Digital Signal Processing 118 (2021) 103224

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

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

Jia Liu [a], Fabing Duan [a,*], François Chapeau-Blondeau [b], Derek Abbott [c]

[a] Institute of Complexity Science, Qingdao University, Qingdao 266071, PR China
[b] Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France
[c] Centre for Biomedical Engineering (CBME) and School of Electrical & Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Article history:
Available online 6 September 2021

Keywords:
Distributed Bayesian estimation
Low-resolution sensor
Optimal added noise
Noise-enhanced vector estimator
Stochastic resonance

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-

* Corresponding author.
E-mail address: fabing.duan@gmail.com (F. Duan).

https://doi.org/10.1016/j.dsp.2021.103224
1051-2004/

Table 1
Summary of notation.

Symbol	Notation	Symbol	Notation
x	scalar measurement	$\boldsymbol{x}$	observation vector
θ	random parameter	$\boldsymbol{\theta}$	parameter vector
$\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$	parameter mean vector	$\boldsymbol{C}_{\boldsymbol{\theta\theta}}$	covariance matrix of $\boldsymbol{\theta}$
$\boldsymbol{h}$	row vector of $\boldsymbol{H}$	$\boldsymbol{H}$	observation matrix
ξ	background noise	$\boldsymbol{\xi}$	background noise vector
g	sensor transfer function	w_0	bias weight coefficient
η	added noise	$\boldsymbol{\eta}$	added noise vector
y	sensor output	$\boldsymbol{y}$	node output vector
ω	weight coefficient for each sensor	$\boldsymbol{\omega}$	network weight vector
w	weight coefficient for each node	$\boldsymbol{w}$	network weight vector
$\boldsymbol{z}$	network output vector	f_η	PDF of added noise η
$\boldsymbol{C}_{\theta\boldsymbol{z}}$	correlation vector between θ and $\boldsymbol{z}$	$\boldsymbol{C}_{\boldsymbol{zz}}$	covariance matrix of $\boldsymbol{z}$
$\widehat{\theta}_i$	Bayesian estimator	$B_{\widehat{\theta}_i}$	Bayesian MSE of $\widehat{\theta}_i$
$\overline{y}$	average output of node	$\overline{\boldsymbol{y}}$	network output vector
M	sensor number	N	network node number
g^*	convergence of $\overline{y}$ for $M \to \infty$	$\boldsymbol{g}^*$	network output vector for $M \to \infty$
$\widehat{\boldsymbol{\theta}}^M$	Bayesian vector estimator with M identical sensors	$\boldsymbol{B}_{\mathrm{NE}}^M$	Bayesian MSE matrix of $\widehat{\boldsymbol{\theta}}^M$
$\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$	vector estimator of $\boldsymbol{\theta}$ with infinite identical sensors	$\boldsymbol{B}_{\mathrm{NE}}^*$	Bayesian MSE matrix of $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$
$\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$	LMMSE estimator on the measurement $\boldsymbol{x}$	$\boldsymbol{B}_{\mathrm{LMMSE}}$	Bayesian MSE matrix of $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$
$\mathrm{E}_x(\cdot)$	expectation with respect to the PDF $f_x(x)$	$\mathrm{E}_{x,y}(\cdot)$	expectation with respect to the joint PDF $f_{x,y}(x, y)$
$\boldsymbol{1}_M$	$M \times 1$-dimensional unit vector	σ^2	variance of the random variable
$\boldsymbol{I}_N$	$N \times N$-dimensional unit matrix	γ	quantizer threshold

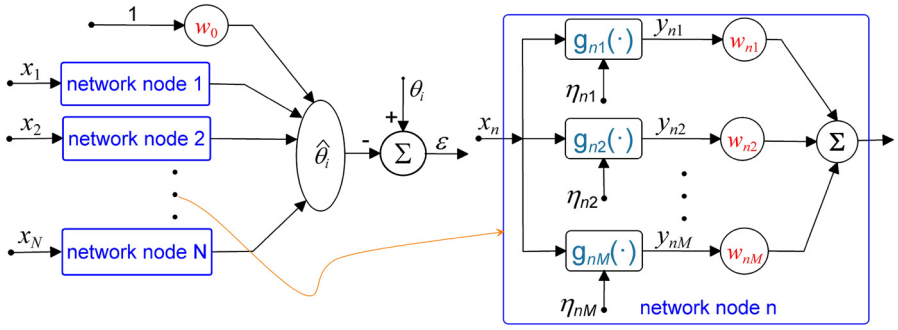

Fig. 1. Block diagram representation of the noise-enhanced estimator $\widehat{\theta}_i$ $(i = 1, 2, \ldots, p)$ with N network nodes, and each node composed of M sensors and M mutually independent added noise components η_{nm} for $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$.

lustrating 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 simplified 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

$$x_n = \boldsymbol{h}_n \boldsymbol{\theta} + \xi_n, \quad n = 1, 2, \ldots, N. \tag{1}$$

Here, $\boldsymbol{h}_n$ is the $1 \times p$ row vector of the $N \times p$ $(N > p)$ known deterministic observation matrix $\boldsymbol{H} = [\boldsymbol{h}_1^\top, \boldsymbol{h}_2^\top, \cdots, \boldsymbol{h}_N^\top]^\top$, $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^\top$ is a $p \times 1$ unknown random vector parameter with mean $\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}}$, and the zero-mean background noise variable ξ_n is uncorrelated with $\boldsymbol{\theta}$. Let $\boldsymbol{x} = [x_1, x_2, \cdots, x_N]^\top$ and $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_N]^\top$, we get the linear vector measurement model $\boldsymbol{x} = \boldsymbol{H}\boldsymbol{\theta} + \boldsymbol{\xi}$ [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

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

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 η_{nm} are with a common PDF f_η, which is assumed to be designable. Then, we consider the problem of estimating the unknown random vector parameter $\boldsymbol{\theta}$ from sensor observation vectors $\boldsymbol{y}_n = [y_{n1}, y_{n2}, \cdots, y_{nM}]^\top$ for $n = 1, 2, \ldots, N$.

First, as indicated in Fig. 1, we assign an adjustable weight ω_{nm} to each sensor output y_{nm} and denote the weight vector as $\boldsymbol{\omega}_n = [\omega_{n1}, \omega_{n2}, \ldots, \omega_{nM}]^\top$, and then the n-th network output is given by $\boldsymbol{\omega}_n^\top \boldsymbol{y}_n$. An intermediate linear estimator is designed as

$$\widehat{\theta}_i = w_0 + \boldsymbol{\omega}^\top \boldsymbol{z} \tag{3}$$

to estimate the i-th parameter θ_i $(i = 1, 2, \ldots, p)$, where w_0 is the biasing weight, the $MN \times 1$ weight vector $\boldsymbol{\omega} = [\boldsymbol{\omega}_1^\top, \boldsymbol{\omega}_2^\top, \ldots, \boldsymbol{\omega}_N^\top]^\top$ and the $MN \times 1$ network output vector $\boldsymbol{z} = [\boldsymbol{y}_1^\top, \boldsymbol{y}_2^\top, \ldots, \boldsymbol{y}_N^\top]^\top$. Then, the error $\varepsilon_i = \theta_i - \widehat{\theta}_i$ and the Bayesian MSE of the intermediate estimate $\widehat{\theta}_i$ becomes

$$\mathrm{E}_{\boldsymbol{x},\eta}(\varepsilon_i^2) = \mathrm{E}_{\boldsymbol{x},\eta}\big[(\theta_i - w_0 - \boldsymbol{\omega}^\top \boldsymbol{z})^2\big]. \tag{4}$$

Setting the derivative $\partial \mathrm{E}_{\boldsymbol{x},\eta}(\varepsilon^2)/\partial w_0 = 0$, we find the optimum biasing weight as $w_0^{\mathrm{opt}} = \mathrm{E}_{\theta_i}(\theta_i) - \boldsymbol{\omega}^\top \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})$. Then, Eq. (3) can be rewritten as $\widehat{\theta}_i = \mathrm{E}_{\theta_i}(\theta_i) + \boldsymbol{\omega}^\top[\boldsymbol{z} - \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})]$ and the Bayesian MSE of $\widehat{\theta}_i$ in Eq. (4) is obtained as

$$\begin{aligned}\mathrm{E}_{\boldsymbol{x},\eta}(\varepsilon_i^2) &= \mathrm{E}_{\boldsymbol{x},\eta}\Big[\Big(\theta_i - \mathrm{E}_{\theta_i}(\theta_i) - \boldsymbol{\omega}^\top[\boldsymbol{z} - \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})]\Big)^2\Big] \\ &= C_{\theta_i} - 2\boldsymbol{\omega}^\top \boldsymbol{C}_{\boldsymbol{z}\theta_i} + \boldsymbol{\omega}^\top \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}} \boldsymbol{\omega},\end{aligned} \tag{5}$$

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

$$\widehat{\theta}_i = \mathrm{E}_{\theta_i}(\theta_i) + \boldsymbol{C}_{\theta_i \boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1}[\boldsymbol{z} - \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})]. \tag{6}$$

The corresponding Bayesian MSE of $\widehat{\theta}_i$ is given by $B_{\widehat{\theta}_i} = \mathrm{E}_{\boldsymbol{x},\eta}[(\theta_i - \widehat{\theta}_i)^2] = C_{\theta_i} - \boldsymbol{C}_{\theta_i \boldsymbol{z}} \boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1} \boldsymbol{C}_{\boldsymbol{z}\theta_i}$. It is seen that, when the sensor number M is large, the performance evaluation of the estimator $\widehat{\theta}_i$ in Eq. (6) requires the costly computations of the $MN \times 1$ correlation vector $\boldsymbol{C}_{\theta_i \boldsymbol{z}}$ and the $MN \times MN$ inverse matrix $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1}$. 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 $\widehat{\theta}_i$ and its performance can be devised.

