})\right) + \mathbf{R}^H\mathbf{W}\mathbf{R}\right)\boldsymbol{\Delta\lambda} \\ &\quad - (\mathbf{U}+\mathbf{Q})^H\mathbf{W}\left((\mathbf{U}+\mathbf{Q})\boldsymbol{\lambda} - (\mathbf{u}+\mathbf{q})\right) \\ &\approx \left((\mathbf{U}+\mathbf{Q})^H\mathbf{W}'(\mathbf{I}_M \otimes ((\mathbf{U}+\mathbf{Q})\boldsymbol{\lambda} - \mathbf{u} - \mathbf{q})) + (\mathbf{U}+\mathbf{Q})^H\mathbf{W}(\mathbf{U}+\mathbf{Q})\right)\boldsymbol{\Delta\lambda}\end{aligned} \tag{25}$$

where $\mathbf{Q}$ and $\mathbf{q}$ are the noise components of $\mathbf{R}$ and $\mathbf{r}$, respectively. With the use of $\mathbf{U}\boldsymbol{\lambda} - \mathbf{u} = \mathbf{0}_{(N-2M)\times 1}$ and ignoring the second-order terms whose values are much smaller than those of the first-order terms, we obtain:

$$\begin{aligned}&-\mathbf{U}^H\mathbf{W}\left(\mathbf{Q}\boldsymbol{\lambda} - \mathbf{q}\right) \approx \left(\mathbf{U}^H\mathbf{W}'\left(\mathbf{I}_M \otimes (\mathbf{U}\boldsymbol{\lambda} - \mathbf{u})\right) + \mathbf{U}^H\mathbf{W}\mathbf{U}\right)\boldsymbol{\Delta\lambda} \\ &\Rightarrow \boldsymbol{\Delta\lambda} = -\left\{\left(\mathbf{U}^H\mathbf{W}'\left(\mathbf{I}_M \otimes (\mathbf{U}\boldsymbol{\lambda} - \mathbf{u})\right) + \mathbf{U}^H\mathbf{W}\mathbf{U}\right)\right\}^{-1}\mathbf{U}^H\mathbf{W}\left(\mathbf{Q}\boldsymbol{\lambda} - \mathbf{q}\right)\end{aligned} \tag{26}$$

where $\mathbf{W}' = [\mathbf{W}'_{\lambda_1}\ \mathbf{W}'_{\lambda_2}\ \ldots\ \mathbf{W}'_{\lambda_m}]$ and $\mathbf{W}'_{\lambda_i} = \frac{\partial\mathbf{W}}{\partial\lambda_i}$. As all elements in $\mathbf{Q}$ and $\mathbf{q}$ are of zero-mean, we have $E\{\boldsymbol{\Delta\lambda}\} \approx \mathbf{0}_{M\times 1}$. That is, $\hat{\boldsymbol{\lambda}}$ is an asymptotically unbiased estimate of $\boldsymbol{\lambda}$.

Employing $\{\hat{k}_m\}$ and (5)–(6), the LLS estimates of $\{A_m\}$ and $\{B_m\}$, denoted by $\{\hat{A}_m\}$ and $\{\hat{B}_m\}$, are [10]:

$$\left[\hat{A}_1\quad \hat{B}_1\quad \ldots\quad \hat{A}_M\quad \hat{B}_M\right] = \boldsymbol{r}^T\mathbf{G}^*\left(\mathbf{G}^T\mathbf{G}^*\right)^{-1} \tag{27}$$

where

$$\mathbf{G} = \begin{bmatrix} \exp(j\hat{k}_1) & \exp(j2\hat{k}_1) & \ldots & \exp(jN\hat{k}_1) \\ \exp(-j\hat{k}_1) & \exp(-2j\hat{k}_1) & \ldots & \exp(-jN\hat{k}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \exp(j\hat{k}_M) & \exp(j2\hat{k}_M) & \ldots & \exp(jN\hat{k}_M) \\ \exp(-j\hat{k}_M) & \exp(-j2\hat{k}_M) & \ldots & \exp(-jN\hat{k}_M) \end{bmatrix}^T \tag{28}$$

which are optimum for independent and identically distributed (IID) $\{Q(n)\}$. Note that when there is numerical instability in computing

(18) or (27), we can use singular value decomposition (SVD) to calculate the matrix inverse or add a small number to the diagonal elements in case of rank deficiency.

## 3. ESTIMATION FOR MULTI-DIMENSIONAL SCENARIOS

In this Section, the multi-dimensional wave equation is addressed. Based on uniformly sampling a generalized version of (3), the signal model is:

$$\begin{aligned} R(n_1,n_2,\ldots,n_P) &= U(n_1,n_2,\ldots,n_P)+Q(n_1,n_2,\ldots,n_P),\\ &\qquad n_p=1,2,\ldots,N_p \end{aligned} \tag{29}$$

where

$$U(n_1,n_2,\ldots,n_P)=A\exp\left(j\sum_{p=1}^{P}k_pn_p\right)+B\exp\left(-j\sum_{p=1}^{P}k_pn_p\right) \tag{30}$$

