IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 16, 2019, accepted January 15, 2020, date of publication January 21, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968420

Exploring *Bona Fide* Optimal Noise for Bayesian Parameter Estimation

JIA LIU[1], FABING DUAN[1], FRANÇOIS CHAPEAU-BLONDEAU[2], AND DEREK ABBOTT[3], (Fellow, IEEE)
[1]Institute of Complexity Science, Qingdao University, Qingdao 266071, China
[2]Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), University of Angers, 49000 Angers, France
[3]Centre for Biomedical Engineering (CBME), School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

Corresponding author: Fabing Duan (fabing.duan@hotmail.com)

This work was supported by the National Natural Science Foundation of China under Grant 61573202.

ABSTRACT In this paper, we investigate the benefit of intentionally added noise to observed data in various scenarios of Bayesian parameter estimation. For optimal estimators, we theoretically demonstrate that the Bayesian Cramér-Rao bound for the case with added noise is never smaller than for the original data, and the updated minimum mean-square error (MSE) estimator performs no better. This motivates us to explore the feasibility of noise benefit in some useful suboptimal estimators. Several Bayesian estimators established from one-bit-quantizer sensors are considered, and for different types of pre-existing background noise, optimal distributions are determined for the added noise in order to improve the performance in estimation. With a single sensor, it is shown that the optimal added noise for reducing the MSE is actually a constant bias. However, with parallel arrays of such sensors, *bona fide* optimal added noise, no longer a constant bias, is shown to reduce the MSE. Moreover, it is found that the designed Bayesian estimators can benefit from the optimal added noise to effectively approach the performance of the minimum MSE estimator, even when the assembled sensors possess different quantization thresholds.

INDEX TERMS Bayesian estimator, Bayesian information, *bona fide* optimal noise, noise benefit, parameter estimation.

I. INTRODUCTION

Adding some random noise to a signal before quantization has been shown to be beneficial for analog to digital converters resulting in smaller signal distortion and wide system dynamic range [1]–[3]. The technique of adding noise or *dithering* was perhaps the first that recognized a beneficial role for noise in a signal processing context [1]–[5]. Then the term *stochastic resonance* was initially coined to describe the possible mechanism for maximizing the response of a bistable system to a small periodic force by optimizing the noise intensity to a non-zero level [6]. Stochastic resonance attracted much attention in physics and biology [7]–[17] soon afterwards. Gammaitoni [1] first pointed out that stochastic resonance, far from being limited to a resonant phenomenon, can also be interpreted as a special case of dithering and is related to the notion of noise-induced threshold crossings. Similarly, Collins *et al*. [18], [19] coined the term of *aperiodic stochastic resonance* for characterizing the noise-induced behavior in excitable systems with aperiodic inputs, and Stocks [20] defined the *suprathreshold stochastic resonance* using Shannon's average mutual information measure between the input and the output of a summing network of threshold devices. These widened concepts of stochastic resonance that are closely relate to the field of statistical signal processing, are now widely referred to as noise enhancement or noise benefit [21]–[50].

There are two main situations whereby the noise benefit has been exploited in signal estimation: one is implementing suboptimal estimators in practical estimation problems to avoid too complex or intractable optimal estimators in general [34], [51]. The other is estimating a signal from observed data of a number of low-cost sensors (e.g. quantizers) in a fusion center. These sensors with a few bits are often deployed over a sensing field to compose wireless sensor networks in distributed estimation problems [38], [52], [53]. For the first situation, the performances of some easily implemented suboptimal estimators were shown to be substantially

The associate editor coordinating the review of this manuscript and approving it for publication was Cesar Vargas-Rosales.

improved by exploiting the benefits of added noise [22], [32], [34]–[38], [42], [46]–[49]. In the second situation, rich results from utilizing various kinds of noise have been reported for quantized observations [1], [2], [12]–[14], [20], [21], [23]–[25], [27], [30], [31], [33], [36]–[43], [45]–[49]. For instance, Papadopoulos *et al*. developed a methodology of additive control input before signal quantization at the sensor to achieve the maximum possible performance for quantizer-based networks [23]. They also noticed the option of using feedback from past observations for efficient estimators in terms of mean-square error (MSE) [23]. Modeling the suprathreshold stochastic resonance as stochastic quantization, McDonnell *et al*. systematically studied the optimal linear and nonlinear decoding schemes associated with the information bound on the MSE [12]. The optimal Bayesian estimators constructed by the quantizer outputs were also explicitly derived, and a basic mechanism was provided for the performance improvement of optimal Bayesian estimator by increasing the noise level [24], [25], [30], [31].