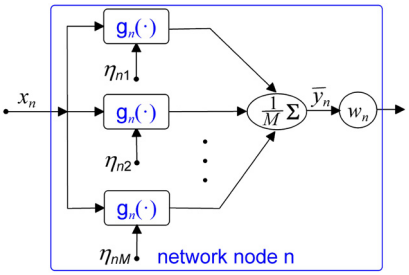

Fig. 2. Block diagram representation of the n-th network nodes with M identical sensors $g_{nm}(\cdot) = g_n(\cdot)$ and M mutually independent added noise components η_{nm} for $m = 1, 2, \ldots, M$. Only one weight coefficient w_n is assigned to the average node output $\overline{y}_n$.

Theorem 1. *For M identical transfer functions $g_{nm}(\cdot) = g_n(\cdot)$ of sensors in the n-th network node, the $M \times 1$ optimum weight vector $\boldsymbol{\omega}_n^{\mathrm{opt}}$ 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 $g_{nm}(\cdot)$ of sensors are identical in the n-th network node, M weight coefficients w_{nm} are reduced to only one weight coefficient w_n^{opt} indicated in Eq. (A.10). Thus, the estimator $\widehat{\theta}_i$ in Eq. (6) and its performance evaluation of $B_{\widehat{\theta}_i}$ can be calculated with low computational complexities as follows.

Corollary 1. *When the transfer functions $g_{nm}(\cdot) = g_n(\cdot)$ 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*

$$\overline{y}_n = \frac{1}{M}\sum_{m=1}^{M} g_n(x_n + \eta_{nm}) \tag{7}$$

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

$$\widehat{\theta}_i^M = \mathrm{E}_{\theta_i}(\theta_i) + \boldsymbol{C}_{\theta_i \overline{\boldsymbol{y}}} \boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1}[\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})] \tag{8}$$

with its Bayesian MSE

$$B_{\widehat{\theta}_i^M} = \mathrm{E}_{\boldsymbol{x},\theta_i}[(\theta_i - \widehat{\theta}_i^M)^2] = C_{\theta_i} - \boldsymbol{C}_{\theta_i \overline{\boldsymbol{y}}} \boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1} \boldsymbol{C}_{\overline{\boldsymbol{y}}\theta_i}, \tag{9}$$

where the $1 \times N$ correlation vector $\boldsymbol{C}_{\theta_i \overline{\boldsymbol{y}}} = \mathrm{E}_{\boldsymbol{x},\eta}\{[\theta_i - \mathrm{E}_{\theta_i}(\theta_i)][\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})]^\top\}$ and the $N \times N$ covariance matrix $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}} = \mathrm{E}_{\boldsymbol{x},\eta}\{[\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})] \times [\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})]^\top\}$. The p designed estimator $\widehat{\theta}_i^M$ in Eq. (8) can be combined into a vector estimator as

$$\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M = [\widehat{\theta}_1^M, \widehat{\theta}_2^M, \ldots, \widehat{\theta}_p^M]^\top = \mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) + \boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}} \boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1}[\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})], \tag{10}$$

where the $p \times N$ correlation matrix $\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}} = \boldsymbol{C}_{\overline{\boldsymbol{y}}\boldsymbol{\theta}}^\top = [\boldsymbol{C}_{\theta_1 \overline{\boldsymbol{y}}}^\top, \boldsymbol{C}_{\theta_2 \overline{\boldsymbol{y}}}^\top, \ldots, \boldsymbol{C}_{\theta_p \overline{\boldsymbol{y}}}^\top]^\top$. The Bayesian matrix of the vector estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M$ is given by

$$\boldsymbol{B}_{\mathrm{NE}}^M = \mathrm{E}_{\boldsymbol{x},\eta}[(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M)(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M)^\top] = \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}} - \boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}} \boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1} \boldsymbol{C}_{\overline{\boldsymbol{y}}\boldsymbol{\theta}}, \tag{11}$$

whose diagonal element $[\boldsymbol{B}_{\mathrm{NE}}^M]_{ii} = B_{\widehat{\theta}_i^M}$ 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 $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1}$ of Eqs. (8) and (9) in comparison with the $MN \times MN$ matrix $\boldsymbol{C}_{\boldsymbol{z}\boldsymbol{z}}^{-1}$ in Eq. (6).

From Eqs. (B.1)–(B.4), it is interesting to note that the node outputs $\overline{y}_n$ and the calculations of N diagonal elements of $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}$ involve the sensor number M. However, the correlation matrix $\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}}$ and

the $N(N-1)$ non-diagonal elements of $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}$ are unrelated to the sensor number M. In the following, it will be seen that the designed vector estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^{M}$ in Eq. (10) can be further simplified.

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

$$\widehat{\theta}_i^* = \mathrm{E}_{\theta_i}(\theta_i) + \boldsymbol{C}_{\theta_i \boldsymbol{g}^*} \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}^{-1} \{\boldsymbol{g}^*(\boldsymbol{x}) - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{g}^*(\boldsymbol{x})]\}, \tag{12}$$

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

$$B_{\widehat{\theta}_i^*} = C_{\theta_i} - \boldsymbol{C}_{\theta_i \boldsymbol{g}^*} \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}^{-1} \boldsymbol{C}_{\boldsymbol{g}^*\theta_i}, \tag{13}$$

where the $1 \times N$ correlation vector $\boldsymbol{C}_{\theta_i \boldsymbol{g}^*} = \mathrm{E}_{\boldsymbol{x}}\{[\theta_i - \mathrm{E}_{\theta_i}(\theta_i)](\boldsymbol{g}^*(\boldsymbol{x}) - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{g}^*(\boldsymbol{x})])^\top\}$ and the $N \times N$ covariance matrix $\boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*} = \mathrm{E}_{\boldsymbol{x}}\{(\boldsymbol{g}^*(\boldsymbol{x}) - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{g}^*(\boldsymbol{x})])(\boldsymbol{g}^*(\boldsymbol{x}) - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{g}^*(\boldsymbol{x})])^\top\}$. The vector estimator is given by

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^* &= [\widehat{\theta}_1^*, \widehat{\theta}_2^*, \ldots, \widehat{\theta}_p^*]^\top \\ &= \mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) + \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{g}^*} \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}^{-1} \{\boldsymbol{g}^*(\boldsymbol{x}) - \mathrm{E}_{\boldsymbol{x}}[\boldsymbol{g}^*(\boldsymbol{x})]\},\end{aligned} \tag{14}$$

where the $p \times N$ correlation matrix $\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{g}^*} = \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{\theta}}^\top = [\boldsymbol{C}_{\theta_1 \boldsymbol{g}^*}^\top, \boldsymbol{C}_{\theta_2 \boldsymbol{g}^*}^\top, \ldots, \boldsymbol{C}_{\theta_p \boldsymbol{g}^*}^\top]^\top$. The Bayesian matrix of the vector estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ is given by

$$\boldsymbol{B}_{\mathrm{NE}}^* = \mathrm{E}_{\boldsymbol{x}}[(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*)(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*)^\top] = \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}} - \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{g}^*} \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}^{-1} \boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{\theta}} \tag{15}$$

whose diagonal element $[\boldsymbol{B}_{\mathrm{NE}}^*]_{ii} = B_{\widehat{\theta}_i^*}$ for $i = 1, 2, \ldots, p$.

Proof of Corollary 2 is presented in Appendix C. It will be seen in Fig. 7 that, for given background noise and added noise, the designed vector estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ in Eq. (14) can achieve the minimum Bayesian MSE that sets the lower MSE bound for $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M$ in Eq. (8).

