Distributed Bayesian vector estimation using noise-optimized low-resolution sensor observations

Jia Liu, Fabing Duan, François Chapeau-Blondeau, Derek Abbott

Abstract

The distributed Bayesian vector parameter estimation problem based on low-resolution observations is investigated in a network, where each node represents an ensemble of estimates from a large number of sensors. A noise-enhanced Bayesian vector estimator that benefits from artificially added noise is proposed. For a network node composed of a sufficiently large number of identical low-resolution sensors, a lemma governing the weight coefficients is proven, and low-cost calculation expressions of the designed estimator and its Bayesian mean square error (MSE) are derived by avoiding costly computations due to high-dimensional matrix inversions. Experimental results show that by intentionally adding an appropriate amount of noise to networks of the low-resolution sensors, the MSE of the designed Bayesian vector estimator can be significantly reduced.

1. Introduction

The widespread use of low-power and low-complexity sensors in wireless networks and digital communication systems offers an answer to the demanding requirements of cost constraints and bandwidth limitations [1–4], but poses the challenge of achieving efficient processing from low-resolution observations of many sensors [5–7]. In recent years, increasing interest has focused on studying efficient parameter estimation based on low-resolution observations. In order to reduce the signal distortion and improve the accuracy of parameter estimation, the operation of adding optimal noise before sensors is often employed in practical applications such as audio coding [8], image compression [9], distributed estimation [2–4,10,11], direction-of-arrival [12], and multiple-input multiple-output communications [13], where it is also commonly referred to as dithering [1,5,7,14]. Dithering has been widely investigated for reducing the mean square error (MSE) of deterministic or random parameter estimation [7,10,13,15], or for minimizing the Cramér-Rao lower bound (CRLB) [3,5,6], by means of an optimal amount of added noise. Stochastic resonance [16,17] also represents a broader class of noise-aided phenomena, and has been exploited for noise-improved information processing for various operations, including signal transmission [18–20], detection [21–24] and estimation [19,25,26].

In the field of signal estimation, Zeitler et al. [5] proved that, for a single-bit dithered quantizer and a Gaussian prior, any estimator using the quantizer output sequence is asymptotically at least $10 \log (\pi/2) \approx 1.97$ dB worse than a minimum MSE (MMSE) estimator using unquantized observations. From quantized observations in multiplicative noise environments, the enhancements of the estimation accuracy by multiplicative noise were reported in [27,28]. For the linear MMSE (LMMSE) case, the optimal added noise for the initial rate of the noise benefit in binary quantizer arrays [29] and for minimizing the MSE of the nonlinear transformation [30] were derived. Using quantized observations of an M-level quantizer, the optimal added noise is proven to be a constant signal level in terms of minimizing the Bayesian CRLB [31]. Recently, an interesting approach to adding noise for signal estimation has been proposed employing an ensemble of estimates from a sufficiently large number of sensors [32–34], whose inputs contain the same signal perturbed by mutually independent noise components but with low-resolution outputs. This ensemble integration method of sensor outputs is asymptotically equivalent to the expectation of estimates with respect to the intentionally added noise distribution [32]. Based on this deduction, these noise-enhanced estimators [32,34] demonstrated effective improvement in the estimation accuracy of a random parameter and in experiments with trained feedforward neural networks [33,35] thus il-
Table 1
Summary of notation.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Notation</th>
<th>Symbol</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>scalar measurement</td>
<td>$\mathbf{x}$</td>
<td>observation vector</td>
</tr>
<tr>
<td>$\theta$</td>
<td>random parameter</td>
<td>$\theta$</td>
<td>parameter vector</td>
</tr>
<tr>
<td>$E_\theta(\theta)$</td>
<td>parameter mean vector</td>
<td>$C_{\theta\theta}$</td>
<td>covariance matrix of $\theta$</td>
</tr>
<tr>
<td>$h$</td>
<td>row vector of $H$</td>
<td>$H$</td>
<td>observation matrix</td>
</tr>
<tr>
<td>$\xi$</td>
<td>background noise</td>
<td>$\xi$</td>
<td>background noise vector</td>
</tr>
<tr>
<td>$g$</td>
<td>sensor transfer function</td>
<td>$w_0$</td>
<td>bias weight coefficient</td>
</tr>
<tr>
<td>$\eta$</td>
<td>added noise</td>
<td>$\eta$</td>
<td>added noise vector</td>
</tr>
<tr>
<td>$y$</td>
<td>sensor output</td>
<td>$y$</td>
<td>node output vector</td>
</tr>
<tr>
<td>$\omega$</td>
<td>weight coefficient for each sensor</td>
<td>$\omega$</td>
<td>network weight vector</td>
</tr>
<tr>
<td>$w$</td>
<td>weight coefficient for each node</td>
<td>$w$</td>
<td>network weight vector</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>network output vector</td>
<td>$f_\eta$</td>
<td>PDF of added noise $\eta$</td>
</tr>
<tr>
<td>$\zeta_{xz}$</td>
<td>correlation vector between $\theta$ and $z$</td>
<td>$C_{xz}$</td>
<td>covariance matrix of $\theta$ and $z$</td>
</tr>
<tr>
<td>$\bar{\theta}_i$</td>
<td>Bayesian estimator</td>
<td>$\mathbf{B}_\theta$</td>
<td>Bayesian MSE of $\theta_i$</td>
</tr>
<tr>
<td>$\mathbf{y}$</td>
<td>average output of node</td>
<td>$\mathbf{y}$</td>
<td>network output vector</td>
</tr>
<tr>
<td>$M$</td>
<td>sensor number</td>
<td>$N$</td>
<td>network node number</td>
</tr>
<tr>
<td>$g^*$</td>
<td>convergence of $\mathbf{y}$ for $M \to \infty$</td>
<td>$G^*$</td>
<td>network output vector for $M \to \infty$</td>
</tr>
<tr>
<td>$\bar{\theta}^M$</td>
<td>Bayesian vector estimator with $M$ identical sensors</td>
<td>$\mathbf{B}^M_{\theta\theta}$</td>
<td>Bayesian MSE matrix of $\bar{\theta}^M$</td>
</tr>
<tr>
<td>$\bar{\theta}^M_{\text{LMMSE}}$</td>
<td>LMMSE estimator on the measurement $\mathbf{x}$</td>
<td>$\mathbf{B}^M_{\theta\theta}$</td>
<td>Bayesian MSE matrix of $\bar{\theta}^M_{\text{NE}}$</td>
</tr>
<tr>
<td>$\mathbb{E}_x(\cdot)$</td>
<td>expectation with respect to the PDF $f_x(x)$</td>
<td>$\mathbb{E}_{x,y}(\cdot)$</td>
<td>expectation with respect to the joint PDF $f_{x,y}(x, y)$</td>
</tr>
<tr>
<td>$\mathbf{1}_M$</td>
<td>$M \times 1$-dimensional unit vector</td>
<td>$\sigma^2$</td>
<td>variance of the random variable</td>
</tr>
<tr>
<td>$\mathbf{1}_N$</td>
<td>$N \times N$-dimensional unit matrix</td>
<td>$\gamma$</td>
<td>quantizer threshold</td>
</tr>
</tbody>
</table>

![Fig. 1](image1.png)

Fig. 1. Block diagram representation of the noise-enhanced estimator $\bar{\theta}_i$ $(i = 1, 2, \ldots, p)$ with $N$ network nodes, and each node composed of $M$ sensors and $M$ mutually independent added noise components $\eta_{mn}$ for $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$.illustrating the benefit of an optimal amount of added noise to the function approximation.