Here, the single-mode is considered for simplicity. The noise-free signal $U(n_1,n_2,\ldots,n_P)$ corresponds to a $P$-D lossy wave with $P\geq 2$, where $k_p$ represents the $p$th dimension wave number while $A$ and $B$ are unknown complex constants, and we drop the subscript $_m$ for notation simplicity. The size in the $p$th dimension is denoted by $N_p$, $p=1,2\ldots,P$. The $\{Q(n_1,n_2,\ldots,n_P)\}$ are $P$-D additive white Gaussian noises with identical variance of $\sigma^2$ and the powers of all $P$ dimensions are the same. The task is to find $k_1,k_2,\ldots,k_P$, $A$ and $B$, given the $N_1\times N_2\times\ldots\times N_P$ measurements in $\{R(n_1,n_2,\ldots,n_P)\}$. Following (7)–(9), we rewrite $U(n_1,n_2,\ldots,n_P)$ as

$$U(n_1,n_2,\ldots,n_P)=\alpha\cos\left(\sum_{p=1}^{P}k_pn_p+\phi\right) \tag{31}$$

where

$$\alpha=2\sqrt{AB} \tag{32}$$

and

$$\phi=\tan^{-1}\left(\frac{j(B-A)}{B+A}\right) \tag{33}$$

Let

$$\begin{aligned} \tilde{\boldsymbol{u}}_{n_1,\ldots,n_{m-1},n_{m+1},\ldots,n_P} = \Big[&U(n_1,\ldots,n_{m-1},1,n_{m+1},\ldots,n_P)\ \ \ldots\\ &U(n_1,\ldots,n_{m-1},N_m,n_{m+1},\ldots,n_P)\Big]^T \end{aligned} \tag{34}$$

Furthermore, we construct $\mathbf{U}_m \in \mathbb{C}^{N_m \times N_{-m}}$ by stacking $\tilde{\boldsymbol{u}}_{n_1,\ldots,n_{m-1},n_{m+1},\ldots,n_P}$, $n_p = 1,2,\ldots,N_p$, $p = 1,2,\ldots,m-1,m+1,\ldots,P$ where $N_{-m} = \prod_{p=1,p\neq m}^{P} n_p$. From (30), any column of $\mathbf{U}_m$ can be expressed as

$$\tilde{\boldsymbol{u}}_{n_1,\ldots,n_{m-1},n_{m+1},\ldots,n_P} = \boldsymbol{\Psi}_m \boldsymbol{\eta}_m \tag{35}$$

where

$$\boldsymbol{\Psi}_m = \begin{bmatrix} \exp(jk_m) & \exp(j2k_m) & \ldots & \exp(jN_m k_m) \\ \exp(-jk_m) & \exp(-j2k_m) & \ldots & \exp(-jN_m k_m) \end{bmatrix}^T \tag{36}$$

$$\boldsymbol{\eta}_m = \left[ A \exp\left( j \sum_{p=1,p\neq m}^{P} k_p n_p \right) \;\; B \exp\left( -j \sum_{p=1,p\neq m}^{P} k_p n_p \right) \right]^T \tag{37}$$

As a result, $\mathbf{U}_m$ is of rank 2. On the other hand, factorizing $\mathbf{U}_m$ with SVD yields

$$\mathbf{U}_m = \boldsymbol{\mathcal{U}}_{m,s} \boldsymbol{\mathcal{S}}_{m,s} \boldsymbol{\mathcal{V}}_{m,s}^H + \boldsymbol{\mathcal{U}}_{m,n} \boldsymbol{\mathcal{S}}_{m,n} \boldsymbol{\mathcal{V}}_{m,n}^H \tag{38}$$

where $\boldsymbol{\mathcal{S}}_{m,s} = \mathrm{diag}(s_{m,1}, s_{m,2})$ and $\boldsymbol{\mathcal{S}}_{m,n} = \mathrm{diag}(s_{m,3}, s_{m,4}, \ldots, s_{m,N_m})$ with singular values $s_{m,1} > s_{m,2} > s_{m,3} = \ldots = s_{m,N_m} = 0$ while $\boldsymbol{\mathcal{U}}_{m,s} = [\mathfrak{u}_{m,1} \; \mathfrak{u}_{m,2}] \in \mathbb{C}^{N_m \times 2}$ and $\boldsymbol{\mathcal{V}}_{m,s} = [\mathfrak{v}_{m,1} \; \mathfrak{v}_{m,2}] \in \mathbb{C}^{N_{-m} \times 2}$ are orthonormal matrices whose columns are the corresponding left and right singular vectors of the signal subspace, respectively. In addition, $\boldsymbol{\mathcal{U}}_{m,n} = [\mathfrak{u}_{m,3} \; \mathfrak{u}_{m,4} \; \ldots \; \mathfrak{u}_{m,N_m}] \in \mathbb{C}^{N_m \times (N_m - 2)}$ and $\boldsymbol{\mathcal{V}}_{m,s} = [\mathfrak{v}_{m,3} \; \mathfrak{v}_{m,4} \; \ldots \; \mathfrak{v}_{m,N_{-m}}] \in \mathbb{C}^{N_{-m} \times (N_{-m} - 2)}$ are orthonormal matrices whose columns are the corresponding left and right singular vectors of the noise subspace, respectively. As $\boldsymbol{\mathcal{U}}_{m,s}$ and $\boldsymbol{\Psi}_m$ span the same subspace, we have [9]

$$\boldsymbol{\mathcal{U}}_{m,s} = \boldsymbol{\Psi}_m \boldsymbol{\Omega}_m \tag{39}$$

where $\boldsymbol{\Omega}_m \in \mathbb{C}^{2\times 2}$ is unknown. It can be deduced from (39) that $[\mathfrak{u}_{m,i}]_l$ is $U(l)$ in (6) with $M = 1$, $A = [\boldsymbol{\Omega}_m]_{1,i}$ and $B = [\boldsymbol{\Omega}_m]_{2,i}$, $i = 1,2$. Nevertheless, in the presence of noise, $U(n_1, n_2, \ldots, n_P)$ is replaced by $R(n_1, n_2, \ldots, n_P)$ and the equal sign in (39) is replaced by approximately equal sign, leading to