Furthermore, in order to improve the performance of $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M$ including $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ for $M \to \infty$ by optimally tuning and exploiting the added noise, we face the minimization problem of the trace of the Bayesian matrix, i.e.

$$f_\eta^{\mathrm{o}}(\eta) = \arg\min_{f_\eta} \mathrm{tr}(\boldsymbol{B}_{\mathrm{NE}}^M), \tag{16}$$

to find the optimal added noise PDF $f_\eta^{\mathrm{o}}(\eta)$. 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

$$\widetilde{f}_\eta^{\mathrm{o}}(\eta) = \sum_{k=1}^{K} \lambda_k \phi(\eta, \mu_k, \sigma_k), \tag{17}$$

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 μ_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 $\widetilde{f}_\eta^{\mathrm{opt}}$ can converge to the optimal noise PDF f_η^{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

$$\begin{aligned}&\min_{\lambda_k, \mu_k, \sigma_k} \quad \mathrm{tr}(\boldsymbol{B}_{\mathrm{NE}}^M) \\ &\text{s.t.} \quad \lambda_k \geq 0, \ \sum_{k=1}^{K} \lambda_k = 1, \ \sigma_k \geq 0,\end{aligned} \tag{18}$$

with respect to parameters λ_k, μ_k and σ_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 $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M$ by the purposeful addition of artificial noise. For reference, we consider the case where the parameter vector $\boldsymbol{\theta}$ is directly estimated based on the measurement vector $\boldsymbol{x}$, by the LMMSE unbiased estimator given by [39,40]

$$\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}} = \mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) + \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{x}} \boldsymbol{C}_{\boldsymbol{x}}^{-1} [\boldsymbol{x} - \mathrm{E}_{\boldsymbol{x}}(\boldsymbol{x})], \tag{19}$$

with its Bayesian MSE matrix [39,40]

$$\begin{aligned}\boldsymbol{B}_{\mathrm{LMMSE}} &= \mathrm{E}_{\boldsymbol{x}}[(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}})(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}})^\top] \\ &= \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}} - \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{x}} \boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}}^{-1} \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{x}}^\top,\end{aligned} \tag{20}$$

where the vector mean $\mathrm{E}_{\boldsymbol{x}}(\boldsymbol{x}) = \boldsymbol{H}\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$, the cross-covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{x}} = \boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}}\boldsymbol{H}^\top$, the covariance matrix $\boldsymbol{C}_{\boldsymbol{x}\boldsymbol{x}} = \boldsymbol{H}\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}}\boldsymbol{H}^\top + \boldsymbol{C}_{\boldsymbol{\xi}\boldsymbol{\xi}}$ and the covariance matrix $\boldsymbol{C}_{\boldsymbol{\xi}\boldsymbol{\xi}}$ of the background noise vector $\boldsymbol{\xi}$ are given *a priori*. Note that the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ 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 $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ of Eq. (14) by the purposeful addition of optimal noise η_{nm}, wherein the measurement vector $\boldsymbol{x}$ has a joint Gaussian distribution.

Example 1. Consider a random vector parameter $\boldsymbol{\theta} = [\theta_1, \theta_2]^\top$ in the linear vector model. Parameters θ_1 and θ_2 are independent and follow the Gaussian PDF $f_{\theta_i}(\theta_i) = \exp[-(\theta_i - u_{\theta_i})^2/2\sigma_{\theta_i}^2]/\sqrt{2\pi\sigma_{\theta_i}^2}$ with variance $\sigma_{\theta_i}^2$ for $i = 1, 2$, respectively. Assume that the vector mean $\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = [0, 0]^T$ and the covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{\theta}} = [0.5, 0; 0, 0.1]$, where the semicolon denotes the end of a row of matrix. The zero-mean Gaussian background vector noise $\boldsymbol{\xi}$ is with the covariance matrix $\boldsymbol{C}_{\boldsymbol{\xi}\boldsymbol{\xi}} = 0.1\boldsymbol{I}_N$ and the known $N \times 2$ observation matrix

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

Consider all network nodes have the identical sensor transfer function

$$g(u) = \begin{cases} 1, & u > \gamma, \\ 0, & u \leq \gamma, \end{cases} \tag{21}$$

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

Without the added noise η_{nm}, the node output of M identical binary quantizers in each node equals to an arbitrary quantizer output $g(x_n)$. Thus, the MSE matrix $\boldsymbol{B}_{\mathrm{NE}}^*$ in Eq. (15) of the designed estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ of Eq. (14) can only be optimized by the quantizer threshold γ. It is shown in Fig. 3 that, at the optimal threshold $\gamma = 0$ and without the added noise ($\sigma_\eta = 0$), the Bayesian

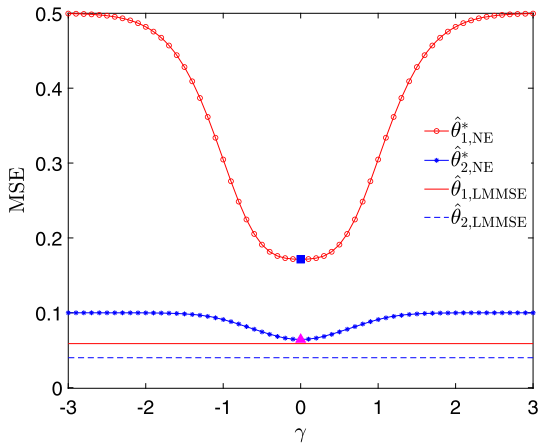

Fig. 3. Bayesian MSEs $[\boldsymbol{B}^*_{\mathrm{NE}}]_{11}$ and $[\boldsymbol{B}^*_{\mathrm{NE}}]_{22}$ of the designed estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}} = [\widehat{\theta}^*_{1,\mathrm{NE}}, \widehat{\theta}^*_{2,\mathrm{NE}}]^\top$ in Eq. (14) versus the quantizer threshold γ without the added noise. The number of measurements x_n is $N = 3$. For comparison, the Bayesian MSEs $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{11} = 0.0588$ and $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{22} = 0.0400$ of the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}} = [\widehat{\theta}_{1,\mathrm{LMMSE}}, \widehat{\theta}_{2,\mathrm{LMMSE}}]^\top$ in Eq. (19) are also plotted by lines. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

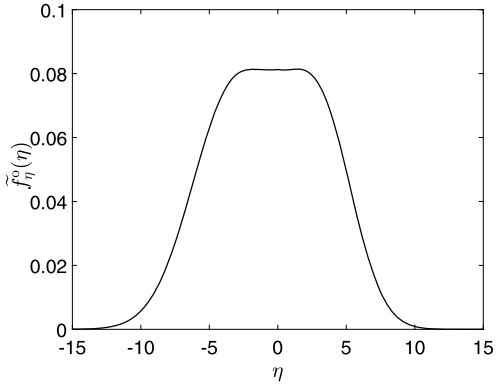

Fig. 4. Approximate optimal PDF $\widetilde{f}^{\mathrm{o}}_{\eta}(\eta)$ of the added noise for the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ with the kernel window number $K = 10$ in Eq. (17). The other parameters are the same as in Fig. 3.

MSEs of the estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}} = [\widehat{\theta}^*_{1,\mathrm{NE}}, \widehat{\theta}^*_{2,\mathrm{NE}}]^\top$ are minimized as $[\boldsymbol{B}^*_{\mathrm{NE}}]_{11} = 0.1715$ (■) and $[\boldsymbol{B}^*_{\mathrm{NE}}]_{22} = 0.0644$ (▲), respectively. Here, the number of measurements x_n is $N = 3$. For comparison, using Eq. (20), the Bayesian MSE matrix $\boldsymbol{B}_{\mathrm{LMMSE}} = [0.0588, 0; 0, 0.0400]$ of the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}} = [\widehat{\theta}_{1,\mathrm{LMMSE}}, \widehat{\theta}_{2,\mathrm{LMMSE}}]^\top$ can be calculated, and the corresponding Bayesian MSEs $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{11}$ and $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{22}$ are also plotted in Fig. 3 by lines. It is seen in Fig. 3 that, without the help of added noise, the performance of $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ is significantly worse than that of $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$, even when the threshold γ is optimized.