Most of studies of noise-enhanced parameter estimation focus on estimating a scalar parameter $[5,27–30,32,34]$. In many problems of interest we need to estimate a random vector of parameters, e.g., estimating a flow vector from the measurement of hydraulic fields [36–38], the impulse response of a linear system model [39] and target tracking [40]. Then, the noise-enhanced vector parameter or multiple-parameter estimator also triggered various studies [31,41,42] in an M-level or binary quantizer. Another practical vector parameter estimation problem arising from sensor network applications is based on a set of spatially distributed sensors, i.e. distributed parameter estimation $[2–4]$. For this estimation problem, there are two meaningful questions worthy to be answered: how to exploit the artificially added noise to design a vector estimator based on the spatially distributed low-resolution sensors, and to what extent the Bayesian MSE of the designed estimator can be improved by the optimal added noise.

In this paper, we design a Bayesian vector parameter estimator, as represented in Fig. 1, to exploit mutually independent added noise components in the low-resolution network for improving estimation accuracy effectively. We first prove that, for any two identical sensors in a network node, the optimum weight coefficients assigned to two sensors are equal to each other. Then, when each network node is composed of a sufficiently large number of identical sensors, this property leads to significantly simplifed representations of the estimator and its Bayesian MSE matrix by avoiding the costly computation of high-dimensional matrix inversions. For a prior Gaussian joint probability distribution of the random parameters and background noise, the optimal probability density function (PDF) of added noise that minimizes the trace of the Bayesian MSE matrix is theoretically characterized by an approximate expression. When the joint probability distribution is non-Gaussian, the improvement of the designed estimator by the added Gaussian noise is numerically evaluated by simulated realizations of the time averaging algorithm. Theoretical and numerical analyses show that, for an optimal added noise level or optimal added noise type, the Bayesian MSE of the designed estimator is very close to that obtained by the MMSE (or LMMSE) Bayesian estimator on the original measurements over a wide range of the input signal-to-noise ratio (SNR). The main contributions of this paper include: proposing a distributed Bayesian vector estimator based on a low-complexity sensor network and confirming the extent to which intentionally artificially added noise reduces the Bayesian MSE of the proposed estimator. List of symbols in this paper is shown in Table 1.

2. Model and formulation

Consider a sensor network, as shown in Fig. 1, where each network node receives a scalar measurement sequence

$$x_n = h_n \theta + \xi_n, \quad n = 1, 2, \ldots, N.$$ (1)
Here, $h_n$ is the $1 \times p$ row vector of the $N \times p$ ($N > p$) known deterministic observation matrix $H = [h_1, h_2, \ldots, h_N]^\top$, $\theta = [\theta_1, \theta_2, \ldots, \theta_p]^\top$ is a $p \times 1$ unknown random vector parameter with mean $E_{\theta}(\theta)$ and covariance matrix $C_{\theta\theta}$, and the zero-mean background noise variable $\xi_n$ is uncorrelated with $\theta$. Let $x = [x_1, x_2, \ldots, x_N]^\top$ and $\xi = [\xi_1, \xi_2, \ldots, \xi_N]^\top$, we get the linear vector measurement model $x = H\theta + \xi$ [39,40].

As illustrated in Fig. 1, the $n$-th network node has a common input $x_n$, but consists of an ensemble of $M$ sensors with low-resolution outputs

$$y_{nm} = g_{nm}(x_n + \eta_{nm}), \quad m = 1, 2, \ldots, M,$$

(2)

where $g_{nm}$ is the transfer function (e.g. quantizer in Eq. (21)) or hard-limiter in Eq. (22)) of the $m$-th sensor in the node $n$, and $M$ mutually independent added noise components $\eta_{nm}$ are with a common PDF $f_{\eta}$, which is assumed to be designable. Then, we consider the problem of estimating the unknown random vector parameter $\theta$ from sensor observation vectors $y_n = [y_{n1}, y_{n2}, \ldots, y_{nm}]^\top$ for $n = 1, 2, \ldots, N$.

First, as indicated in Fig. 1, we assign an adjustable weight $w_{nm}$ to each sensor output $y_{nm}$ and denote the weight vector as $w = [w_1, w_2, \ldots, w_M]^\top$ and the $MN \times 1$ network output vector $z = [y_{11}, y_{12}, \ldots, y_{MN}]^\top$. Then, the error $\bar{\epsilon}_i = \theta_i - \hat{\theta}_i$ and the Bayesian MSE of the intermediate estimate $\hat{\theta}_i$ becomes

$$E_{x,n}(\epsilon_i^2) = E_{x,n}\left[\left(\theta_i - w_0 - \omega^\top z\right)^2\right].$$

(4)

Setting the derivative $\partial E_{x,n}(\epsilon_i^2)/\partial w_0 = 0$, we find the optimum biasing weight as $w_0^{opt} = E_{x,n}(\theta_i) - \omega^\top E_{x,n}(z)$. Then, Eq. (3) can be rewritten as $\hat{\theta}_i = E_{x,n}(\theta_i) + \omega^\top [z - E_{x,n}(z)]$ and the MN MSE of $\hat{\theta}_i$ in Eq. (4) is obtained as

$$E_{x,n}(\epsilon_i^2) = E_{x,n}\left[\left(\theta_i - E_{x,n}(\theta_i) - \omega^\top [z - E_{x,n}(z)]\right)^2\right] = C_{\theta_i} - 2\omega^\top C_{\theta z} + \omega^\top C_{zz} \omega,$$

(5)

where the variance of the parameter $\theta_i$ is $C_{\theta_i} = E_{x,n}(\theta_i^2) - E_{x,n}(\theta_i)^2$, the covariance matrix $C_{\theta z} = E_{x,n}(z - E_{x,n}(z))(\theta_i - E_{x,n}(\theta_i))^\top$ and the correlation vector $C_{\theta z i} = E_{x,n}(z - E_{x,n}(z))(\theta_i - E_{x,n}(\theta_i))$. It is clear from this expression that the Bayesian MSE $E_{x,n}(\epsilon_i^2)$ is precisely a quadratic function of the weight vector $\omega$ when the vector parameter $\theta$, the background vector noise $\xi$ and sensor noise components $\eta_{nm}$ are stationary stochastic variables. Thus, to obtain the minimum $E_{x,n}(\epsilon_i^2)$ with respect to $\omega$, we set $\partial E_{x,n}(\epsilon_i^2)/\partial \omega = -2C_{\theta z} + 2\omega C_{zz} = 0$ and find the optimum weight vector, i.e. Wiener weight vector $\omega^{opt} = C_{\theta z}^{-1}C_{zz}$. Substituting $\omega^{opt}$ into Eq. (3) and noting $C_{\theta z i} = C_{\theta z i}^{-1}$, we can re-express the estimator $\hat{\theta}_1$ in Eq. (3) as

$$\hat{\theta}_i = E_{x,n}(\theta_i) + C_{\theta z i}^{-1} [z - E_{x,n}(z)].$$

(6)

The corresponding Bayesian MSE of $\hat{\theta}_i$ is given by $B_{\theta_i} = E_{x,n}(\epsilon_i^2)$, and the performance evaluation of the estimator $\hat{\theta}_i$ in Eq. (6) requires the costly computations of the $MN \times 1$ correlation vector $C_{\theta z}$ and the $MN \times MN$ inverse matrix $C_{\theta z}^{-1}$. In order to reduce the calculation cost, we will show that, for a number of identical sensors in each network node, simplifications of the designed estimator $\hat{\theta}_i$ and its performance can be devised.