$$\mathscr{U}_m \lambda_m \approx \mathfrak{u}_m \tag{40}$$

where

$$\mathscr{U}_m = [\mathscr{U}_{m,1}^T \quad \mathscr{U}_{m,2}^T]^T \tag{41}$$

$$\mathscr{U}_{m,i} = \left[ [\mathfrak{u}_{m,i}]_2 \quad [\mathfrak{u}_{m,i}]_3 \quad \cdots \quad [\mathfrak{u}_{m,i}]_{N_m - 1} \right]^T \tag{42}$$

$$\mathfrak{u}_m = \left[ \mathfrak{u}_{m,1}^T \quad \mathfrak{u}_{m,2}^T \right]^T \tag{43}$$

$$\mathfrak{u}_{m,i} = -\left[ [\mathfrak{u}_{m,i}]_1 + [\mathfrak{u}_{m,i}]_3 \; [\mathfrak{u}_{m,i}]_2 + [\mathfrak{u}_{m,i}]_4 \cdots [\mathfrak{u}_{m,i}]_{N_m - 2} + [\mathfrak{u}_{m,i}]_{N_m} \right]^T \tag{44}$$

$$\lambda_m = -2\cos(k_m) \tag{45}$$

As a result, the WLS estimate of $\lambda_m$, denoted by $\hat{\lambda}_m$, is

$$\hat{\lambda}_m = \arg\min_{\breve{\lambda}_m} \left(\boldsymbol{\mathscr{U}}_m \breve{\lambda}_m - \mathbb{u}_m\right)^H \mathbf{W}_m \left(\boldsymbol{\mathscr{U}}_m \breve{\lambda}_m - \mathbb{u}_m\right)$$
$$= (\boldsymbol{\mathscr{U}}_m^H \mathbf{W}_m \boldsymbol{\mathscr{U}}_m)^{-1} \boldsymbol{\mathscr{U}}_m^H \mathbf{W}_m \mathbb{u}_m \tag{46}$$

In the Appendix A, we have derived the optimal weighting matrix $\mathbf{W}_m$ as

$$\mathbf{W}_m = \left[\mathrm{diag}\left(s_{m,1}^{-2}, s_{m,2}^{-2}\right) \otimes \boldsymbol{\Lambda}_m \boldsymbol{\Lambda}_m^H\right]^{-1} \tag{47}$$

where

$$\boldsymbol{\Lambda}_m = \mathrm{Toeplitz}\left(\begin{bmatrix}1 & \mathbf{0}_{1\times(N_m-3)}\end{bmatrix}^T, \begin{bmatrix}1 & \lambda_m & 1 & \mathbf{0}_{1\times(N_m-3)}\end{bmatrix}\right) \tag{48}$$

which is parameterized by the unknown parameter $\lambda_m$. As (47) is block diagonal, (46) can be simplified to

$$\hat{\lambda}_m = \frac{\sum_{i=1}^{2} s_{m,i}^2 \boldsymbol{\mathscr{U}}_{m,i}^H \boldsymbol{\mathcal{W}}_m \mathbb{u}_{m,i}}{\sum_{i=1}^{2} s_{m,i}^2 \boldsymbol{\mathscr{U}}_{m,i}^H \boldsymbol{\mathcal{W}}_m \boldsymbol{\mathscr{U}}_{m,i}} \tag{49}$$

where

$$\boldsymbol{\mathcal{W}}_m = \left(\boldsymbol{\Lambda}_m \boldsymbol{\Lambda}_m^H\right)^{-1} \tag{50}$$

Similar to Section 2, we estimate $\lambda_m$, $m = 1, 2, \ldots, P$, in a separable and iterative manner as follows:

**(i)** Set $\boldsymbol{\mathcal{W}}_m = \mathbf{I}_{N_m-2}$

**(ii)** Calculate $\hat{\lambda}_m$ using (49)

**(iii)** Compute an updated version of $\boldsymbol{\mathcal{W}}_m$ using (50) with $\lambda_m = \hat{\lambda}_m$

**(iv)** Repeat Steps (ii)–(iii) until a stopping criterion is reached

**(v)** Use the finalized $\hat{\lambda}_m$ to compute the estimate of $k_m$ as $\hat{k}_m = \cos^{-1}(-\hat{\lambda}_m/2)$

After all $\{\hat{k}_m\}$ are available, the optimal estimates of $A$ and $B$, denoted by $\hat{A}$ and $\hat{B}$, are then computed using LLS as

$$\begin{bmatrix}\hat{A} & \hat{B}\end{bmatrix} = \mathbf{r}^T \boldsymbol{\mathcal{G}}^* \left(\boldsymbol{\mathcal{G}}^T \boldsymbol{\mathcal{G}}^*\right)^{-1} \tag{51}$$

where

$$\boldsymbol{\mathcal{G}} = \begin{bmatrix} \exp(j\mathbf{n}_1^T \hat{\boldsymbol{k}}) & \exp(j\mathbf{n}_2^T \hat{\boldsymbol{k}}) & \ldots & \exp(j\mathbf{n}_{\prod_{p=1}^P N_p}^T \hat{\boldsymbol{k}}) \\ \exp(-j\mathbf{n}_1^T \hat{\boldsymbol{k}}) & \exp(-j\mathbf{n}_2^T \hat{\boldsymbol{k}}) & \ldots & \exp(-j\mathbf{n}_{\prod_{p=1}^P N_p}^T \hat{\boldsymbol{k}}) \end{bmatrix}^T \tag{52}$$

Here, $\hat{\boldsymbol{k}} = [\hat{k}_1\ \hat{k}_2\ \ldots\ \hat{k}_P]^T$, $\mathbf{r} \in \mathbb{C}^{\prod_{p=1}^P N_p}$ denotes the vector containing all elements in $\{R(n_1, n_2, \ldots, n_P)\}$ and $\mathbf{n}_l \in \mathbb{Z}^P$ contains the indices with $[\mathbf{n}_l]_p$ equals $n_p$ in $[\mathbf{r}]_l$.