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 λ_k, μ_k and σ_k for $k = 1, 2, \ldots, K$. Substituting these parameters into Eq. (17), the approximate optimal noise PDF $\widetilde{f}^{\mathrm{o}}_{\eta}$ can be established as illustrated in Fig. 4. Using this approximate optimal PDF $\widetilde{f}^{\mathrm{o}}_{\eta}$, we can calculate the corresponding Bayesian MSE matrix as $\boldsymbol{B}^*_{\mathrm{NE}} = [0.0589, 0; 0, 0.0400]$, which is almost equal to the minimum MSE matrix $\boldsymbol{B}_{\mathrm{LMMSE}} = [0.0588, 0; 0, 0.0400]$ obtained by the estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ and far better than the Bayesian MSE matrix $\boldsymbol{B}^*_{\mathrm{NE}} = [0.1715, 0; 0, 0.0644]$ obtained without the added noise. The distribution of noise components η_{nm} according to the approximate optimal PDF $\widetilde{f}^{\mathrm{o}}_{\eta}(\eta)$ 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 result is that the optimization of the Bayesian MSE matrix $\boldsymbol{B}^*_{\mathrm{NE}}$ by the optimal noise is independent of the threshold value of γ. Whatever the quantizer threshold $\gamma = 1$ or 0, the same minimum Bayesian MSE matrix $\boldsymbol{B}^*_{\mathrm{NE}} = [0.0589, 0; 0, 0.0400]$ can be achieved by the corresponding optimal PDF solutions of $\widetilde{f}^{\mathrm{o}}_{\eta}(\eta)$. The reason is that the optimal PDF $\widetilde{f}^{\mathrm{o}}_{\eta}(\eta)$ contains the adjustable mean parameters μ_k that can eliminate the influence of the quantizer threshold on the parameter vector estimation, so that the performance of the designed estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ can be maintained at a stable and high level.

Another interesting problem is whether or not the designed estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ of Eq. (14) based on the binary data approaches the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ of Eq. (19) that uses the complete analog data $\boldsymbol{x}$, no matter how the background noise intensity σ_ξ varies. In Fig. 5, we illustratively show the Bayesian MSEs of the estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ versus the background noise intensity $\sigma_\xi \in [0.1, 4]$ (the input SNR $\sigma^2_\theta/\sigma^2_\xi$). It is seen in Fig. 5 (a) and (b) that, for a wide range of input SNRs, the estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ 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 $\boldsymbol{\theta}$ and background noise $\boldsymbol{\xi}$, the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ is just the MMSE estimator [39]. As illustrated in Fig. 5, for both parameters θ_1 and θ_2, the Bayesian MSEs of $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ are very close to the minimum Bayesian MSEs achieved by $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ 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 $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ of Eq. (14) is illustrated in Fig. 6. With the help of the optimal added noise, the Bayesian MSEs of $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ all nearly approach the minimum Bayesian MSEs achieved by $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ 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 BCRLB matrix at the low power allocation [3]. It is known that the BCRLB is just the Bayesian MSE achieved by the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ 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 $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ in Eq. (14) improved by the optimal added noise is very close to the BCRLB indicated by the primary data $\boldsymbol{x}$. Therefore, compared with the LMMSE estimator [3,45] without exploiting the added noise, the designed estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ in Eq. (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 $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ of Eq. (8) approaches that of the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ of Eq. (19).

In Fig. 7, it is seen that the Bayesian MSEs of the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ in Eq. (8) monotonically decrease as the sensor number M increases. When the sensor number $M \geq 3000$, it is shown in Fig. 7 that the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ in Eq. (8), assisted by the optimal added noise, has the corresponding minimum Bayesian MSEs given by diagonal elements $[\boldsymbol{B}^M_{\mathrm{NE}}]_{11} = 0.0598$ and $[\boldsymbol{B}^M_{\mathrm{NE}}]_{22} = 0.0403$, which is very close to the minimum Bayesian MSEs $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{11} = 0.0588$ and $[\boldsymbol{B}_{\mathrm{LMMSE}}]_{22} = 0.0400$ of the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$. For different background noise intensities $\sigma_\xi \in$

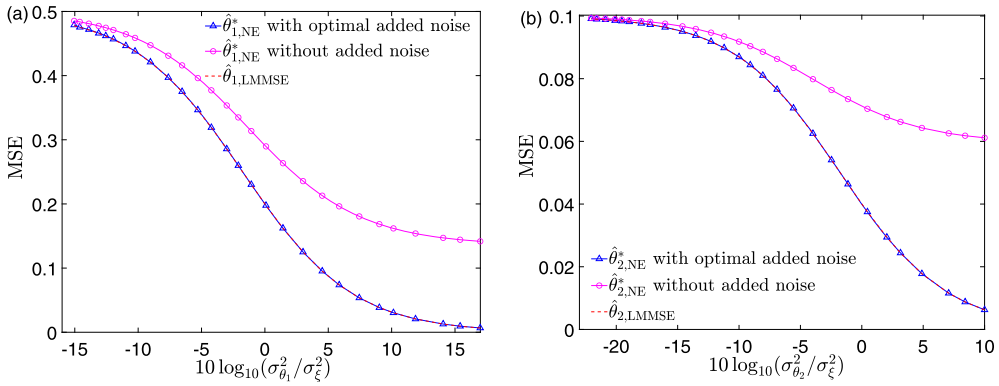

Fig. 5. Bayesian MSEs of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ in Eq. (14) with the optimal added noise, $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ without added noise and the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ in Eq. (19) as a function of the input SNR $\sigma^2_\theta/\sigma^2_\xi$ for estimating (a) the parameter θ_1 and (b) the parameter θ_2. Here, the parameters θ_1 and θ_2 are Gaussian distributed with the variances $\sigma^2_{\theta_1}=0.5$ and $\sigma^2_{\theta_1}=0.1$, respectively.

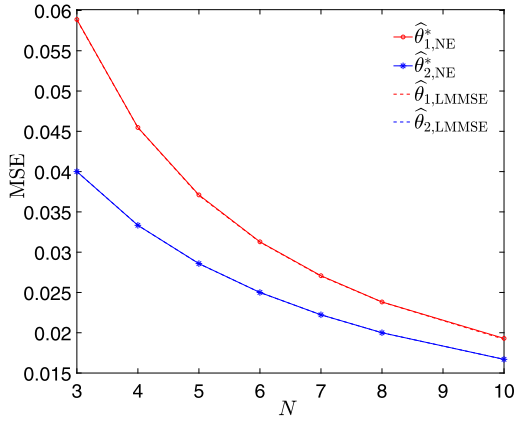

Fig. 6. Bayesian MSEs of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ in Eq. (14) with the approximate optimal PDF $\widetilde{f}^{\mathrm{opt}}_\eta(\eta)$ versus the number N of the measurements x_n. The other parameters are the same as in Fig. 4.

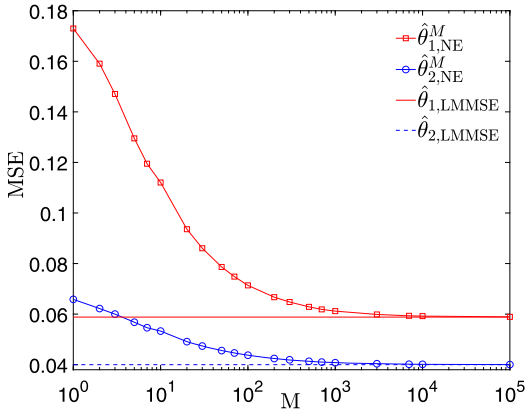

Fig. 7. Bayesian MSEs of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ in Eq. (8) with the approximate optimal PDF $\widetilde{f}^{\mathrm{opt}}_\eta(\eta)$ versus the sensor number M in each network node. The other parameters are the same as in Fig. 4.

$[0.1, 4]$, the Bayesian MSEs of the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ are also observed to be comparable with that of $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ 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 $10\log(\pi/2) \approx 1.97$ 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 $[\boldsymbol{B}^M_{\mathrm{NE}}]_{11}/[\boldsymbol{B}_{\mathrm{LMMSE}}]_{11} > 1.57$ and $[\boldsymbol{B}^M_{\mathrm{NE}}]_{22}/[\boldsymbol{B}_{\mathrm{LMMSE}}]_{22} \approx 1.57$ 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 $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ can break through this restriction of $10\log(\pi/2) \approx 1.97$ 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 $\widetilde{f}^{\mathrm{o}}_\eta$, the corresponding Bayesian MSE matrix $\boldsymbol{B}^*_{\mathrm{NE}} = [0.0589, 0; 0, 0.0401]$ can be achieved. Compared with the Bayesian MSEs obtained in Example 1 with identical quantizers over all nodes, the designed estimator $\widehat{\boldsymbol{\theta}}^*_{\mathrm{NE}}$ with non-identical quantizers performs almost the same. Secondly, 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 $\boldsymbol{B}^M_{\mathrm{NE}} = [0.0605, 0; 0, 0.0410]$. Compared with the Bayesian MSEs $[\boldsymbol{B}^M_{\mathrm{NE}}]_{11} = 0.0598$ and $[\boldsymbol{B}^M_{\mathrm{NE}}]_{22} = 0.0403$ obtained by $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ with $M = 3000$ identical quantizers in Example 3, the 2% performance degradation of $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ 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 $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ in Eq. (8).