![Fig. 2. Block diagram representation of the $n$-th network node with $M$ identical sensors $g_{nm}()$ and $M$ mutually independent added noise components $\eta_{nm}$ for $m = 1, 2, \ldots, M$. Only one weight coefficient $w_0$ is assigned to the average node output $y_n$.](image)

**Theorem 1.** For $M$ identical transfer functions $g_{nm}() = g_n()$ of sensors in the $n$-th network node, the $M \times 1$ optimum weight vector $\omega_{i}^{opt}$ of the network node has $M$ equivalent weight coefficients.

Proof of Theorem 1 is presented in Appendix A. Based on Theorem 1, when the transfer functions $g_{nm}()$ of sensors are identical in the $n$-th network node, $M$ weight coefficients $w_{nm}$ are reduced to only one weight coefficient $w_0^{opt}$ indicated in Eq. (A.10). Thus, the estimator $\hat{\theta}_i$ in Eq. (6) and its performance evaluation of $B_{\theta_i}$ can be calculated with low computational complexities as follows.

**Corollary 1.** When the transfer functions $g_{nm}() = g_n()$ of sensors are identical in the $n$-th network node, each network node model can be simplified as Fig. 2 with only one weight coefficient $w_n$ assigned to the average node output

$$\bar{y}_n = \frac{1}{M} \sum_{m=1}^{M} y_{nm}(x_n + \eta_{nm})$$

(7)

for $n = 1, 2, \ldots, N$. Letting the weight vector $w = [w_1, w_2, \ldots, w_N]^\top$ and the network output vector $\bar{y} = [\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_N]^\top$, the designed estimator $\hat{\theta}_i$ in Eq. (6) can be simplified as

$$\hat{\theta}_i^M = E_{x,n}(\theta_i) + C_{\theta z i}^{-1} [\bar{y} - E_{x,n}(\bar{y})]$$

(8)

with its Bayesian MSE

$$B_{\theta_i}^M = E_{x,n}(\epsilon_i^2) = C_{\theta_i} - C_{\theta z i}^{-1} C_{\theta z i}$$

(9)

where the $N \times 1$ correlation vector $C_{\theta z i} = \bar{y} - E_{x,n}(\bar{y})$ and the $N \times N$ covariance matrix $C_{\bar{y} y} = E_{x,n}([\bar{y} - E_{x,n}(\bar{y})]([\bar{y} - E_{x,n}(\bar{y})])^\top)$. The p designed estimator $\hat{\theta}_1^M$ in Eq. (8) can be combined into a vector estimator as

$$\hat{\theta}_1^M = [\hat{\theta}_1^M, \hat{\theta}_2^M, \ldots, \hat{\theta}_p^M]^\top = E_{\theta}(\theta) + C_{\theta z i}^{-1}C_{\theta z i}^{-1} [\bar{y} - E_{x,n}(\bar{y})]$$

(10)

where the $p \times N$ correlation matrix $C_{\theta z i} = C_{\bar{y} y}$ and the $N \times N$ covariance matrix $C_{\bar{y} y} = [C_{\bar{y} y}, C_{\bar{y} y}, \ldots, C_{\bar{y} y}]^\top$. The Bayesian matrix of the vector estimator $\hat{\theta}_1^M$ is given by

$$B_{\theta}^M = E_{x,n}(\epsilon_i^2) = C_{\theta_i} - C_{\theta z i}^{-1} C_{\theta z i}$$

(11)

whose diagonal element $[B_{\theta}^M]_{ii} = B_{\theta_i}^M$ for $i = 1, 2, \ldots, p$.

Proof of Corollary 1 is presented in Appendix B. It is seen that the computation cost of the solution involving matrix inversion is much reduced with the $N \times N$ matrix $C_{\theta z i}^{-1}$ of Eqs. (8) and (9) in comparison with the $MN \times MN$ matrix $C_{\theta z i}^{-1}$ in Eq. (6).

From Eqs. (B.1)–(B.4), it is interesting to note that the node outputs $\bar{y}_n$ and the calculations of $N$ diagonal elements of $C_{\theta z i}$ involve the sensor number $M$. However, the correlation matrix $C_{\theta z i}$ and
the $N(N-1)$ non-diagonal elements of $C_{FF}$ are unrelated to the sensor number $M$. In the following, it will be seen that the designed vector estimator $\hat{\theta}_M^{\text{BNE}}$ in Eq. (10) can be further simplified.

**Corollary 2.** For a sufficiently large number $M$ of identical transfer functions $g_k(\cdot)$ in each network node, the designed estimator $\hat{\theta}_M^{\text{B}}$ of Eq. (8) has an explicit form

$$
\hat{\theta}_i^{\text{B}} = E_\theta(\theta_i) + C_{\theta g}^{-1} g^\top \{ g^*(x) - E_x[g^*(x)] \},
$$

where the network output vector $g^*(x) = [g_1^*(x_1), g_2^*(x_2), \ldots, g_N^*(x_N)]^\top$ and $g_n^*(x_n) = E_x[g_n(x_n + \eta)]$ for $n = 1, 2, \ldots, N$. The Bayesian MSE of the estimator $\hat{\theta}_i^{\text{B}}$ of Eq. (12) can be computed as

$$
B_0^{\text{B}} = C_{\theta} - C_{\theta g} C_{g g}^{-1} C_{g g}^\top C_{\theta g},
$$

where the $1 \times N$ correlation vector $C_{\theta g} = E_\theta[(\theta_i - E_\theta(\theta_i))(g^*(x) - E_x[g^*(x)])^\top]$ and the $N \times N$ covariance matrix $C_{g g} = E_x[(g^*(x))(g^*(x))^\top] - E_x[g^*(x)](g^*(x))^\top$. The vector estimator is given by

$$
\hat{\theta}_M^{\text{BNE}} = [\hat{\theta}_1^{\text{B}}, \hat{\theta}_2^{\text{B}}, \ldots, \hat{\theta}_N^{\text{B}}]^\top
$$

$$
= E_\theta(\theta) + C_{\theta g}^{-1} g^\top \{ g^*(x) - E_x[g^*(x)] \},
$$

where the $p \times N$ correlation matrix $C_{\theta g} = C_{\theta g}^\top = [C_{\theta g_1}, C_{\theta g_2}, \ldots, C_{\theta g_N}]^\top$. The Bayesian matrix of the vector estimator $\hat{\theta}_M^{\text{BNE}}$ is given by

$$
B_0^{\text{BNE}} = E_x[(\theta - \hat{\theta}_M^{\text{BNE}})(\theta - \hat{\theta}_M^{\text{BNE}})^\top] = C_{\theta \theta} - C_{\theta g} C_{g g}^{-1} C_{g g}^\top C_{\theta g},
$$

whose diagonal element $B_{ii}^{\text{BNE}} = B_i^{\text{B}}$ for $i = 1, 2, \ldots, p$.

Proof of Corollary 2 is presented in Appendix 3. It will be seen in Fig. 7 that, for given background noise and added noise, the designed vector estimator $\hat{\theta}_M^{\text{BNE}}$ in Eq. (14) can achieve the minimum Bayesian MSE that sets the lower MSE bound for $\hat{\theta}_M^{\text{B}}$ in Eq. (8).