Since the addition of noise can be artificially designed, then finding the optimal probability density function (PDF) of added noise becomes an interesting question [26], [28], [29], [32], [33], [36]–[43], [45]–[49]. Especially, Chen *et al*. considered all possible PDFs of added noise to optimize an arbitrary fixed or variable estimator and proved that the optimal noise, if it exists, is just a finite number of (no more than two) constant vectors by using the properties of convex hull and Caratheodory theorem [28], [29], [32], [38], [39]. Then, this kind of optimal noise PDFs inspired a series of theoretical improvability of estimation under various estimation criteria [26], [33], [37], [40]–[43], [45]–[49]. An interesting question is whether optimal *bona fide* noise, rather than a constant bias, exists for enhancing the estimator performance or not. Interestingly, Uhlich [42] proposed a new estimator constructed by bagged estimators that are modified by mutually independent noise samples, and derived the necessary and sufficient conditions for the existence of the optimal noise. Uhlich [42] also found that the optimal noise PDF, not limited to the noise type revealed in [28], [29], [32], [38], [39], has non-trivial complicated shapes. For minimizing the MSE of a combiner of identical estimators, we also found that solving the optimal noise PDF is a constrained nonlinear functional optimization problem, and approximate optimal PDFs of the optimal noise are also found to be complicated [46], [48].

Although many important results for noise benefits in estimators have been obtained, there are still some unsolved questions. For instance, it is known that, for random parameter estimation, a lower bound on the MSE of any estimator is called the Bayesian Cramér-Rao bound (BCRB) that is directly calculated from the primary observations [54]. Then, two interesting questions need to be addressed: after artificially injecting noise into the primary observed data, resulting updated data—so, does the corresponding new BCRB calculated on the updated data increase or decrease? Can the minimum MSE (MMSE) estimator deduced from the updated data achieve a lower MSE than that of the original MMSE estimator based on the primary observations?

In this paper, we will theoretically provide the solutions to aforementioned crucial questions, and elucidate the possibility of exploiting the noise benefits in some easily implemented suboptimal estimators. We argue that the noise-enhanced Bayesian estimators proposed by [22]–[43], [45]–[49] in recent years can be mainly classified into four categories: (i) the noise-modified estimator established on a single sensor [32], (ii) a linear minimum MSE (LMMSE) estimator based on a single sensor, (iii) the noise-enhanced Bayesian estimator as the average of outputs of an ensemble of identical sensors [42] and (iv) the linear combination estimator executing the LMMSE transform on an array of identical or nonidentical sensors [48].

For a noise-modified estimator, it was proved that the optimal added noise is just a constant bias [32]. However, the optimized MSE achieved by the noise-modified estimator has a long way to catch up the MSE of the original MMSE estimator. In order to provide useful parameter estimation, it is useful to incorporate both the statistical properties of the original background noise as well as the prior knowledge of the random parameter. This design principle leads to the LMMSE estimator with adaptively adjustable weights that depend only on the first two moments of the joint PDF. We demonstrate that, based on a single sensor, the LMMSE estimator can obtain a lower MSE than the noise-modified estimator does, but the minimum MSE achieved by the LMMSE estimator is still larger than that of the MMSE estimator. Moreover, the optimal added noise is still a constant bias for minimizing the MSE of the LMMSE estimator, and not *bona fide* random noise.