Example 5. Let mutually independent random parameters θ_1 and θ_2 uniformly distribute in the intervals $[0, 1]$ and $[0, 2]$, respectively. The sensor transfer function is a hard-limiting function

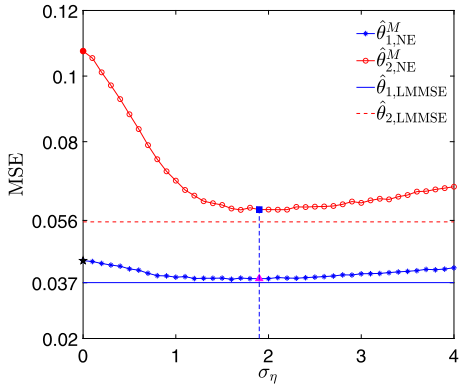

Fig. 8. Plots of Bayesian MSEs of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ as a function of the added Gaussian noise level σ_η. Here, the random parameters θ_1 and θ_2 uniformly distribute in the intervals $[0,1]$ and $[0,2]$, respectively. The sensor number $M=10^3$ in each node, and the number of simulation trials is 10^5.

$$g(u)=\begin{cases}0, & u<0,\\ u, & 0\le u<1,\\ 1, & u\ge 1,\end{cases} \tag{22}$$

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 $\boldsymbol{x}$ is $N=3$, and the zero-mean background noise $\boldsymbol{\xi}$ has the covariance matrix $\boldsymbol{C}_\xi=0.1\boldsymbol{I}_3$. For a given added noise level σ_η, we add the zero-mean Gaussian noise components η_{nm} to network nodes. Using the Algorithm 1, the first-order and second-order moments of Eqs. (B.2)–(B.4) are numerically calculated, and the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ in Eq. (10) and its Bayesian MSE matrix $\boldsymbol{B}^M_{\mathrm{NE}}$ of Eq. (11) are also obtained.

Using the Algorithm 1 and without the added noise ($\sigma_\eta=0$), the Bayesian MSE matrix of the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ is numerically computed as $\boldsymbol{B}^M_{\mathrm{NE}}=[0.0434, 0.0141;\ 0.0141, 0.1076]$. Here, the sensor number $M=10^3$ in each network node. Therefore, as illustrated in Fig. 8, the Bayesian MSEs of $[\boldsymbol{B}^M_{\mathrm{NE}}]_{11}$ and $[\boldsymbol{B}^M_{\mathrm{NE}}]_{22}$ start from 0.0434 and 0.1076 at $\sigma_\eta=0$, respectively. It is shown in Fig. 8 that, upon increasing the added noise level σ_η, the Bayesian MSE matrix can be minimized as $\boldsymbol{B}^M_{\mathrm{NE}}=[0.0382, 0.0008; 0.0008, 0.0593]$ at the optimal added noise level $\sigma_\eta^{\mathrm{opt}}=1.9$. Compared to the case of no added noise ($\sigma_\eta=0$), the MSE of $\widehat{\theta}^M_{1,\mathrm{NE}}$ is reduced from 0.0434 (★) to 0.0382 (▲), and the MSE of $\widehat{\theta}^M_{2,\mathrm{NE}}$ can be reduced from 0.1076 (●) to 0.0593 (■), as shown in Fig. 8. For comparison, the Bayesian MSEs of $\widehat{\theta}_{1,\mathrm{LMMSE}}$ and $\widehat{\theta}_{2,\mathrm{LMMSE}}$ are also illustrated in Fig. 8 by lines, and the Bayesian MSE matrix is $\boldsymbol{B}_{\mathrm{LMMSE}}=[0.0370, 0; 0, 0.0556]$ achieved by the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ based on the complete analog data $\boldsymbol{x}$.

Furthermore, for different input SNRs of $\sigma_\theta^2/\sigma_\xi^2$, the corresponding Bayesian MSEs upon increasing the noise intensity σ_ξ of the background noise $\boldsymbol{\xi}$ 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 σ_ξ, the designed estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ aided by the optimal added Gaussian noise does outperform the one without the added noise, and is also very close to the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ based on the complete analog data $\boldsymbol{x}$ at low input SNRs of $\sigma_\theta^2/\sigma_\xi^2$. Note that there still is a visible difference between the MSEs of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ and the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ at high input SNRs of $\sigma_\theta^2/\sigma_\xi^2$, because the added noise type is restricted to a given Gaussian distribution and the only parameter to be optimized is the noise level σ_η under the considered conditions in Example 5.

Input: Data number N, sensor number M in every node, sensor function g, parameter vectors $\boldsymbol{\theta}^t$ with known mean vector $\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta})$, background noise vectors $\boldsymbol{\xi}^t$, measurement vectors $\boldsymbol{x}^t=\boldsymbol{H}\boldsymbol{\theta}^t+\boldsymbol{\xi}^t$ and the added noise vectors $\boldsymbol{\eta}^t_n$ with a given noise level σ_η of added Gaussian noise. t is the times of simulation and T is the total simulation number.

Output: $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$, $\boldsymbol{B}^M_{\mathrm{NE}}$.

for n=1...N **do**
  for t=1...T **do**
    $\overline{y}^t_n=\frac{1}{M}\sum_{m=1}^{M} g(x^t_n\mathbf{1}_M+\boldsymbol{\eta}^t_n)$;
  end
  $\mathrm{E}(\overline{y}_n)\approx\frac{1}{T}\sum_{t=1}^{T}\overline{y}^t_n$;
end
for t=1...T **do**
  for n=1...N **do**
    for i=1...p **do**
      $[\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}}]^t_{in}=[\theta^t_i-\mathrm{E}_{\theta_i}(\theta_i)][\overline{y}^t_n-\mathrm{E}(\overline{y}_n)]$;
    end
    for ℓ=n...N **do**
      $[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]^t_{n\ell}=[\overline{y}^t_n-\mathrm{E}(\overline{y}_n)][\overline{y}^t_\ell-\mathrm{E}(\overline{y}_\ell)]$;
    end
  end
  $[\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}}]_{in}=\frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}}]^t_{in}$;
  $[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{n\ell}=\frac{1}{T}\sum_{t=1}^{T}[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]^t_{n\ell}$;
  $\widehat{\boldsymbol{\theta}}^{M,t}_{\mathrm{NE}}\approx\mathrm{E}_{\boldsymbol{\theta}}(\boldsymbol{\theta})+\boldsymbol{C}_{\boldsymbol{\theta}\overline{\boldsymbol{y}}}\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}\begin{pmatrix}\overline{y}^t_1-\mathrm{E}(\overline{y}_1)\\ \overline{y}^t_2-\mathrm{E}(\overline{y}_2)\\ \vdots\\ \overline{y}^t_N-\mathrm{E}(\overline{y}_N)\end{pmatrix}$;
end
$\boldsymbol{B}^M_{\mathrm{NE}}\approx\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\theta}^t-\widehat{\boldsymbol{\theta}}^{M,t}_{\mathrm{NE}})(\boldsymbol{\theta}^t-\widehat{\boldsymbol{\theta}}^{M,t}_{\mathrm{NE}})^\top$.

Algorithm 1: Numerical calculations of $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ of Eq. (10) and $\boldsymbol{B}^M_{\mathrm{NE}}$ 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 the 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 $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ and the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ is evident at high input SNRs, as shown in Fig. 9. How to further reduce the MSE of the noise-enhanced estimator $\widehat{\boldsymbol{\theta}}^M_{\mathrm{NE}}$ 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

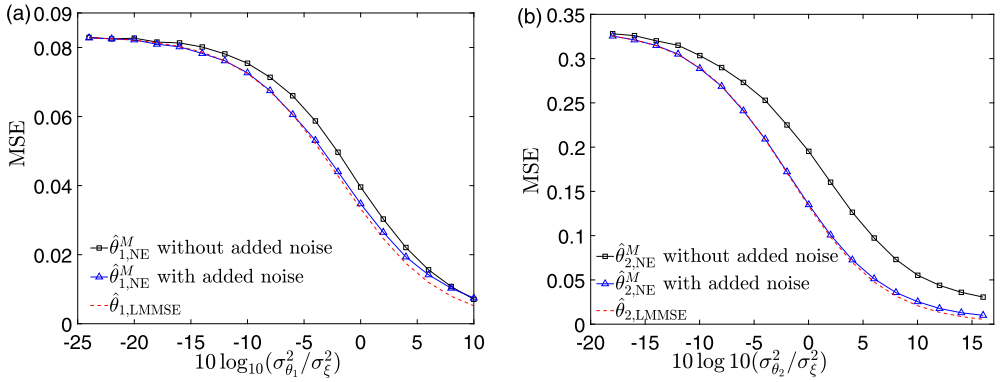

Fig. 9. Bayesian MSEs of the designed noise-enhanced estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ in Eq. (11) with the optimal added Gaussian noise, $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^M$ without added noise and the LMMSE estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{LMMSE}}$ in Eq. (19) as a function of the input SNR $\sigma_\theta^2/\sigma_\xi^2$ for (a) the parameter θ_1 and (b) the parameter θ_2. The other parameters are the same as in Fig. 8.