Furthermore, in order to improve the performance of $\hat{\theta}_M^{\text{BNE}}$ including $\hat{\theta}_M^{\text{BNE}}$ for $M \rightarrow \infty$ by optimally tuning and exploiting the added noise, we face the minimization problem of the trace of the Bayesian matrix, i.e.,

$$
f_\eta^p(\eta) = \arg \min_{\eta} \text{tr}(B_0^{\text{BNE}}),
$$

to find the optimal added noise PDF $\tilde{f}_\eta^p(\eta)$. This functional optimization problem of Eq. (16) is usually analytically intractable. For a numerical resolution, we use a kernel method [32,34] to find an approximate optimal solution as

$$
\tilde{f}_\eta^p(\eta) = \sum_{k=1}^K \lambda_k \phi(\eta, \mu_k, \sigma_k),
$$

where the normalization coefficients $\lambda_k \geq 0$ satisfy the constraint $\sum_{k=1}^K \lambda_k = 1$, and the Gaussian kernel function $\phi(\mu) = \exp(-(\mu - \mu_k)^2/2\sigma_k^2)/\sqrt{2\pi\sigma_k^2}$ with parameters $\mu_k$ and $\sigma_k \geq 0$. Moreover, it has been theoretically proved that, as the number of kernel functions $K$ increases in Eq. (17), the approximate optimal noise PDF $f_\eta^{\text{opt}}$ can converge to the optimal noise PDF $f_\eta^{\text{opt}}$ if it exists [32,43,44]. Thus, for a given kernel function number $K$, the minimization problem of Eq. (16) reduces to a finite-dimensional nonlinear constrained optimization

$$
\begin{align*}
\min_{\lambda_k, \mu_k, \sigma_k} \text{tr}(B_0^{\text{BNE}}) \\
\text{s.t.} \quad & \lambda_k \geq 0, \quad \sum_{k=1}^K \lambda_k = 1, \quad \sigma_k \geq 0, \quad k = 1, 2, \ldots, K.
\end{align*}
$$

with respect to parameters $\lambda_k$, $\mu_k$ and $\sigma_k$ for $k = 1, 2, \ldots, K$. The sequential quadratic programming method [44] has been implemented to solve this nonlinear constrained optimization of Eq. (18). At each major iteration, an approximation is made of the Hessian of the Lagrangian function using a quasi-Newton (BFGS) updating method [44]. This is then used to generate a quadratic programming subproblem whose solution is used to form a search direction for a line search procedure [44].

In the following parts, we will present some illustrative examples to show improvement of the designed estimator $\hat{\theta}_M^{\text{B}}$ by the purposeful addition of artificial noise. For reference, we consider the case where the parameter vector $\theta$ is directly estimated based on the measurement vector $x$, by the LMMSE unbiased estimator given by [39,40]

$$
\hat{\theta}_{\text{LMMSE}} = E_\theta(\theta) + C_{\theta x}^{-1}[x - E_x(x)],
$$

with its Bayesian MSE matrix $[39,40]$

$$
B_{\text{LMMSE}} = E_x[(\theta - \hat{\theta}_{\text{LMMSE}})(\theta - \hat{\theta}_{\text{LMMSE}})^\top] = C_{\theta \theta} - C_{\theta x} C_{x x}^{-1} C_{x \theta},
$$

where the vector mean $E_x(x) = H E_\theta(\theta)$, the cross-covariance matrix $C_{\theta x} = C_{\theta x}^\top$ and the covariance matrix $C_{x x} = H C_{\theta \theta} H^\top + C_{\xi \xi}$ of the background noise vector $\xi$ are given a priori. Note that the LMMSE estimator $\hat{\theta}_{\text{LMMSE}}$ in Eq. (19) is identical in form to the MMSE estimator for jointly Gaussian distribution of parameters and background noise [39,40].

3. Results

We will first discuss some illustrative examples to show the improvement of $\hat{\theta}_M^{\text{B}}$ of Eq. (14) by the purposeful addition of optimal noise $\eta_{\text{opt}}$, wherein the measurement vector $x$ has a joint Gaussian distribution.

**Example 1.** Consider a random vector parameter $\theta = [\theta_1, \theta_2]^\top$ in the linear vector model. Parameters $\theta_1$ and $\theta_2$ are independent, and follow the Gaussian PDF $\tilde{f}_\eta^p(\eta) = \exp[-(\eta - \mu_\eta)^2/2\sigma_\eta^2]/\sqrt{2\pi}\sigma_\eta$. The vector mean $E_\theta(\theta) = [0, 0]^\top$ and the covariance matrix $C_{\theta \theta} = [0.5, 0; 0, 0.1]$, where the semicolon denotes the end of a row of matrix. The zero-mean Gaussian background vector noise $\xi$ is with the covariance matrix $C_{\xi\xi} = 0.1 I_N$ and the known $N \times 2$ observation matrix

$$
H = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
\cos(\frac{2\pi}{N}) & \sin(\frac{2\pi}{N}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\cos(\frac{2\pi(N-1)}{N}) & \sin(\frac{2\pi(N-1)}{N}) & \cdots & 1
\end{bmatrix}.
$$

Consider all network nodes have the identical sensor transfer function

$$
g(u) = \begin{cases} 
1, & u > \gamma, \\
0, & u \leq \gamma,
\end{cases}
$$

which is also known as a binary quantizer with the threshold $\gamma [7,18,19,29,39]$.

Without the added noise $\eta_{\text{opt}}$, the node output of $M$ identical binary quantizers in each node equals to an arbitrary quantizer output $g(x_n)$. Thus, the MSE matrix $B_{\eta_{\text{opt}}}$ in Eq. (15) of the designed estimator $\hat{\theta}_M^{\text{B}}$ of Eq. (14) can only be optimized by the quantizer threshold $\gamma$. It is shown in Fig. 3 that, at the optimal threshold $\gamma = 0$ and without the added noise ($\sigma_\eta = 0$), the Bayesian
MSEs of the estimator $\hat{\theta}_{\text{NE}} = [\hat{\theta}_{1,\text{NE}}, \hat{\theta}_{2,\text{NE}}]$ are minimized as $[B_{\text{NE}}]_{11} = 0.1715$ (■) and $[B_{\text{NE}}]_{22} = 0.0644$ (▲), respectively. Here, the number of measurements $x_n$ is $N = 3$. For comparison, using Eq. (20), the Bayesian MSE matrix $B_{\text{LMMSE}} = [0.0588; 0; 0, 0.0400]$ of the LMMSE estimator $\hat{\theta}_{\text{LMMSE}} = [\hat{\theta}_{1,\text{LMMSE}}, \hat{\theta}_{2,\text{LMMSE}}]$ can be calculated, and the corresponding Bayesian MSEs $[B_{\text{LMMSE}}]_{11}$ and $[B_{\text{LMMSE}}]_{22}$ are also plotted in Fig. 3 by lines. It is seen in Fig. 3 that, without the help of added noise, the performance of $\hat{\theta}_{\text{NE}}$ is significantly worse than that of $\hat{\theta}_{\text{LMMSE}}$, even when the threshold $\gamma$ is increased. However, with the help of the optimal added noise, the much lower Bayesian MSE of the designed estimator can be expected. For a given window number $K = 10$, we use the sequential quadratic programming algorithm [44] to numerically solve the nonlinear constrained optimization problem in Eq. (18), and obtain the optimized parameters $\lambda_k, \mu_k$ and $\sigma_k$ for $k = 1, 2, \ldots, K$. Substituting these parameters into Eq. (17), the approximate optimal noise PDF $f_\eta^* (\eta)$ can be established as illustrated in Fig. 4. Using this approximate optimal PDF $f_\eta^* (\eta)$, we can calculate the corresponding Bayesian MSE matrix as $B^*_{\text{NE}} = [0.0589; 0; 0, 0.0400]$, which is almost equal to the minimum MSE matrix $B_{\text{LMMSE}} = [0.0588; 0; 0, 0.0400]$ obtained by the estimator $\hat{\theta}_{\text{LMMSE}}$ and far better than the Bayesian MSE matrix $B_{\text{NE}} = [0.1715; 0; 0, 0.0644]$ obtained without the added noise. The distribution of noise components $\eta_m$ according to the approximate optimal PDF $f_\eta^* (\eta)$ of Fig. 4 is nontrivial, and its beneficial role in the sensor network for vector parameter estimation is clearly manifested through greatly reducing the Bayesian MSEs of the designed estimator.