Furthermore, based on a sufficiently large number of identical sensors, it is shown that the noise-enhanced estimator can efficiently approach the MMSE estimator by the *bona fide* optimal noise that is not restricted to a constant bias. However, the noise-enhanced estimator is inapplicable to an ensemble of nonidentical sensors. For the general case of nonidentical sensors, we theoretically demonstrate that the linear combination estimator always outperforms the noise-enhanced estimator, and is able to perform as efficiently as the MMSE estimator. From observations of one-bit-quantizer sensors, we illustratively confirm the aforementioned conclusions of the performance comparison between four considered estimators. The complicated PDFs of *bona fide* optimal noise for the linear combination estimator are also presented. These interesting results of mutually independent added noise components in sensors manifest their potential benefits to the parameter estimation problems.

II. PARAMETER ESTIMATION MODEL AND PROBLEM FORMULATION

Consider a parameter estimation scenario with the scalar observation

$$x_n = s(\theta) + \xi_n, \tag{1}$$

where $s(\theta)$ is a function of an unknown random parameter θ with a prior PDF f_θ, and the mutually independent samples ξ_n of background noise, uncorrelated with θ, are with zero mean and common PDF f_ξ for $n = 1, 2, \cdots, N$. Letting $\mathbf{x} = [x_1, x_2, \cdots, x_N]^\top$, the statistical characteristic of observation data can be described by the joint PDF $f_{\mathbf{x},\theta}(\mathbf{x}, \theta) = f_{\mathbf{x}|\theta}(\mathbf{x}|\theta)f_\theta(\theta) = f_\xi(\mathbf{x}-s(\theta))f_\theta(\theta)$. Here, $f_{\mathbf{x}|\theta}$ is the conditional PDF. It is well known [54] that the MSE $\mathcal{R}$ of any estimator $\hat{\theta}(\mathbf{x})$ satisfies the inequality

$$\mathcal{R} = \mathrm{E}_{\mathbf{x},\theta}[(\hat{\theta}(\mathbf{x}) - \theta)^2] = \mathrm{E}_{\mathbf{x},\theta}[\varepsilon^2] \geq J_\mathrm{B}^{-1} = \mathrm{BCRB}, \quad (2)$$

where the error of estimator is $\varepsilon = \hat{\theta}(\mathbf{x}) - \theta$, and the Bayesian information J_B is defined as

$$J_\mathrm{B} = \mathrm{E}_{\mathbf{x},\theta}\Big[\Big(\frac{\partial \ln f_{\mathbf{x},\theta}(\mathbf{x}, \theta)}{\partial\theta}\Big)^2\Big] = \mathrm{E}_\theta[J_\mathrm{F}(\theta)] + J_\theta \quad (3)$$

with the prior Fisher information $J_\theta = \mathrm{E}_\theta[(\partial \ln f_\theta(\theta)/\partial\theta)^2]$ of the prior PDF f_θ and the Fisher information $J_\mathrm{F}(\theta) = \mathrm{E}_{\mathbf{x}|\theta}[(\partial \ln f_{\mathbf{x}|\theta}(\mathbf{x}|\theta)/\partial\theta)^2]$ of observation data $\mathbf{x}$ with respect to the parameter θ [54]. This lower bound of J_B^{-1} in (2) on the MSE $\mathcal{R}$ of any estimator is also called BCRB [54]. Here, $\mathrm{E}_{\mathbf{x},\theta}(\cdot)$, $\mathrm{E}_{\mathbf{x}|\theta}(\cdot)$ and $\mathrm{E}_\theta(\cdot)$ denote expectations with respect to the joint PDF $f_{\mathbf{x},\theta}$, the conditional PDF $f_{\mathbf{x}|\theta}$ and the prior PDF f_θ, respectively.

Theorem 1: After the injection of added noise η_n into x_n, the updated observation data $\bar{x}_n = x_n + \eta_n = s(\theta) + \xi_n + \eta_n = s(\theta) + z_n$, and the updated BCRB is not less than the original one.