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

Fabing 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 $g_{1m}(\cdot) = g_1(\cdot)$ of sensors in the first network node. Letting $\boldsymbol{\omega} = [\boldsymbol{\omega}_1^\top, \widetilde{\boldsymbol{\omega}}^\top]$ ($\widetilde{\boldsymbol{\omega}} = [\boldsymbol{\omega}_2^\top, \boldsymbol{\omega}_3^\top, \ldots, \boldsymbol{\omega}_N^\top]^\top$) and $\boldsymbol{y} = [\boldsymbol{y}_1^\top, \widetilde{\boldsymbol{y}}^\top]$ ($\widetilde{\boldsymbol{y}} = [\boldsymbol{y}_2^\top, \boldsymbol{y}_3^\top, \ldots, \boldsymbol{y}_N^\top]^\top$), the covariance matrix $\boldsymbol{C}_{\boldsymbol{zz}}$ can be partitioned as

$$\boldsymbol{C}_{\boldsymbol{zz}} = \mathrm{E}_{\boldsymbol{x},\eta}\{[\boldsymbol{z} - \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})][\boldsymbol{z} - \mathrm{E}_{\boldsymbol{x},\eta}(\boldsymbol{z})]^\top\} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1} & \boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}} \\ \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\boldsymbol{y}_1} & \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}} \end{bmatrix}, \quad \text{(A.1)}$$

where the partitioned matrices $\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1} = \mathrm{E}_{x_1,\eta}\{[\boldsymbol{y}_1 - \mathrm{E}_{x_1,\eta}(\boldsymbol{y_1})][\boldsymbol{y}_1 - \mathrm{E}_{x_1,\eta}(\boldsymbol{y_1})]^\top\}$, $\boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}} = \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\boldsymbol{y}_1}^\top = \mathrm{E}_{\boldsymbol{x},\eta}\{[\boldsymbol{y}_1 - \mathrm{E}_{x_1,\eta}(\boldsymbol{y_1})][\widetilde{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\widetilde{\boldsymbol{y}})]^\top\}$ and $\boldsymbol{C}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}} = \mathrm{E}_{\boldsymbol{x},\eta}\{[\widetilde{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\widetilde{\boldsymbol{y}})][\widetilde{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\widetilde{\boldsymbol{y}})]^\top\}$. Similarly, the correlation vector $\boldsymbol{C}_{\boldsymbol{z}\theta_i}$ can be partitioned as

$$\boldsymbol{C}_{\boldsymbol{z}\theta_i} = [\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}^\top, \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\theta_i}^\top]^\top \quad \text{(A.2)}$$

with the partitioned correlation vector $\boldsymbol{C}_{\boldsymbol{y}_1\theta_i} = \mathrm{E}_{x_1,\eta}\{[\boldsymbol{y}_1 - \mathrm{E}_{x_1,\eta}(\boldsymbol{y_1})][\theta_i - \mathrm{E}_{\theta_i}(\theta_i)]\}$ and $\boldsymbol{C}_{\widetilde{\boldsymbol{y}}\theta_i} = \mathrm{E}_{\boldsymbol{x},\eta}\{[\widetilde{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\widetilde{\boldsymbol{y}})][\theta_i - \mathrm{E}_{\theta_i}(\theta_i)]\}$.

Since the optimum weight vector $\boldsymbol{\omega}^{\mathrm{opt}} = \boldsymbol{C}_{\boldsymbol{zz}}^{-1}\boldsymbol{C}_{\boldsymbol{z}\theta_i}$ can be rewritten as

$$\begin{bmatrix} \boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1} & \boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}} \\ \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\boldsymbol{y}_1} & \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\widetilde{\boldsymbol{y}}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega}_1 \\ \widetilde{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{y}_1\theta_i} \\ \boldsymbol{C}_{\widetilde{\boldsymbol{y}}\theta_i} \end{bmatrix} \quad \text{(A.3)}$$

thus we find

$$\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}\boldsymbol{\omega}_1 + \boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}}\widetilde{\boldsymbol{\omega}} = \boldsymbol{C}_{\boldsymbol{y}_1\theta_i}. \quad \text{(A.4)}$$

For M identical transfer functions $g_{1m}(\cdot) = g_1(\cdot)$ in the first network node and the identically distributed noise components η_{1m}, the covariance matrix $\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}$ has M equivalent diagonal elements

$$[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mm} = \mathrm{E}_{x_1,\eta}[g_1^2(x_1+\eta)] - \mathrm{E}_{x_1,\eta}^2[g_1(x_1+\eta)] \quad \text{(A.5)}$$

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

$$[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mj} = \mathrm{E}_{x_1}\{\mathrm{E}_\eta^2[g_1(x_1+\eta)]\} - \mathrm{E}_{x_1,\eta}^2[g_1(x_1+\eta_n)] \quad \text{(A.6)}$$

for $m, j = 1, 2, \ldots, M$ $(m \neq j)$. Its inverse matrix $\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}$ also has M equivalent diagonal elements $[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}]_{mm}$ and $M(M-1)$ equivalent non-diagonal elements $[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}]_{mj}$. Therefore, the equality $\mathbf{1}_M^\top \boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}\mathbf{1}_M = \mathbf{1}_M^\top\mathbf{1}_M = M$ yields

$$[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}]_{mm} + (M-1)[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}]_{mj} = \left([\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mm} + (M-1)[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mj}\right)^{-1}. \quad \text{(A.7)}$$

The correlation vector $\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}$ also has M equivalent elements

$$[\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}]_m = \mathrm{E}_{x_1,\eta}(\widetilde{y}_{1m}\widetilde{\theta}_i) = \mathrm{E}_{x_1}\{\theta_i \mathrm{E}_\eta[g_1(x_1+\eta)]\} - \mathrm{E}_{\theta_i}(\theta_i)\mathrm{E}_{x_1}\{\mathrm{E}_\eta[g_1(x_1+\eta)]\}. \quad \text{(A.8)}$$

Similarly, for M identical transfer functions $g_{1m}(\cdot) = g_1(\cdot)$, the $M \times (N-1)M$ partitioned matrix $\boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}}$ has M identical row vectors, which can be expressed as $\boldsymbol{d} = \mathrm{E}_{\boldsymbol{x},\eta}\{[g_1(x_1+\eta) - \mathrm{E}_{x_1,\eta}(g_1(x_1+\eta))][\widetilde{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\widetilde{\boldsymbol{y}})]^\top\}$. Therefore, the vector $\boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}}\widetilde{\boldsymbol{\omega}}$ in Eq. (A.4) can be written as $\alpha\mathbf{1}_M$ with $\alpha = \boldsymbol{d}\widetilde{\boldsymbol{\omega}}$.

Thus, from Eq. (A.4), we find $\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}\boldsymbol{\omega}_1 = \boldsymbol{C}_{\boldsymbol{y}_1\theta_i} - \boldsymbol{C}_{\boldsymbol{y}_1\widetilde{\boldsymbol{y}}}\widetilde{\boldsymbol{\omega}} = ([\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}]_m - \alpha)\mathbf{1}_M$ and the optimum weight vector $\boldsymbol{\omega}_1^{\mathrm{opt}}$

$$\boldsymbol{\omega}_1^{\mathrm{opt}} = \boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}^{-1}([\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}]_m - \alpha)\mathbf{1}_M = w_1^{\mathrm{opt}}\mathbf{1}_M, \quad \text{(A.9)}$$

with

$$w_1^{\mathrm{opt}} = \frac{[\boldsymbol{C}_{\boldsymbol{y}_1\theta_i}]_m - \alpha}{[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mm} + (M-1)[\boldsymbol{C}_{\boldsymbol{y}_1\boldsymbol{y}_1}]_{mj}}. \quad \text{(A.10)}$$

For any of network node with M identical sensors, Eq. (A.10) holds.