An interesting problem is whether or not the designed estimator $\hat{\theta}^*_\text{NE}$ of Eq. (14) based on the binary data approaches the LMMSE estimator $\hat{\theta}^*_\text{LMMSE}$ of Eq. (19) that uses the complete analog data, no matter how the background noise intensity $\sigma_r$ varies. In Fig. 5, we illustrate a number of the Bayesian MSEs of the estimator $\hat{\theta}_{\text{NE}}$ versus the background noise intensity $\sigma_r(\eta)$ is greater than 1.4 (the input SNR $\sigma_r^2/\sigma_n^2$). It is seen in Fig. 5 (a) and (b) that, for a wide range of input SNRs, the estimator $\hat{\theta}^*_\text{NE}$ in Eq. (14) improved by the optimal added noise has a better estimation performance than the case without added noise. Especially, it is well known that, for the jointly Gaussian vector parameter $\theta$ and background noise $\xi$, the LMMSE estimator $\hat{\theta}^*_\text{LMMSE}$ is just the MMSE estimator [39]. As illustrated in Fig. 5, for both parameters $\theta_1$ and $\theta_2$, the Bayesian MSEs of $\hat{\theta}_{\text{NE}}$ are very close to the minimum Bayesian MSEs achieved by $\hat{\theta}_{\text{LMMSE}}$ over the considered range of input SNRs.

Example 2. As the observation number $N$ of $x_n$ increases, the Bayesian MSEs of the noise-enhanced estimator $\hat{\theta}^*_{\text{NE}}$ of Eq. (14) is illustrated in Fig. 6. With the help of the optimal added noise, the Bayesian MSEs of $\hat{\theta}^*_\text{NE}$ all nearly approach the minimum Bayesian MSEs achieved by $\hat{\theta}^*_\text{LMMSE}$ for different observation numbers.

It is noted that, based on the M-level uniformly quantized data, the trace of the Bayesian MSE matrix of the LMMSE estimator is much higher than that of the BCRBL matrix at the low power allocation [3]. It is known that the BCRBL is just the Bayesian MSE achieved by the LMMSE estimator $\hat{\theta}^*_\text{LMMSE}$ in the circumstance of the joint Gaussian distribution of parameters and background noise. In Figs. 5 and 6, based on binary quantized data, the Bayesian MSE of the designed estimator $\hat{\theta}^*_\text{NE}$ in Eq. (14) improved by the optimal added noise is very close to the BCRBL indicated by the primary data $x$. Therefore, compared with the LMMSE estimator [3,45] without exploiting the added noise, the designed estimator $\hat{\theta}^*_\text{NE}$ in Fig. (14) has improved performance when the artificially added noise is optimized.

Example 3. In Example 1, the sensor number $M$ in each sensor network node is assumed to be sufficiently large. However, large sensor number $M$ may not be practical to implement. An important question is how sufficiently large the sensor number $M$ in each node needs to be as the performance of the designed estimator $\hat{\theta}^*_{\text{NE}}$ of Eq. (8) approaches that of the LMMSE estimator $\hat{\theta}^*_\text{LMMSE}$ of Eq. (19).

In Fig. 7, it is seen that the Bayesian MSEs of the designed estimator $\hat{\theta}^*_M_{\text{NE}}$ in Eq. (8) almost completely decrease as the sensor number $M$ increases. When the sensor number $M \geq 3000$, it is shown in Fig. 7 that the designed estimator $\hat{\theta}^*_M_{\text{NE}}$ of Eq. (8), assisted by the optimal added noise, has the corresponding minimum Bayesian MSEs given by diagonal elements $[B^*_{\text{NE}}]_{11} = 0.0598$ and $[B^*_{\text{NE}}]_{22} = 0.0403$, which is very close to the minimum Bayesian MSEs $[B^*_{\text{LMMSE}}]_{11} = 0.0588$ and $[B^*_{\text{LMMSE}}]_{22} = 0.0400$. Here, $f_{\text{LMMSE}}$ is the variance of the noise.
Fig. 5. Bayesian MSEs of the noise-enhanced estimator $\hat{\theta}_{NE}$ in Eq. (14) with the optimal added noise, $\hat{\theta}_{NE}$ without added noise and the LMMSE estimator $\hat{\theta}_{LMMSE}$ in Eq. (19) as a function of the input SNR $\sigma^2_\theta/\sigma^2_\epsilon$ for estimating (a) the parameter $\theta_1$ and (b) the parameter $\theta_2$. Here, the parameters $\theta_1$ and $\theta_2$ are Gaussian distributed with the variances $\sigma^2_\theta = 0.5$ and $\sigma^2_\epsilon = 0.1$, respectively.

Fig. 6. Bayesian MSEs of the noise-enhanced estimator $\hat{\theta}_{NE}$ in Eq. (14) with the approximate optimal PDF $f^{opt}_{\theta_{NE}}(\theta)$ versus the number $N$ of the measurements $x_n$. The other parameters are the same as in Fig. 4.

Fig. 7. Bayesian MSEs of the noise-enhanced estimator $\hat{\theta}_{NE}^M$ in Eq. (8) with the approximate optimal PDF $f^{opt}_{\theta_{NE}}(\theta)$ versus the sensor number $M$ in each network node. The other parameters are the same as in Fig. 4.

[0.1, 4], the Bayesian MSEs of the designed estimator $\hat{\theta}_{NE}^M$ are also observed to be comparable with that of $\hat{\theta}_{LMMSE}$ for a finite sensor number $M = 3000$ (not shown here).

It was proven [5] that, for a single-bit dithered quantizer and a Gaussian prior, any estimator using the binary quantizer output sequence is asymptotically at least $10\log_2(\pi/2) \approx 1.97$ dB worse than MMSE estimator using unquantized observations. In Fig. 7, it is seen that, for $M = 1$ (i.e. the single-bit dithered quantizer), the Bayesian MSEs $[B_{NE}^M]_{11}/[B_{LMMSE}]_{11} > 1.57$ and $[B_{NE}^M]_{12}/[B_{LMMSE}]_{12} \approx 1.57$ also accord with the conclusion in [5]. However, upon the increasing the quantizer number $M$ in each network node, it is shown in Fig. 7 that the designed estimator $\hat{\theta}_{NE}^M$ can break through this restriction of $10\log_2(\pi/2) \approx 1.97$ dB with the cost of establishing a large scale network. The reason is that, in the presence of the optimal added noise, a number of identical quantizers is demonstrated to be equivalent to an M-level quantizer [18,19].