Proof: Letting j_F be the Fisher information of one sample x_n, we have

$$\begin{aligned} j_\mathrm{F}(\theta) &= \mathrm{E}_{x|\theta}\Big[\Big(\frac{\partial \ln f_{x|\theta}(x|\theta)}{\partial\theta}\Big)^2\Big] \\ &= \mathrm{E}_{x|\theta}\Big[\Big(\frac{\partial \ln f_\xi(x - s(\theta))}{\partial\theta}\Big)^2\Big] \\ &= \mathrm{E}_\xi\Big[\Big(\frac{\partial \ln f_\xi(x)}{\partial x}\Big)^2\Big]\Big(\frac{\partial s}{\partial\theta}\Big)^2 = j_\xi\Big(\frac{\partial s}{\partial\theta}\Big)^2 \end{aligned} \quad (4)$$

with the Fisher information $j_\xi = \mathrm{E}_\xi[(\partial \ln f_\xi(x)/\partial x)^2]$ of the PDF f_ξ. Then, for the independent identically distributed (i.i.d.) noise samples ξ_n, the Fisher information of the observation vector $\mathbf{x}$ is $J_\mathrm{F}(\theta) = N j_\mathrm{F}(\theta)$. Similarly, the Fisher information of the updated data $\bar{x}_n$ can be expressed as $\bar{j}_\mathrm{F}(\theta) = j_z(\partial s/\partial\theta)^2$ with the Fisher information $j_z = \mathrm{E}_z[(\partial \ln f_z(z)/\partial z)^2]$ of the PDF $f_z(z) = \int f_\xi(z - \eta)f_\eta(\eta)d\eta$ of the composite noise z_n. Since the Fisher information quantities $j_z, j_\xi, j_\eta > 0$ satisfy the convolution inequality $j_z^{-1} \geq j_\xi^{-1} + j_\eta^{-1}$ [55], resulting in

$$j_z \leq j_\xi\Big(\frac{j_\eta}{j_\xi + j_\eta}\Big) \leq j_\xi. \quad (5)$$

Thus, for i.i.d. noise samples z_n, the Fisher information of the updated data $\bar{\mathbf{x}} = [\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_N]^\top$ becomes $\bar{J}_\mathrm{F}(\theta) = N\bar{j}_\mathrm{F}(\theta) = N j_z(\partial s/\partial\theta)^2$ and the updated Bayesian information satisfies

$$\bar{J}_\mathrm{B} = \mathrm{E}_\theta[\bar{J}_\mathrm{F}(\theta)] + J_\theta \leq J_\mathrm{B} = \mathrm{E}_\theta[J_\mathrm{F}(\theta)] + J_\theta. \quad (6)$$

Substituting (6) into (2) proves *Theorem* 1.

Theorem 1 only tells us the increase of the updated BCRB of the updated data vector $\bar{\mathbf{x}}$. However, based on the MSE criterion and among all estimators, the minimum MSE $\mathcal{R}_\mathrm{ms}$ is achieved by the MMSE estimator $\hat{\theta}_\mathrm{ms}(\mathbf{x}) = \mathrm{E}_{\theta|\mathbf{x}}(\theta|\mathbf{x}) = \int \theta f_{\mathbf{x}|\theta}(\mathbf{x}|\theta)f_\theta(\theta)d\theta / \int f_{\mathbf{x}|\theta}(\mathbf{x}|\theta)f_\theta(\theta)d\theta$ [51], [54]. Therefore, an interesting question is whether the updated MMSE estimator $\hat{\vartheta}_\mathrm{ms}(\bar{\mathbf{x}}) = \mathrm{E}_{\theta|\bar{\mathbf{x}}}(\theta|\bar{\mathbf{x}})$ can achieve a lower MSE $\bar{\mathcal{R}}_\mathrm{ms}$ than that of the original MMSE estimator $\hat{\theta}_\mathrm{ms}$ or not. The answer is given in *Theorem* 2.

Theorem 2: It is impossible to design an updated MMSE estimator $\hat{\vartheta}_\mathrm{ms}(\bar{\mathbf{x}})$ to achieve a lower MSE $\bar{\mathcal{R}}_\mathrm{ms}$ than the orignial MMSE estimator $\hat{\theta}_\mathrm{ms}(\mathbf{x})$ does.