Appendix B. Proof of Corollary 1

For M identical transfer functions $g_{nm}(\cdot) = g_n(\cdot)$ of sensors in the n-th network node, each node has a scalar output $\overline{y}_n$ as shown in Fig. 2. Based on N outputs $\overline{y}_n$, the intermediate linear estimator can be expressed as $\widehat{\theta}_i^M = w_0 + \boldsymbol{w}^\top \overline{\boldsymbol{y}}$, where w_0 is the biasing weight. In the same way, setting the derivatives $\partial \mathrm{E}_{\boldsymbol{x},\eta}[(\theta_i - \widehat{\theta}_i^M)^2]/\partial w_0 = 0$ and $\partial \mathrm{E}_{\boldsymbol{x},\eta}[(\theta_i - \widehat{\theta}_i^M)^2]/\partial \boldsymbol{w} = 0$, we find the optimum bias $w_0 = \mathrm{E}_{\theta_i}(\theta_i) - \boldsymbol{w}^\top \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})$ and the optimal weight vector $\boldsymbol{w}^{\mathrm{opt}} = \boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}^{-1} \boldsymbol{C}_{\overline{\boldsymbol{y}}\boldsymbol{\theta}}$. Substituting w_0 and $\boldsymbol{w}^{\mathrm{opt}}$ into $\widehat{\theta}_i^M = w_0 + \boldsymbol{w}^\top \overline{\boldsymbol{y}}$, we find the designed estimator $\widehat{\theta}_i^M$ given by Eq. (8).

More specifically, the expectation of the node output $\overline{y}_n$ in Eq. (7) can be computed as

$$\begin{aligned}\mathrm{E}_{x_n,\eta}(\overline{y}_n) &= \frac{1}{M}\sum_{m=1}^{M} \mathrm{E}_{x_n,\eta}[g_n(x_n+\eta_{nm})] \\ &= \mathrm{E}_{x_n}\{\mathrm{E}_\eta[g_n(x_n+\eta)]\},\end{aligned} \tag{B.1}$$

because the added noise components η_{nm} have the common PDF f_η. Therefore, the correlation vector $\boldsymbol{C}_{\theta_i\overline{\boldsymbol{y}}} = \boldsymbol{C}_{\overline{\boldsymbol{y}}\theta_i}^\top = \mathrm{E}_{\boldsymbol{x},\eta}\{[\theta_i - \mathrm{E}_{\theta_i}(\theta_i)][\overline{\boldsymbol{y}} - \mathrm{E}_{\boldsymbol{x},\eta}(\overline{\boldsymbol{y}})]^\top\}$ has elements

$$\begin{aligned}[\boldsymbol{C}_{\overline{\boldsymbol{y}}\theta_i}]_n &= \mathrm{E}_{x_n,\eta}\{[\theta_i - \mathrm{E}_{\theta_i}(\theta_i)][\overline{y}_n - \mathrm{E}_{x_n,\eta}(\overline{y}_n)]\} \\ &= \mathrm{E}_{x_n}\{\theta_i \mathrm{E}_\eta[g_n(x_n+\eta)]\} - \mathrm{E}_{\theta_i}(\theta_i)\mathrm{E}_{x_n}\{\mathrm{E}_\eta[g_n(x_n+\eta)]\},\end{aligned} \tag{B.2}$$

which is independent of the sensor number M. The covariance matrix $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}$ has N diagonal elements

$$\begin{aligned}[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{nn} &= \mathrm{E}_{x_n,\eta}(\overline{y}_n^2) - \mathrm{E}_{x_n,\eta}^2(\overline{y}_n) \\ &= \mathrm{E}_{x_n,\eta}\Big[\Big(\frac{1}{M}\sum_{m=1}^{M} g_n(x_n+\eta_{nm})\Big)^2\Big] - \mathrm{E}_{x_n,\eta}^2[g_n(x_n+\eta)] \\ &= \frac{1}{M^2}\Big\{\sum_{m=1}^{M} \mathrm{E}_{x_n,\eta}[g_n^2(x_n+\eta_{nm})] \\ &\quad + \mathrm{E}_{x_n}\Big(\sum_{m=1}^{M} \mathrm{E}_\eta[g_n(x_n+\eta_{nm})] \\ &\quad \times \sum_{j=1}^{M} \mathrm{E}_\eta[g_n(x_n+\eta_{nj})]\Big)\Big\}(m\neq j)\Big\} - \mathrm{E}_{x_n,\eta}^2[g_n(x_n+\eta)] \\ &= \frac{1}{M}\mathrm{E}_{x_n}\{\mathrm{E}_\eta[g_n^2(x_n+\eta)]\} + \frac{M-1}{M}\mathrm{E}_{x_n}\{\mathrm{E}_\eta^2[g_n(x_n+\eta)]\} \\ &\quad - \mathrm{E}_{x_n}^2\{\mathrm{E}_\eta[g_n(x_n+\eta)]\},\end{aligned} \tag{B.3}$$

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

$$\begin{aligned}[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{n\ell} &= \mathrm{E}_{x_n,x_\ell,\eta}(\overline{y}_n\overline{y}_\ell) - \mathrm{E}_{x_n,\eta}(\overline{y}_n)\mathrm{E}_{x_\ell,\eta}(\overline{y}_\ell)(n\neq\ell) \\ &= \frac{1}{M^2}\mathrm{E}_{x_n,x_\ell}\Big(\sum_{m=1}^{M} \mathrm{E}_\eta[g_n(x_n+\eta_{nm})]\sum_{j=1}^{M} \mathrm{E}_\eta[g_\ell(x_\ell+\eta_{\ell j})]\Big)\Big\} \\ &\quad - \mathrm{E}_{x_n,\eta}[g_n(x_n+\eta)]\mathrm{E}_{x_\ell,\eta}[g_\ell(x_\ell+\eta)] \\ &= \mathrm{E}_{x_n,x_\ell}\{\mathrm{E}_\eta[g_n(x_n+\eta)]\mathrm{E}_\eta[g_\ell(x_\ell+\eta)]\} \\ &\quad - \mathrm{E}_{x_n}\{\mathrm{E}_\eta[g_n(x_n+\eta)]\}\mathrm{E}_{x_\ell}\{\mathrm{E}_\eta[g_\ell(x_\ell+\eta)]\}.\end{aligned} \tag{B.4}$$

It is noted that the $N(N-1)$ non-diagonal elements $[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{n\ell}$ are also independent of the sensor number M. With Eqs. (B.1)–(B.4), the p designed estimator $\widehat{\theta}_i^M$ 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

$$\begin{aligned}\lim_{M\to\infty} \overline{y}_n &= \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g_n(x_n+\eta_{nm}) = \mathrm{E}_\eta[g_n(x_n+\eta)] \\ &= g_n^*(x_n),\end{aligned} \tag{C.1}$$

and the diagonal elements $[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{nn}$ in Eq. (B.3) of $\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}$ converge to

$$\begin{aligned}\lim_{M\to\infty} [\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{nn} &= \mathrm{E}_{x_n}\{\mathrm{E}_\eta^2[g_n(x_n+\eta)]\} - \mathrm{E}_{x_n}^2\{\mathrm{E}_\eta[g_n(x_n+\eta)]\} \\ &= \mathrm{E}_{x_n}[g_n^{*2}(x_n)] - \mathrm{E}_{x_n}^2[g_n^*(x_n)] \\ &= [\boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}]_{nn},\end{aligned} \tag{C.2}$$

because the term $\mathrm{E}_{x_n}\{\mathrm{E}_\eta[g_n^2(x_n+\eta)]\} < \infty$ in Eq. (B.3). It is also noted that, using Eq. (C.1), Eq. (B.2) and Eq. (B.4), we find that the elements $[\boldsymbol{C}_{\theta_i\overline{\boldsymbol{y}}}]_n = \mathrm{E}_{x_n}[\theta_i g_n^*(x_n)] - \mathrm{E}_{\theta_i}(\theta_i)\mathrm{E}_{x_n}[g_n^*(x_n)] = [\boldsymbol{C}_{\boldsymbol{\theta}\boldsymbol{g}^*}]_n$ and the non-diagonal elements $[\boldsymbol{C}_{\overline{\boldsymbol{y}}\overline{\boldsymbol{y}}}]_{n\ell} = \mathrm{E}_{x_n,x_\ell}[g_n^*(x_n)g_\ell^*(x_\ell)] - \mathrm{E}_{x_n}[g_n^*(x_n)]\mathrm{E}_{x_\ell}[g_\ell^*(x_\ell)] = [\boldsymbol{C}_{\boldsymbol{g}^*\boldsymbol{g}^*}]_{n\ell}$.

Substituting Eqs. (C.1) and (C.2) into Eq. (8), we have the explicit form of the estimator $\widehat{\theta}_i^*$ in Eq. (12) and its Bayesian MSE $B_{\widehat{\theta}_i^*}$ in Eq. (13). Combining p estimators of $\widehat{\theta}_i^*$ in Eq. (12), we have the vector estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{NE}}^*$ in Eq. (14) and its Bayesian MSE matrix $\boldsymbol{B}_{\mathrm{NE}}^*$ in Eq. (15).