For the case of a multi-mode multi-dimensional wave, we can combine the previous developments for (6) and (30) to achieve parameter estimation. When the number of modes is $M$, the rank of $\mathbf{U}_m$ in (38) becomes $2M$ and we need to utilize the first $2M$ left singular vectors and singular values for estimating the wave numbers where each column of $\boldsymbol{\mathcal{U}}_{m,s}$ now corresponds to a $M$-mode 1-D wave. It is worthy to point out that the major difficulty lies in pairing the wave numbers in the $P$-dimension [11]. Using $P = 3$ as an illustration, we have to determine the triplets of $(k_{x,m}, k_{y,m}, k_{z,m})$ after independently obtaining $k_{x,m}$, $k_{y,m}$ and $k_{z,m}$, $m = 1, 2, \ldots, M$.

## 4. DERIVATION OF CRAMÉR-RAO LOWER BOUND

In this Section, the CRLBs for parameters in (5)–(6) and (29)–(30) are derived. We first address the 1-D multiple waves. As $Q(n)$, $n = 1, 2, \ldots, N$ are IID, the CRLB, denoted by $\mathbf{H}$, is computed from the inverse of the corresponding Fisher information matrix (FIM):

$$\mathbf{H}^{-1} = \frac{2}{\sigma^2}\Re\left(\mathbf{F}\mathbf{F}^H\right) \tag{53}$$

where the columns of $\mathbf{F}$ contain $\frac{\partial U(n)}{\partial \boldsymbol{\theta}}$ and $\boldsymbol{\theta} = \left[\boldsymbol{\theta}_k^T \quad \boldsymbol{\theta}_A^T \quad \boldsymbol{\theta}_B^T\right]^T$ with $\boldsymbol{\theta}_k = [\Re(k_1) \quad \Im(k_1) \quad \ldots \quad \Re(k_M) \quad \Im(k_M)]^T$, $\boldsymbol{\theta}_A = [\Re(A_1) \quad \Im(A_1) \quad \ldots \quad \Re(A_M) \quad \Im(A_M)]^T$ and $\boldsymbol{\theta}_k = [\Re(B_1) \quad \Im(B_1) \quad \ldots \quad \Re(B_M) \quad \Im(B_M)]^T$. The values of $\frac{\partial U(n)}{\partial \boldsymbol{\theta}}$ are calculated as:

$$\frac{\partial U(n)}{\partial \Re(k_m)} = jnA_m \exp(jk_m n) - jnB_m \exp(-jk_m n) \tag{54}$$

$$\frac{\partial U(n)}{\partial \Im(k_m)} = -nA_i \exp(jk_m n) + nB_i \exp(-jk_m n) \tag{55}$$

$$\frac{\partial U(n)}{\partial \Re(A_m)} = \exp(jk_m n) \tag{56}$$

$$\frac{\partial U(n)}{\partial \Im(A_m)} = j\exp(jk_m n) \tag{57}$$

$$\frac{\partial U(n)}{\partial \Re(B_m)} = \exp(-jk_m n) \tag{58}$$

$$\frac{\partial U(n)}{\partial \Im(B_m)} = j\exp(-jk_m n) \tag{59}$$

The CRLB of $k_m$, $A_m$, $B_m$ are $[\mathbf{H}]_{2m-1,2m-1} + [\mathbf{H}]_{2m,2m}$, $[\mathbf{H}]_{2M+2m-1,2M+2m-1} + [\mathbf{H}]_{2M+2m,2M+2m}$, $[\mathbf{H}]_{3M+2m-1,3M+2m-1} + [\mathbf{H}]_{3M+2m,3M+2m}$, $m = 1, 2, \ldots, M$, respectively.

For the signal model of (29)–(30), the expression of FIM is equal to (53) except that the elements in $\mathbf{F}$ are now modified as $\frac{\partial U(n_1,n_2,\ldots,n_P)}{\partial \boldsymbol{\theta}}$ with $\boldsymbol{\theta} = [\Re(k_1)\ \Im(k_1)\ \ldots\ \Re(k_P)\ \Im(k_P)\ \Re(A)\ \Im(A)\ \Re(B)\ \Im(B)]^T$, $n_p = 1, 2, \ldots, N_p$, $p = 1, 2, \ldots, P$. The values for $\frac{\partial U(n_1,n_2,\ldots,n_P)}{\partial \boldsymbol{\theta}}$ are computed as:

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Re(k_p)} = j n_p A \exp\left(j \sum_{p=1}^{P} k_p n_p\right) - j n_p B \exp\left(-j \sum_{p=1}^{P} k_p n_p\right) \tag{60}$$

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Im(k_p)} = -n_p A \exp\left(j \sum_{p=1}^{P} k_p n_p\right) + n_p B \exp\left(-j \sum_{p=1}^{P} k_p n_p\right) \tag{61}$$

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Re(A_m)} = \exp\left(j \sum_{p=1}^{P} k_p n_p\right) \tag{62}$$

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Im(A_m)} = j \exp\left(j \sum_{p=1}^{P} k_p n_p\right) \tag{63}$$

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Re(B_m)} = \exp\left(-j \sum_{p=1}^{P} k_p n_p\right) \tag{64}$$

$$\frac{\partial U(n_1, n_2, \ldots, n_P)}{\partial \Im(B_m)} = j \exp\left(-j \sum_{p=1}^{P} k_p n_p\right) \tag{65}$$

The CRLB of $k_p$ is $[\mathbf{H}]_{p,p} + [\mathbf{H}]_{p+P,p+P}$, $p = 1, 2, \ldots, P$, and those of $A$ and $B$ are $[\mathbf{H}]_{2P+1,2P+1} + [\mathbf{H}]_{2P+2,2P+2}$ and $[\mathbf{H}]_{2P+3,2P+3} + [\mathbf{H}]_{2P+4,2P+4}$, respectively.