Proof of *Theorem* 2 is presented in Appendix A. Although this theorem leads to a negative aspect of the added noise to the optimal MMSE estimator $\hat{\theta}_\mathrm{ms}(\mathbf{x})$, it also indicates the possibility of noise benefits in some suboptimal estimators beyond the restricted conditions of [12], [20], [22]–[25], [27], [30], [31], [33], [36]–[43], [45]–[49]. In practice, the MMSE estimator $\hat{\theta}_\mathrm{ms}(\mathbf{x})$ is usually too computationally intensive to implement [51], [54], thus we will exploit the optimal added noise in some easily implementable suboptimal estimators as follows.

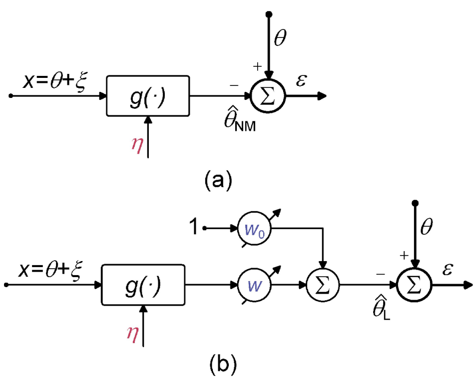

FIGURE 1. **Block diagram representations of (a) the noise-modified estimator $\hat{\theta}_\mathrm{NM}$ in (7) and (b) the LMMSE estimator $\hat{\theta}_\mathrm{L}$ in (13). The optimal noise η is intentionally injected into a given sensor g for the improvement of the MSE of the designed estimator.**

III. NOISE BENEFITS IN SUBOPTIMAL ESTIMATORS

A. NOISE BENEFITS IN A NOISE-MODIFIED ESTIMATOR

Consider the scalar-parameter observation model $x = \theta + \xi$, and the observation x plus the added noise η is applied to a fixed sensor g, as shown in Fig. 1 (a). Then, the noise-modified estimator

$$\hat{\theta}_\mathrm{NM} = g(x + \eta) \quad (7)$$

is established on the updated sensor output $g(x + \eta)$. Then, the artificially added noise η is optimized to minimize the MSE $\mathcal{R}_\mathrm{NM} = \mathrm{E}_{x,\eta}[(\hat{\theta}_\mathrm{NM} - \theta)^2] = \mathrm{E}_\theta(\theta^2) - \mathrm{E}_\eta\{\mathrm{E}_x[2\theta g(x + \eta) - g^2(x + \eta)]\}$. Since the two-moment $\mathrm{E}_\theta(\theta^2)$ is given, then Chen *et al.* [32] proved

$$\begin{aligned} \min_{f_\eta} \mathcal{R}_\mathrm{NM} &= \mathrm{E}_\theta(\theta^2) - \max_{f_\eta} \mathrm{E}_\eta\{\mathrm{E}_x[2\theta g(x + \eta) - g^2(x + \eta)]\} \\ &\geq \mathrm{E}_\theta(\theta^2) - \max_\eta \mathrm{E}_x[2\theta g(x + \eta) - g^2(x + \eta)], \end{aligned} \quad (8)$$

TABLE 1. MSEs of various estimators with optimal added noise.

Estimator \ noise ξ	Gaussian		Rayleigh		Laplace	
	MSE	f_η^{o}	MSE	f_η^{o}	MSE	f_η^{o}
$\hat{\theta}_{\text{NM}}(x)$	0.2571	/	0.3333	/	0.2643	/
$\hat{\theta}_{\text{NM}}(\bar{x})$	0.1643	$\delta(\eta+0.5)$	0.1242	$\delta(\eta+0.9)$	0.1771	$\delta(\eta+0.5)$
$\hat{\theta}_{\text{L}}(x)$	0.0701	/	0.0833	/	0.0741	/
$\hat{\theta}_{\text{L}}(\bar{x})$	0.0547	$\delta(\eta+0.5)$	0.0396	$\delta(\eta+0.9)$	0.0589	$\delta(\eta+0.5)$
$\hat{\theta}_{\text{NE}}(\bar{x})$	0.0448	Fig. 3(a)	0.0267	Fig. 3(b)	0.0536	Fig. 3(c)
$\hat{\theta}_{\text{LC}}(\bar{x})$	0.0447	Fig. 3(d)	0.0266	Fig. 3(e)	0.0533	Fig. 3(f)
$\hat{\theta}_{\text{ms}}(x)$	0.0446	/	0.0256	/	0.0533	/

and the optimal added noise accords with the PDF $f_\eta^{\text{o}}(\eta) = \delta(\eta - \eta^*)$ and the constant [32]

$$\eta^* = \arg\max_{\eta} \text{E}_x[2\theta g(x+\eta) - g^2(x+\eta)]. \tag{9}$$