References

[1] O. Dabeer, A. Karnik, Signal parameter estimation using 1-bit dithered quantization, IEEE Trans. Inf. Theory 52 (12) (2006) 5389–5405.
[2] J. Xiao, R. Ribeiro, Z. Luo, G. Giannakis, Distributed compression-estimation using wireless sensor networks, IEEE Signal Process. Mag. 23 (4) (2006) 27–41.
[3] M. Shieazi, A. Vosoughi, On Bayesian Fisher information maximization for distributed vector estimation, IEEE Trans. Signal Inf. Process. Netw. 5 (4) (2019) 628–645.
[4] H. Chen, P.K. Varshney, Performance limit for distributed estimation systems with identical one-bit quantizers, IEEE Trans. Signal Process. 58 (1) (2010) 466–471.
[5] G. Zeitler, G. Kramer, A. Singer, Bayesian parameter estimation using single-bit dithered quantization, IEEE Trans. Signal Process. 60 (6) (2012) 2713–2716.
[6] F. Chapeau-Blondeau, S. Blanchard, D. Rousseau, Fisher information and noise-aided power estimation from one-bit quantizers, Digit. Signal Process. 18 (3) (2008) 434–443.
[7] H.C. Papadopoulos, G.W. Wornell, A.V. Oppenheim, Sequential signal encoding from noisy measurements using quantizers with dynamic bias control, IEEE Trans. Inf. Theory 47 (3) (2001) 978–1002.
[8] M. Li, J. Klejsa, W.B. Kleijn, On distribution preserving quantization, IEEE Signal Process. Lett. 17 (12) (2011) 1014–1017.
[9] D. Simon, J. Sulam, Y. Romano, Y. Lu, M. Elad, MMSE approximation for sparse coding algorithms using stochastic resonance, IEEE Trans. Signal Process. 67 (17) (2019) 4597–4610.
[10] A. Sanzi, A. Vosoughi, Distributed vector estimation for power- and bandwidth-constrained wireless sensor networks, IEEE Trans. Signal Process. 64 (15) (2016) 3879–3894.
[11] H. Chen, P.K. Varshney, Nonparametric one-bit quantizers for distributed estimation, IEEE Trans. Signal Process. 58 (7) (2010) 3777–3787.
[12] A. Gershman, J. Böhme, A pseudo-noise approach to direction finding, Signal Process. 71 (1) (1998) 1–13.
[13] O. Dabeer, E. Masry, Multivariate signal parameter estimation under dependent noise from 1-bit dithered quantized data, IEEE Trans. Inf. Theory 54 (4) (2008) 1637–1654.
[14] S. Engelberg, An introduction to dither, IEEE Instrum. Meas. Mag. 15 (6) (2012) 50–54.
[15] J. Liu, F. Duan, F. Chapeau-Blondeau, D. Abbott, Exploring Bona fide optimal noise for Bayesian parameter estimation, IEEE Access 8 (2020) 18822–18831.
[16] R. Benzi, A. Sutera, A. Vulpiani, The mechanism of stochastic resonance, J. Phys. A, Math. Gen. 14 (11) (1981) L453–L457.
[17] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance, Rev. Mod. Phys. 70 (1) (1998) 233–287.
[18] N.G. Stocks, Suprathreshold stochastic resonance in multilevel threshold systems, Phys. Rev. Lett. 84 (11) (2000) 2310–2313.

[19] M.D. McDonnell, N.G. Stocks, C.E.M. Pearce, D. Abbott, Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantization, Cambridge University Press, New York, 2008.
[20] Q. Zhai, Y. Wang, Noise effect on signal quantization in an array of binary quantizers, Signal Process. 152 (11) (2018) 265–272.
[21] S. Kay, Can detectability be improved by adding noise?, IEEE Signal Process. Lett. 7 (1) (2000) 8–10.
[22] H. Chen, P.K. Varshney, S.M. Kay, J.H. Michels, Theory of the stochastic resonance effect in signal detection: part I–fixed detectors, IEEE Trans. Signal Process. 55 (7) (2007) 3172–3184.
[23] H. Chen, P.K. Varshney, Theory of the stochastic resonance effect in signal detection–part II: variable detectors, IEEE Trans. Signal Process. 56 (10) (2008) 5031–5041.
[24] S. Lu, Q. He, F. Kong, Effects of underdamped step-varying second-order stochastic resonance for weak signal detection, Digit. Signal Process. 36 (1) (2015) 93–103.
[25] F. Chapeau-Blondeau, D. Rousseau, Noise-enhanced performance for an optimal Bayesian estimator, IEEE Trans. Signal Process. 52 (5) (2004) 1327–1334.
[26] H. Chen, P.K. Varshney, J.H. Michels, Noise enhanced parameter estimation, IEEE Trans. Signal Process. 56 (10) (2008) 5074–5081.
[27] J. Zhu, X. Li, R.S. Blum, Y. Gu, Parameter estimation from quantized observations in multiplicative noise environments, IEEE Trans. Signal Process. 63 (15) (2015) 4037–4050.
[28] A. Sani, A. Vosoughi, Noise enhanced distributed Bayesian estimation, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017, pp. 4217–4221.
[29] A. Patel, B. Kosko, Optimal mean-square noise benefits in quantizer-array linear estimation, IEEE Signal Process. Lett. 17 (12) (2010) 1005–1009.
[30] T. Yang, S. Liu, H. Liu, M. Tang, X. Tan, X. Zhou, Noise benefits parameter estimation in LMMSE sense, Digit. Signal Process. 73 (2) (2018) 153–163.
[31] G.O. Balkan, S. Gezici, CRLB based optimal noise enhanced parameter estimation using quantized observations, IEEE Signal Process. Lett. 17 (5) (2010) 477–480.
[32] S. Uhlich, Bayes risk reduction of estimators using artificial observation noise, IEEE Trans. Signal Process. 63 (20) (2015) 5535–5545.
[33] S. Ikemoto, F. Dallalibera, K. Hosoda, Noise-modulated neural networks as an application of stochastic resonance, Neurocomputing 277 (2) (2017) 29–37.
[34] F. Duan, Y. Pan, F. Chapeau-Blondeau, D. Abbott, Noise benefits in combined nonlinear Bayesian estimators, IEEE Trans. Signal Process. 67 (17) (2019) 4611–4623.
[35] X. Liu, L. Duan, F. Duan, F. Chapeau-Blondeau, D. Abbott, Enhancing threshold neural network via suprathreshold stochastic resonance for pattern classification, Phys. Lett. A 403 (2021) 127387.
[36] A. Ribeiro, G. Giannakis, Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case, IEEE Trans. Signal Process. 54 (3) (2006) 1131–1143.
[37] A. Ribeiro, G. Giannakis, Bandwidth-constrained distributed estimation for wireless sensor networks-part II: unknown probability density function, IEEE Trans. Signal Process. 54 (7) (2006) 2784–2796.
[38] K. You, Recursive algorithms for parameter estimation with adaptive quantizer, Automatica 52 (2015) 192–201.
[39] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, vol. 1, Prentice-Hall, Upper Saddle River, New Jersey, 1998.
[40] H.L. van Trees, K.L. Bell, Z. Tan, Detection, Estimation and Modulation Theory–Part I: Detection, Estimation, and Filtering Theory, Wiley-Interscience, 2002. Wiley, River Street, Hoboken, New Jersey, 2013.
[41] D.K.W. Ho, B.D. Rao, Antithetic dithered 1-bit massive MIMO architecture: Efficient channel estimation via parameter expansion and PML, IEEE Trans. Signal Process. 67 (9) (2019) 2291–2301.
[42] H. Chen, L.R. Varshney, P.K. Varshney, Noise-enhanced information systems, Proc. IEEE 102 (10) (2014) 1607–1621.
[43] S. Bayram, S. Gezici, H.V. Poor, Noise enhanced hypothesis-testing in the restricted Bayesian framework, IEEE Trans. Signal Process. 58 (8) (2010) 3972–3989.
[44] J. Nocedal, S.J. Wright, Numerical Optimization, Springer–Verlag, New York, 2006.
[45] A. Sani, A. Vosoughi, On distributed linear estimation with observation model uncertainties, IEEE Trans. Signal Process. 66 (12) (2018) 3212–3227.

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 ultrasonics. 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.

Derek Abbott was born in South Kensington, London, U.K., in 1960. He received the B.Sc. (Hons.) degree in physics from Loughborough University, Leicestershire, U.K., in 1982 and the Ph.D. degree in electrical and electronic engineering from The University of Adelaide, Adelaide, SA, Australia, in 1995, under K. Eshraghian and B. R. Davis.

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*, *Microelectronics 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.