## 5. SIMULATION RESULTS

Computer simulations have been conducted to evaluate the lossy wave parameter estimation performance of the proposed approach for 1-D dual waves and 3-D single wave in zero-mean white complex Gaussian noise, that is $M = 2$ and $P = 3$. The stopping criterion of the proposed methodology is a fixed number of iterations. In both scenarios, we iterate the algorithms for 3 times as no obvious improvement is observed for more iterations. To illustrate the comparative performance, we include the results of the NLS estimators which are realized by the Newton's method with the true values as initial estimates. We scale $Q(n)$ and $Q(n_1, n_2, n_3)$ to produce different SNR conditions where $\mathrm{SNR} = \sum_{n=1}^{N} |U(n)|^2/(N\sigma^2)$ and $\mathrm{SNR} = \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \sum_{n_3=1}^{N_3} |U(n_1, n_2, n_3)|^2/(N_1 N_2 N_3 \sigma^2)$ in the former and latter cases, respectively. All results provided are averages of 1000 independent runs.

In the first test, the mean square error (MSE) performance of wave numbers versus SNR for 1-D multiple lossy waves is examined. The parameters of $U(n)$ are $k_1 = 1 + j0.02$, $k_2 = 2 - j0.01$, $A_1 = 1 + j0.6$, $A_2 = 0.8 + j0.65$, $B_1 = -0.6 + j0.8$ and $B_2 = 0.9 + j0.6$ while a data length of $N = 10$ is assigned. Figures 1 to 6 plot $E\{|k_1 - \hat{k}_1|^2\}$, $E\{|k_2 - \hat{k}_2|^2\}$, $E\{|A_1 - \hat{A}_1|^2\}$, $E\{|B_1 - \hat{B}_1|^2\}$,