**Example 4.** In Example 3, each network node has $M$ identical sensors receiving the same weight coefficient. However, we may also investigate the impact of non-identical sensors which may arise in practice in a network. First, we consider $N = 3$ network nodes with quantizer thresholds $\gamma = 0, 0.2$, and 0.4, respectively. The other parameters are the same as in Example 1. Using the optimal added noise indicated by the solved PDF $f^\eta_{\theta_{NE}}$, the corresponding Bayesian MSE matrix $B_{NE}^{\eta} = [0.0589, 0; 0, 0.0401]$ can be achieved. Compared with the Bayesian MSEs obtained in Example 1 with identical quantizers over all nodes, the designed estimator $\hat{\theta}_{NE}$ with non-identical quantizers performs almost the same. Second, we settle $M = 3000$ binary quantizers in each network node, but 50 percent of the quantizers is with the threshold $\gamma = 0$ and the other half is with $\gamma = 0.5$. With the injection of optimal added noise, the Bayesian MSE matrix is obtained as $B_{NE}^{M} = [0.0605, 0; 0, 0.0410]$. Compared with the Bayesian MSEs $[B_{NE}^{M}]_{11} = 0.0598$ and $[B_{NE}^{M}]_{12} = 0.0403$ obtained by $\hat{\theta}_{NE}$ with $M = 3000$ identical quantizers in Example 3, the 2% performance degradation of $\hat{\theta}_{NE}^{M}$ with non-identical sensors is at tolerable cost.

In Examples 1–4, using the property of multivariate Gaussian PDF, the minimization problem of the trace of the Bayesian MSE matrix in Eq (18) by the optimal noise becomes possible. However, some parameters to be estimated may have a non-Gaussian distribution in practice. As a consequence, the characterization of the MSE via the second-order moments of Eqs. (B.2)–(B.4) cannot be performed analytically, and a numerical approach has to be implemented. In the next example, we consider another practical vector parameter estimation via numerically computing the first two moments of the sensor outputs, where the added Gaussian noise level is numerically optimized to improve the performance of the designed estimator $\hat{\theta}_{NE}$ in Eq. (8).

**Example 5.** Let mutually independent random parameters $\theta_1$ and $\theta_2$ uniformly distribute in the intervals $[0, 1]$ and $[0, 2]$, respectively. The sensor transfer function is a hard-limiting function
which is also the maximum a posteriori estimator for estimating a scalar uniform parameter buried in Gaussian noise [32, 39, 40]. The size of the measurement vector $x$ is $N = 3$, and the zero-mean background noise $\xi$ has the covariance matrix $C_\xi = 0.11 I_3$. For a given added noise level $\sigma_n$, we add the zero-mean Gaussian noise components $n_{lm}$ to network nodes. Using the Algorithm 1, the first-order and second-order moments of Eqs. (8.2)-(8.4) are numerically calculated, and the designed estimator $\hat{\theta}_{\text{NE}}$ in Eq. (10) and its Bayesian MSE matrix $B^M_{\text{NE}}$ of Eq. (11) are also obtained.

Using the Algorithm 1 and without the added noise ($\sigma_n = 0$), the Bayesian MSE matrix of the designed estimator $\hat{\theta}_{\text{NE}}$ is numerically computed as $B^M_{\text{NE}} = [0.0434, 0.0141; 0.0141, 0.1076]$. Here, the sensor number $M = 10^3$ in each network node. Therefore, as illustrated in Fig. 8, the Bayesian MSEs of $B^M_{\text{NE}}$ and $B^M_{\text{LMMSE}}$ start from 0.0434 and 0.1076 at $\sigma_n = 0$, respectively. It is shown in Fig. 8 that, upon increasing the added noise level $\sigma_n$, the Bayesian MSE matrix can be minimized as $B^M_{\text{NE}} = [0.0382, 0.0008; 0.0008, 0.0593]$ at the optimal added noise level $\sigma^*_n = 1.9$. Compared to the case of no added noise ($\sigma_n = 0$), the MSE of $B^M_{\text{LMMSE}}$ is reduced from 0.0434 (●) to 0.0382 (●), and the MSE of $B^M_{\text{NE}}$ can be reduced from 0.1076 (●) to 0.0593 (●), as shown in Fig. 8. For comparison, the Bayesian MSEs of $\hat{\theta}_{\text{LMMSE}}$ and $\hat{\theta}_{\text{LMMSE}}$ are also illustrated in Fig. 8 by lines, and the Bayesian MSE matrix is $B^M_{\text{LMMSE}} = [0.0370, 0.0; 0.0, 0.0556]$ achieved by the LMMSE estimator $\hat{\theta}_{\text{LMMSE}}$ based on the complete analog data $x$. Furthermore, for different input SNRs of $\sigma^2_\xi/\sigma^2_n$, the corresponding Bayesian MSEs upon increasing the noise intensity $\sigma_n$ of the background noise $\xi$ are also illustrated in Fig. 9. It is seen in Fig. 9 (a) and (b) that, for the considered range of the background noise intensity $\sigma_n$, the designed estimator $\hat{\theta}^M_{\text{NE}}$ aided by the optimal added Gaussian noise does outperform the one without the added noise, and is also very close to the LMMSE estimator $\hat{\theta}_{\text{LMMSE}}$ based on the complete analog data $x$ at low input SNRs of $\sigma^2_\xi/\sigma^2_n$. Note that there is still a visible difference between the MSEs of the noise-enhanced estimator $\hat{\theta}^M_{\text{NE}}$ and the LMMSE estimator $\hat{\theta}_{\text{LMMSE}}$ at high input SNRs of $\sigma^2_\xi/\sigma^2_n$, because the added noise type is restricted to a given Gaussian distribution and the only parameter to be optimized is the noise level $\sigma_n$ under the considered conditions in Example 5.

Algorithm 1: Numerical calculations of $\hat{\theta}^M_{\text{NE}}$ of Eq. (10) and $B^M_{\text{NE}}$ of Eq. (11).

4. Conclusion

In this paper, we design a noise-enhanced Bayesian vector estimator based on low-resolution observations transmitted by sensor networks, where each network node is composed of a large number of low-complexity sensors. When the sensors in the network nodes are identical, we present a simplified calculation method for the design of the noise-enhanced estimator and its Bayesian MSE matrix by reducing the dimensions of the weight vector and network output vector. For illustrative examples with a joint Gaussian distribution of the vector parameter and background noise, we obtain a characterization of the optimal added noise that provides a Bayesian MSE comparable with that of the MMSE estimator built on the primary data. For a joint non-Gaussian case, we intentionally add an optimal amount of Gaussian noise into the designed estimator and also obtain much reduced Bayesian MSE by numerical experiments. The derivations and examples illustrate the feasibility of exploiting added noise in low-complexity sensor networks to enhance the efficiency for information processing.