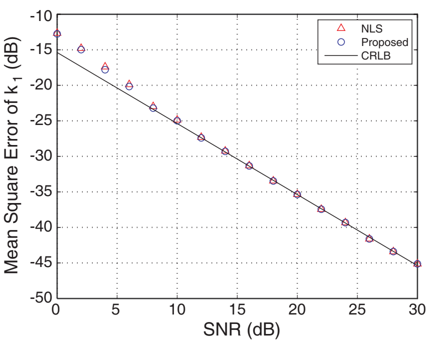


**Figure 1.** Mean square error for $k_1$ versus SNR with 1-D multiple waves.

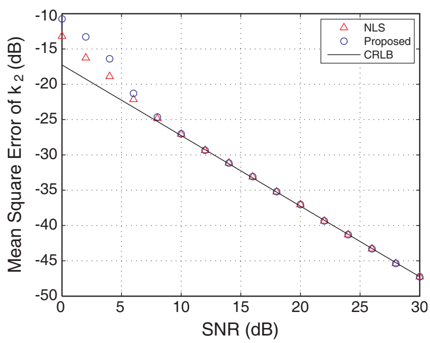


**Figure 2.** Mean square error for $k_2$ versus SNR with 1-D multiple waves.

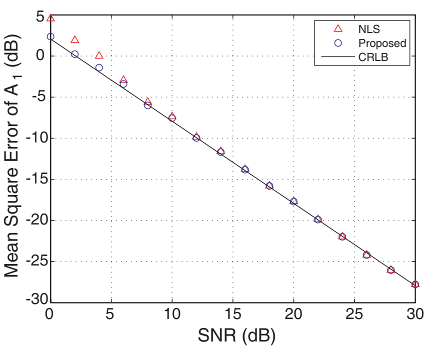


**Figure 3.** Mean square error for $A_1$ versus SNR with 1-D multiple waves.

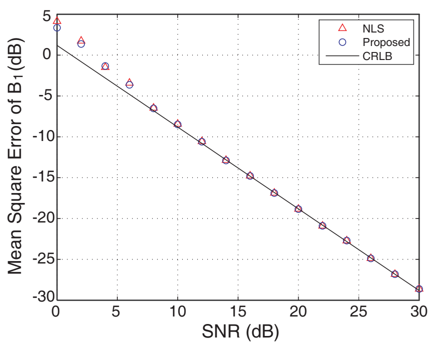


**Figure 4.** Mean square error for $B_1$ versus SNR with 1-D multiple waves.

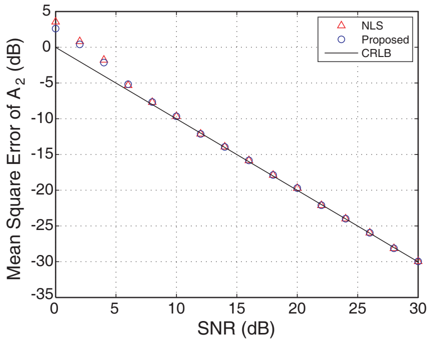


**Figure 5.** Mean square error for $A_2$ versus SNR with 1-D multiple waves.

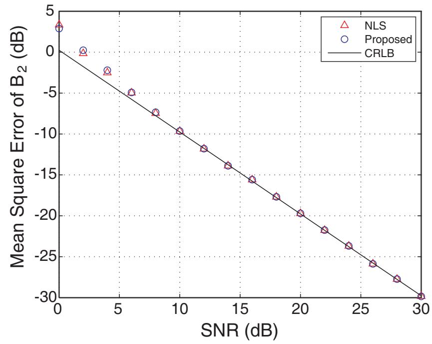


**Figure 6.** Mean square error for $B_2$ versus SNR with 1-D multiple waves.

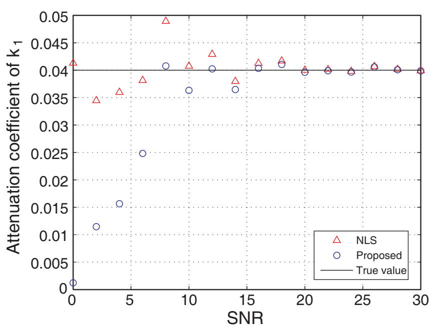


**Figure 7.** Mean attenuation coefficient estimate for $k_1$ versus SNR with 1-D multiple waves.

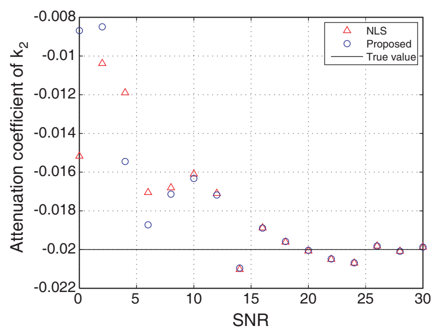


**Figure 8.** Mean attenuation coefficient estimate for $k_2$ versus SNR with 1-D multiple waves.

$E\{|A_2 - \hat{A}_2|^2\}$ and $E\{|B_2 - \hat{B}_2|^2\}$, respectively. It is seen that the accuracy of both approaches attains the CRLB for SNR $\geq 8$ dB. The proposed and NLS methods have almost the same performance except in the estimation of $k_2$ and $A_1$ at SNR $\leq 6$ dB. It is worthy to note that in practice, the NLS method will give poorer performance when its initial estimates are not sufficiently close to the global solution. Regarding computational complexity, the average computation times of the proposed and NLS algorithms for a single trial are measured as $9.23 \times 10^{-4}$ s and $6.83 \times 10^{-2}$ s, respectively, indicating the former is more computationally attractive. The mean attenuation coefficients obtained from $\hat{k}_1$ and $\hat{k}_2$ are plotted in Figures 7 and 8. We see that both algorithms give nearly unbiased estimates when the SNR is sufficiently high and the proposed method exhibits a larger bias in Figure 7 when SNR $\leq$ 6 dB. Comparing Figures 1, 2, 7 and 8, we deduce that the variance in the IQML estimator is smaller than that in the NLS method. This test is repeated with white uniform noise to investigate the algorithm robustness and the MSE results for $k_1$ and $k_2$ are plotted in Figures 9 and 10. It is observed that the performance of the both algorithms in uniform noise is comparable with that of the Gaussian noise case. We have also repeated the first experiment with $N = 30$ to study the effect of larger data length and the MSEs for $k_1$ and $k_2$ are shown in Figures 11 and 12. We see that the performance of both schemes improves and their threshold SNR is reduced to 4 dB. Combining the findings, the proposed estimator is able to achieve optimum performance for sufficiently high SNRs and/or large data lengths.