Some open questions still need to be addressed in future studies. For instance, when the joint PDF of the vector parameter and background noise is non-Gaussian, the theoretical solution of the optimal added noise remains unsolved. We set the type of added noise and only optimize the added noise level to minimize the Bayesian MSE of the designed estimator. Thus, the difference between the MSEs of the noise-enhanced estimator $\hat{\theta}^M_{\text{NE}}$ and the LMMSE estimator $\hat{\theta}_{\text{LMMSE}}$ is evident at high input SNRs, as shown in Fig. 9. How to further reduce the MSE of the noise-enhanced estimator $\hat{\theta}^M_{\text{NE}}$ in the considered case of the joint non-Gaussian PDF of vector parameter and background noise is an open question for future study. In addition, the evaluations of the designed estimator...
and its Bayesian MSE still cannot avoid the matrix inversion, although the matrix dimension is greatly reduced for each network node consisting of identical sensors. Alternatively, in other signal estimation problems the observed data may be on-going in that as time progresses, more data become available and need to be processed sequentially in time [39]. Then, the sequential implementation of the noise-enhanced estimator with no matrix inversion will also be of great interest.

**CRediT authorship contribution statement**

**Fabling Duan:** Conceptualization, Methodology, Creation of models. **Jia Liu:** Data curation, Programming, Writing - Original draft preparation. **François Chapeau-Blondeau and Derek Abbott:** Investigation, Validation, Writing - Reviewing and Editing.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

This work is supported in part by the Taishan Scholar Project of Shandong Province of China (No. ts20190930) and Shandong Key Laboratory of Industrial Control Technology.

**Appendix A. Proof of Theorem 1**

Without loss of generality, we set $M$ identical transfer functions $g_{1m}() = g_1()$ of sensors in the first network node. Letting \( \omega = [\omega_1, \omega_2, \ldots, \omega_M]^T \) and \( \hat{y} = [\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M]^T \), the covariance matrix $C_{yy}$ can be partitioned as

$$
C_{yy} = E_{\eta'}[(y - E_{\eta'}(y))(y - E_{\eta'}(y))^T] = 
\begin{bmatrix}
C_{yy_1} & C_{y_1y_2} \\
C_{y_1y_2}^T & C_{y_2y_2}
\end{bmatrix}, \quad (A.1)
$$

where the partitioned matrices $C_{yy_1} = E_{\eta'}([y_1 - E_{\eta'}(y_1)][y_1 - E_{\eta'}(y_1)])$ and $C_{yy_2} = E_{\eta'}([y_2 - E_{\eta'}(y_2)][y_2 - E_{\eta'}(y_2)])$. Similarly, the correlation vector $C_{y_2}$ can be partitioned as

$$
C_{y_2} = [C_{y_1y_2}, C_{yy_2}]^T, \quad (A.2)
$$

with the partitioned correlation vector $C_{y_1y_2} = E_{\eta'}([y_1 - E_{\eta'}(y_1)][\theta_2 - E_{\theta_2}(\theta_2)])$ and $C_{yy_2} = E_{\eta'}([y_2 - E_{\eta'}(y_2)][\theta_1 - E_{\theta_1}(\theta_1)])$.

Since the optimum weight vector $\omega_{\text{opt}} = C_{yy_2}^{-1}C_{y_2}$ can be rewritten as

$$
\begin{bmatrix}
C_{y_1y_1} & C_{y_1y_2} \\
C_{y_1y_2}^T & C_{y_2y_2}
\end{bmatrix}[\omega_1] = 
\begin{bmatrix}
C_{y_1y_1} \omega_1 \\
C_{y_1y_2} \omega_2
\end{bmatrix}
$$

(A.3)

thus we find

$$
C_{y_1y_1}\omega_1 + C_{y_1y_2}\omega_2 = C_{y_2} \omega_2 \quad (A.4)
$$

For $M$ identical transfer functions $g_{1m}() = g_1()$ in the first network node and the identically distributed noise components $\eta_{1m}$, the covariance matrix $C_{yy_1}$ has $M$ equivalent diagonal elements

$$
[C_{y_1y_1}],_{\text{MMSE}} = E_{\eta_1,\eta}(g_1(x_1 + \eta) - E_\eta\{g_1(x_1 + \eta)\})
$$

and $M (M-1)$ equivalent non-diagonal elements.

For $m, j = 1, 2, \ldots, M$ ($m \neq j$). Its inverse matrix $C_{yy_1}^{-1}$ also has $M$ equivalent diagonal elements $[C_{yy_1}]^{-1},_{\text{MMSE}}$ and $M (M-1)$ equivalent non-diagonal elements $[C_{yy_2}]^{-1},_{\text{MMSE}}$. Therefore, the equality $1_M C_{yy_1}^{-1},_{\text{MMSE}} 1_M = 1_M - M$ yields

$$
[C_{y_1y_1}],_{\text{MMSE}} = (M - 1)\{C_{y_1y_1},_{\text{MMSE}} - 1\}^{-1} \quad (A.7)
$$

The correlation vector $C_{y_1},_{\text{opt}}$ also has $M$ equivalent elements

$$
C_{y_1y_1},_{\text{MMSE}} = E_{\eta_1,\eta}(\hat{y}_1\hat{y}_1) - E_{\eta_1}(\hat{y}_1)E_{\eta_1}(\hat{y}_1) = \omega_{\text{opt}} 1_M \quad (A.8)
$$

Similarly, for $M$ identical transfer functions $g_{1m}() = g_1()$, the $M \times (N - 1)M$ partitioned matrix $C_{yy}$ has $M$ identical row vectors, which can be expressed as $d = E_{\eta_1}(g_1(x_1 + \eta) - E_{\eta_1}(g_1(x_1 + \eta)))$. Therefore, the vector $C_{y_1y_2}\omega_2$ in Eq. (A.4) can be written as $\alpha 1_M$ with $\alpha = d\omega_2$.

Thus, from Eq. (A.4), we find $C_{y_1y_1}\omega_1 = C_{y_1},_{\text{opt}} - C_{y_1y_2}\omega_2 = (C_{y_1},_{\text{MMSE}} - \alpha)1_M$ and the optimum weight vector $\omega_{\text{opt}}$

$$
\omega_{\text{opt}} = C_{y_1y_1},_{\text{MMSE}} - \alpha = \omega_{\text{opt}}^\text{1_M} \quad (A.9)
$$

with

$$
\omega_{\text{opt}}^\text{1_M} = \frac{[C_{y_1y_1},_{\text{MMSE}} - \alpha]}{(M - 1)\{C_{y_1y_1},_{\text{MMSE}} - 1\}} \quad (A.10)
$$

For any of network node with $M$ identical sensors, Eq. (A.10) holds.
Appendix B. Proof of Corollary 1

For $M$ identical transfer functions $g_m(\cdot) = g_0(\cdot)$ of sensors in the $n$-th network node, each node has a scalar output $y_n$ as shown in Fig. 2. Based on $N$ outputs $y_n$, the intermediate linear estimator can be expressed as $\hat{\theta}^M_n = w_0 + w^T_n y_n$, where $w_0$ is the biasing weight. In the same way, setting the derivatives $\partial E_{\hat{x}_n}(\theta,-\hat{\theta})/\partial w_0 = 0$ and $\partial E_{\hat{x}_n}(\theta,-\hat{\theta})/\partial w = 0$, we find the optimum bias $w_0 = E_0(\theta_0) - w^T E_0(y)$ and the optimal weight vector $w^* = C_{y^*}C_{yy}$. Substituting $w_0$ and $w^*$ into $\hat{\theta}^M = w_0 + w^T y$, we find the designed estimator $\hat{\theta}^M_n$ given by Eq. (8).

More specifically, the expectation of the node output $y_n$ in Eq. (7) can be computed as

$$E_{\eta_n}(y_n) = \frac{1}{M} \sum_{m=1}^{M} E_{\eta_n}[g_n(x_n + \eta_m)]$$

$$= E_{\theta_n}[g_n(x_n + \eta)],$$

(B.1)

because the added noise components $\eta_m$ have the common PDF $y_n$. Therefore, the correlation vector $C_{y_n,y_n} = E_{\eta_n}(\theta_0 - E_{\theta_n}(\theta_0))y_n - E_{\theta_n}(y_n) \langle y_n \rangle$ has elements

$$C_{y_n,y_n} = E_{\eta_n}[\theta_0 - E_{\theta_n}(\theta_0)][y_n - E_{\eta_n}(y_n)]$$

$$= E_{\theta_n}[\theta_0 E_{\eta_n}[g_n(x_n + \eta)] - E_{\theta_n}(\theta_0) E_{\eta_n}[g_n(x_n + \eta)]]$$

(B.2)

which is independent of the sensor number $M$. The covariance matrix $C_{y_n,y_n}$ has $N$ diagonal elements

$$C_{y_n,y_n} = E_{\eta_n}[y_n^2] - E_{\eta_n}[y_n]$$

and $N(N-1)$ non-diagonal elements

$$C_{y_n,y_{n'}} = E_{\eta_n}[y_n y_{n'}] - E_{\eta_n}(y_n) E_{\eta_n}(y_{n'}) \langle y_{n'} \rangle$$

$$(n \neq n')$$

$$= \frac{1}{M} \sum_{m=1}^{M} E_{\eta_n}[g_n(x_n + \eta_m)] E_{\eta_n}[g_n(x_{n'} + \eta_m)]$$

$$- E_{\eta_n}[g_n(x_n + \eta)] E_{\eta_n}[g_n(x_{n'} + \eta)]$$

(B.3)

It is noted that the $N(N-1)$ non-diagonal elements $C_{y_n,y_{n'}}$ are also independent of the sensor number $M$. With Eqs. (B.1)–(B.4), the $p$ designed estimator $\hat{\theta}^M_n$ in Eq. (8) and its Bayesian MSEs in Eq. (9) can be obtained.

Appendix C. Proof of Corollary 2

As the sensor number $M \to \infty$ in $N$ network nodes, the node output asymptotically converges to

$$\lim_{M \to \infty} y_n = \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} g_n(x_n + \eta_m) = E_\eta[g_n(x_n + \eta)]$$

(C.1)

and the diagonal elements $[C_{yy}]_{nn}$ in Eq. (B.3) of $C_{y_n,y_n}$ converge to

$$\lim_{M \to \infty} [C_{yy}]_{nn} = E_{\eta_n}[g_n^2(x_n + \eta)] - E_{\eta_n}[g_n(x_n + \eta)]$$

$$= E_{\eta_n}[g_n^2(x_n)] - E_{\theta_n}[g_n^2(x_n)]$$

$$= [g_n^2(x_n)] - [g_n^2(x_n)]$$

(C.2)

because the term $E_{\eta_n}([g_n^2(x_n + \eta)]) < \infty$ in Eq. (B.3). It is also noted that, using Eq. (C.1), Eq. (B.2) and Eq. (4.4), we find that the elements $[C_{y_n,y_n}]_{nn} = E_{\eta_n}[g_n^2(x_n) - E_{\eta_n}(\theta_0) E_{\eta_n}[g_n^2(x_n)] = [C_{yy}]_{nn}$ and the non-diagonal elements $[C_{y_n,y_{n'}}]_{nn'} = E_{\eta_n}[g_n^2(x_n) g_n^2(x_{n'}) - E_{\eta_n}(\theta_0) E_{\eta_n}(\theta_0) E_{\eta_n}[g_n^2(x_n) g_n^2(x_{n'})] = [C_{y_n,y_{n'}}]_{nn'}$.

Substituting Eqs. (C.1) and (C.2) into Eq. (8), we have the explicit form of the estimator $\hat{\theta}^M_n$ in Eq. (12) and its Bayesian MSE $\hat{\theta}^M_n$ in Eq. (13). Combining $p$ estimators of $\hat{\theta}^M_n$ in Eq. (12), we have the vector estimator $\hat{\theta}_{\text{est}}^M$ in Eq. (14) and its Bayesian MSE matrix $\hat{B}_{\text{est}}^M$ in Eq. (15).

References

J. Liu, F. Duan, F. Chapeau-Blondeau et al. Digital Signal Processing 118 (2021) 103224


Jia Liu was born in China in 1995. She received the B. Eng. degree in communication engineering from Qingdao University in 2018 and she is currently working toward the Master degree in system science at Qingdao University. Her research interests are in signal estimation and machine learning.

Fabing Duan was born in China in 1974. He received the Master degree in engineering mechanics from the China University of Mining and Technology (Beijing) in 1999. He received, in 2002, the Ph.D. degree in solid mechanics at Zhejiang University, China. From 2002 to 2003, he was a postdoctoral fellow at the University of Angers, France. Since 2004, he has been with Qingdao University, China, and is currently a professor of system science. His research interests are in nonlinear systems and signal processing.

François Chapeau-Blondeau was born in France in 1959. He received the Engineer Diploma from ESEO, Angers, France, in 1982, the Ph.D. degree in electrical engineering from University Pierre et Marie Curie, Paris 6, France, in 1987, and the Habilitation degree from the University of Angers, France, in 1994. In 1988, he was a research associate in the Department of Biophysics at the Mayo Clinic, Rochester, Minnesota, USA, working on biomedical ultrasounds. Since 1990, he has been with the University of Angers, France, where he is currently a professor of electrical and electronic engineering. His research interests include information theory, signal processing and imaging, and the interactions between physics and information sciences.


From 1978 to 1986, he was a Research Engineer with the GEC Hirst Research Centre, London, U.K. From 1986 to 1987, he was a VLSI Design Engineer with Austek Microsystems, Australia. Since 1987, he has been with The University of Adelaide, where he is currently a full Professor with the School of Electrical and Electronic Engineering. He has authored or coauthored more than 1000 publications and a number of patents. His research interests include multidisciplinary physics and electronic engineering applied to complex systems, networks, game theory, energy policy, stochastics, and biophotonics.

Prof. Abbott has been an invited speaker at more than 100 institutions. He coedited Quantum Aspects of Life (Imperial College Press, 2008), and coauthored Stochastic Resonance (Cambridge Univ. Press, 2008) and Terahertz Imaging for Biomedical Applications (Springer-Verlag, 2012). He is a Fellow of the Institute of Physics. He has been an Editor and/or Guest Editor for a number of journals, including the IEEE Journal of Solid-State Circuits, Journal of Optics B, Microelectronic Journal, PLOS ONE, Proceedings of the IEEE, and the IEEE Photonics Journal. He is currently on the Editorial Boards of IEEE Access, Nature’s Scientific Reports, Royal Society Open Science, and Frontiers in Physics.

Prof. Abbott has received a number of awards, including the Tall Poppy Award for Science (2004), the Premier’s SA Great Award in Science and Technology for outstanding contributions to South Australia (2004), an Australian Research Council Future Fellowship (2012), the David Dewhurst Medal (2015), the Barry Inglis Medal (2018), and the M. A. Sargent Medal (2019) for eminence in engineering.