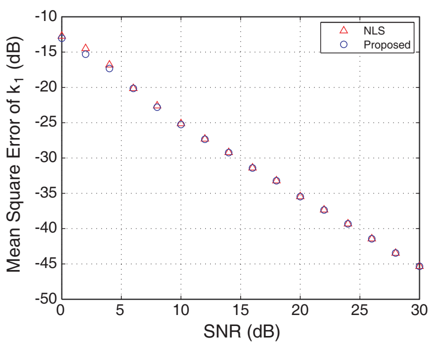


**Figure 9.** Mean square error for $k_1$ versus SNR in uniform noise.

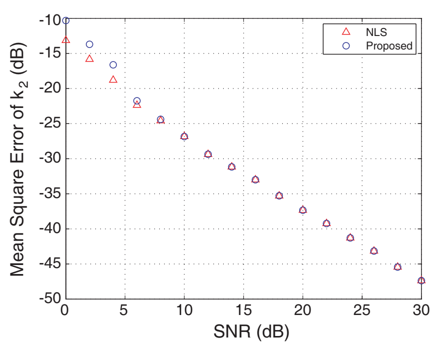


**Figure 10.** Mean square error for $k_2$ versus SNR in uniform noise.

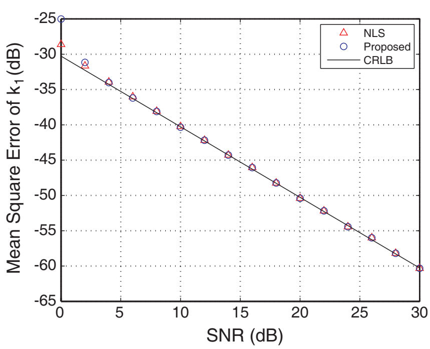


**Figure 11.** Mean square error for $k_1$ versus SNR at $N = 30$.

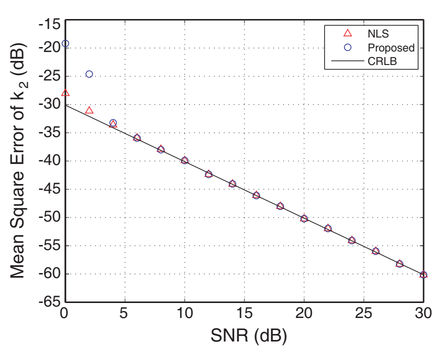


**Figure 12.** Mean square error for $k_2$ versus SNR at $N = 30$.

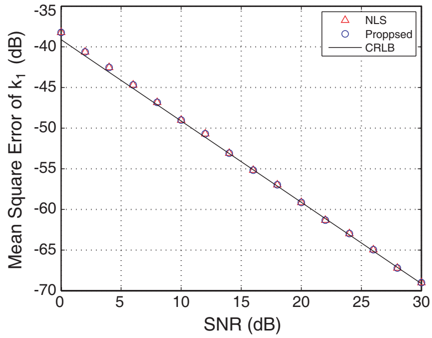


**Figure 13.** Mean square error for $k_1$ versus SNR with 3-D single wave.

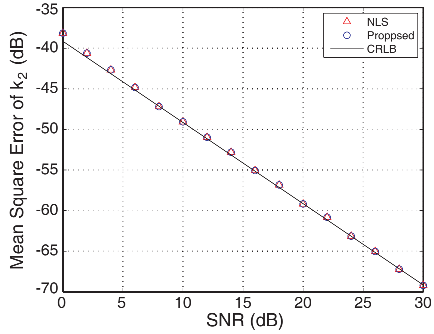


**Figure 14.** Mean square error for $k_2$ versus SNR with 3-D single wave.

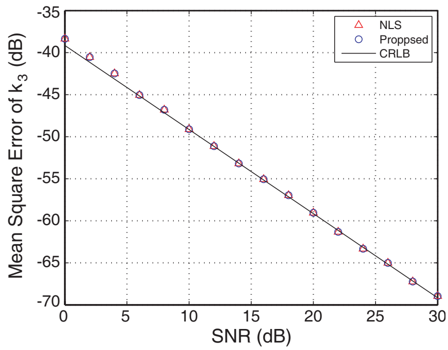


**Figure 15.** Mean square error for $k_3$ versus SNR with 3-D single wave.

In the second test, the wave number estimation performance of the proposed approach for 3-D single-mode is evaluated. Now the parameter settings are $k_1 = 1 + j0.02$, $k_2 = 2 - j0.01$, $k_3 = 1.5 + j0.01$, $A = 1 + j0.6$, $B = -0.6 + j0.8$ and $N_1 = N_2 = N_3 = 10$. Figures 13 to 15 show the MSEs of the wave numbers in the three dimensions. We observe that the wave number estimation performance achieves CRLB for SNR $\geq 4$ dB. Note that the results for $A$ and $B$ are not included as similar findings are obtained.

## 6. CONCLUSION

An iterative parameter estimation approach for two forms of lossy wave equations, which correspond to multi-mode and multi-dimensional waves, has been devised. In the algorithm development, linear prediction and weighted least squares techniques are utilized in both scenarios. Furthermore, we make use of the subspace methodology to achieve parameter estimation from the principal singular vectors of the multi-dimensional data. It is demonstrated that the estimation performance of the proposed algorithms attains Cramér-Rao lower bound for sufficiently high signal-to-noise ratio (SNR) and/or large data size conditions. As an illustration, for a one-dimensional dual-mode wave with an observation length of 10, the threshold SNR is 10 dB.

## ACKNOWLEDGMENT

The work described in this paper was supported by a grant from CityU (Project No. 7002448).

## APPENDIX A.

In this Appendix, the derivation of $\mathbf{W}_m$ in (47) is provided. Let $\tilde{\mathbf{u}}_{m,1,2}$ and $\boldsymbol{\Delta}\mathbf{u}_{m,1,2}$ be the noise-free version of $\mathbf{u}_{m,1,2} = \begin{bmatrix}\mathbf{u}_{m,1}^T & \mathbf{u}_{m,2}^T\end{bmatrix}^T$ and the corresponding perturbation, respectively. Using $(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}_m)\,\tilde{\mathbf{u}}_{m,1,2} = \mathbf{0}_{2N_m-4}$, we first show that the inverse of $\mathbf{W}_m$ is

$$\begin{aligned}\mathbf{W}_m^{-1} =& \mathbb{E}\left\{(\mathcal{U}_m\lambda_m - \mathbf{u}_m)(\mathcal{U}_m\lambda_m - \mathbf{u}_m)^H\right\}\\ =& \mathbb{E}\left\{(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}_m)(\tilde{\mathbf{u}}_{m,1,2}+\boldsymbol{\Delta}\mathbf{u}_{m,1,2})(\tilde{\mathbf{u}}_{m,1,2}+\boldsymbol{\Delta}\mathbf{u}_{m,1,2})^H\left(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}_m^H\right)\right\}\\ =& \mathbb{E}\left\{(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}_m)\,\boldsymbol{\Delta}\mathbf{u}_{m,1,2}\boldsymbol{\Delta}\mathbf{u}_{m,1,2}^H\left(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}_m^H\right)\right\} \qquad \text{(A1)}\end{aligned}$$

As $E\left\{\boldsymbol{\Delta}\mathbf{u}_{m,1,2}\boldsymbol{\Delta}\mathbf{u}_{m,1,2}^H\right\}$ in (A1) involve perturbations of singular vectors, we need to relate it with the noise $Q(n_1, n_2, \ldots, n_P)$. Let $\mathbf{Q}_m$ be the noise counterpart of $\mathbf{U}_m$. The first-order perturbation of $\boldsymbol{\Delta}\mathbf{u}_{m,i}$ is [12]

$$\begin{aligned}\boldsymbol{\Delta}\mathbf{u}_{m,i} =& \boldsymbol{\mathcal{U}}_{m,s}\mathbf{D}_{m,i}\boldsymbol{\mathcal{U}}_{m,s}^H\mathbf{Q}_m\mathbf{v}_{m,i}s_{m,i} + \boldsymbol{\mathcal{U}}_{m,s}\mathbf{D}_{m,i}\boldsymbol{\mathcal{S}}_{m,s}\boldsymbol{\mathcal{V}}_{m,s}^H\mathbf{Q}_m^H\mathbf{u}_{m,i}\\ &+\boldsymbol{\mathcal{U}}_{m,n}\boldsymbol{\mathcal{U}}_{m,n}^H\mathbf{Q}_m\mathbf{v}_{m,i}s_{m,i}^{-1} = \sum_{j=1}^{3} t_{m,i,j} \qquad \text{(A2)}\end{aligned}$$

where

$$t_{m,i,1} = s_{m,i}\left(\mathbf{v}_{m,i}^T \otimes \boldsymbol{\mathcal{U}}_{m,s}\mathbf{D}_{m,i}\boldsymbol{\mathcal{U}}_{m,s}^H\right)\mathrm{vec}(\mathbf{Q}_m) \qquad \text{(A3)}$$

$$t_{m,i,2} = \left(\mathbf{u}_{m,i}^T \otimes \boldsymbol{\mathcal{U}}_{m,s}\mathbf{D}_{m,i}\boldsymbol{\mathcal{S}}_{m,s}\boldsymbol{\mathcal{V}}_{m,s}^H\right)\mathrm{vec}\left(\mathbf{Q}_m^H\right) \qquad \text{(A4)}$$

$$
\begin{aligned}
t_{m,i,3} &= s_{m,i}^{-1}\left(\mathfrak{v}_{m,i}^{T}\otimes\boldsymbol{\mathcal{U}}_{m,n}\boldsymbol{\mathcal{U}}_{m,n}^{H}\right)\operatorname{vec}\left(\mathbf{Q}_m\right)\\
&= s_{m,i}^{-1}\left(\mathfrak{v}_{m,i}^{T}\otimes\mathbf{I}_{N_m}\right)\operatorname{vec}(\mathbf{Q}_m)-s_{m,i}^{-1}\left(\mathfrak{v}_{m,i}^{T}\otimes\boldsymbol{\mathcal{U}}_{m,s}\boldsymbol{\mathcal{U}}_{m,s}^{H}\right)\operatorname{vec}(\mathbf{Q}_m)
\end{aligned}
\tag{A5}
$$

Here, $\mathbf{D}_{m,i}$ is a diagonal matrix with the $(i,i)$ entry being zero and the other diagonal entry equals $1/(s_{m,3-i}^2 - s_{m,i}^2)$. Note that $t_{m,i,j}$, $j = 1, 2, 3$, are obtained by the formula $\operatorname{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\operatorname{vec}(\mathbf{B})$. Using $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD}$ and $\boldsymbol{\Lambda}_m\boldsymbol{\mathcal{U}}_{m,s} = \mathbf{0}_{(N_m-2)\times 2}$, (A1) becomes

$$
\begin{aligned}
&\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m\right)\begin{bmatrix} s_{m,1}^{-1}(\mathfrak{v}_{m,1}^{T}\otimes\mathbf{I}_{N_m})\\ s_{m,2}^{-1}(\mathfrak{v}_{m,2}^{T}\otimes\mathbf{I}_{N_m})\end{bmatrix} E\left(\operatorname{vec}(\mathbf{Q}_m)\operatorname{vec}(\mathbf{Q}_m)^H\right)\\
&\times\left[s_{m,1}^{-1}(\mathfrak{v}_{m,1}^{*}\otimes\mathbf{I}_{N_m})\quad s_{m,2}^{-1}(\mathfrak{v}_{m,2}^{*}\otimes\mathbf{I}_{N_m})\right]\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m^H\right)\\
=\;&\sigma^2\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m\right)\begin{bmatrix} s_{m,1}^{-1}(\mathfrak{v}_{m,1}^{T}\otimes\mathbf{I}_{N_m})\\ s_{m,2}^{-1}(\mathfrak{v}_{m,2}^{T}\otimes\mathbf{I}_{N_m})\end{bmatrix}\\
&\times\left[s_{m,1}^{-1}(\mathfrak{v}_{m,1}^{*}\otimes\mathbf{I}_{N_m})\quad s_{m,2}^{-1}(\mathfrak{v}_{m,2}^{*}\otimes\mathbf{I}_{N_m})\right]\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m^H\right)\\
=\;&\sigma^2\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m\right)\left(\operatorname{diag}\left(s_{m,1}^{-2},s_{m,2}^{-2}\right)\otimes\mathbf{I}_{N_m}\right)\left(\mathbf{I}_2\otimes\boldsymbol{\Lambda}_m^H\right)\\
=\;&\sigma^2\operatorname{diag}\left(s_{m,1}^{-2},s_{m,2}^{-2}\right)\otimes\boldsymbol{\Lambda}_m\boldsymbol{\Lambda}_m^H
\end{aligned}
\tag{A6}
$$

which corresponds to (47).

## REFERENCES


1. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, 1998.
2. Oliphant, T. E., "On parameter estimates for the lossy wave equation," *IEEE Transactions on Signal Processing*, Vol. 56, No. 1, 49–60, Jan. 2008.
3. So, H. C., M. T. Amin, and F. K. W. Chan, "Accurate and simple estimator for lossy wave equation," *IEEE Signal Processing Letters*, Vol. 16, No. 9, 798–801, Sep. 2009.
4. Goodwin, G. C. and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York, 1977.
5. Kumaresan, R., L. L. Scharf, and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Transactions on Acoustic, Speech and Signal Processing*, Vol. 34, 637–640, Jun. 1986.

6. Bresler, Y. and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustic, Speech and Signal Processing*, Vol. 34, 1081–1089, Oct. 1986.
7. Makhoul, J., "Linear prediction: A tutorial review," *Proceedings of the IEEE*, Vol. 63, No. 4, 561–581, 1975.
8. Chan, Y. T., J. M. M. Lavoie, and J. B. Plant, "A parameter estimation approach to estimation of frequencies of sinusoids," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 29, No. 2, 214–219, Apr. 1981.
9. Roy, R. and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 37, No. 7, 984–995, Jul. 1989.
10. Kay, S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.
11. Hua, Y., "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Transactions on Signal Processing*, Vol. 40, No. 9, 2267–2280, Sep. 1992.
12. Liu, J., X. Liu, and X. Ma, "First-order perturbation analysis of singular vectors in singular value decomposition," *IEEE Transactions on Signal Processing*, Vol. 56, No. 7, 3044–3049, Jul. 2008.