Thus, with this optimal bias η^*, the MSE $\mathcal{R}_{\text{NM}}$ of the noise-modified estimator $\hat{\theta}_{\text{NM}}$ has a minimum

$$\mathcal{R}_{\text{NM}}^{\min} = \text{E}_\theta(\theta^2) - \text{E}_x[2\theta g(x+\eta^*) - g^2(x+\eta^*)]. \tag{10}$$

Based on *Theorem* 2 and compared with the minimum MSEs achieved by the MMSE estimators $\hat{\vartheta}_{\text{ms}}(\bar{x})$ and $\hat{\theta}_{\text{ms}}(x)$, $\mathcal{R}_{\text{NM}}^{\min}$ in (10) satisfies

$$\mathcal{R}_{\text{NM}}^{\min} \geq \bar{\mathcal{R}}_{\text{ms}} \geq \mathcal{R}_{\text{ms}}. \tag{11}$$

Example 1: Consider an uniformly distributed parameter θ with its PDF $f_\theta(\theta) = 1/a$ $(a > 0)$ over the interval $(0, a)$ and zero otherwise. A quantizer sensor is given by

$$g(u) = \begin{cases} 1, & u > \gamma, \\ 0, & u \leq \gamma, \end{cases} \tag{12}$$

where γ is the threshold value of the quantizer. In Table 1, consider three background noise types with Gaussian PDF $f_\xi(u) = \exp(-u^2/2\sigma_\xi^2)/\sqrt{2\pi\sigma_\xi^2}$, Rayleigh PDF $f_\xi(u) = (4-\pi)u\exp[-(4-\pi)u^2/4\sigma_\xi^2]/2\sigma_\xi^2$ $(u \geq 0)$, Laplace PDF $f_\xi(u) = \sqrt{2}\exp(-\sqrt{2}|u|/\sigma_\xi)/2\sigma_\xi$ and the standard deviation $\sigma_\xi > 0$. Here, the interval bound $a = 1$, the quantizer threshold $\gamma = 0$ and the standard deviations σ_ξ take $\sqrt{0.1}$, $\sqrt{(4-\pi)/20}$ and $1/\sqrt{5}$ for three types of considered background noise, respectively. It is seen in Table 1 that MSEs of $\hat{\theta}_{\text{NM}}(x)$ without the added noise are 0.2571, 0.3333 and 0.2643, respectively. With the optimal bias $\eta = \eta^*$ given in Table 1, the noise-modified quantizer $\hat{\theta}_{\text{NM}}(x+\eta^*)$ has the optimized MSEs of 0.1643, 0.1242 and 0.1771. However, compared with the MSEs of 0.0446, 0.0256 and 0.0533 achieved by the corresponding MMSE estimators $\hat{\theta}_{\text{ms}}(x)$, the improvement on the MSE of $\hat{\theta}_{\text{NM}}$ by the optimal added noise $\eta = \eta^*$ is limited.

B. NOISE BENEFITS IN A LMMSE ESTIMATOR

In order to further reduce the MSE of the noise-modified estimator, we perform the LMMSE transform on the sensor output $g(x+\eta)$ and establish a LMMSE estimator

$$\hat{\theta}_{\text{L}} = wg(x+\eta) + w_0 \tag{13}$$

with adjustable weights w and w_0, as shown in Fig.1 (b).

Theorem 3: For the LMMSE estimator $\hat{\theta}_{\text{L}}$ of (13), the optimal noise η has the PDF $f_\eta^{\text{o}}(\eta) = \delta(\eta - \eta^\dagger)$ with the constant

$$\eta^\dagger = \arg\max_{\eta} \frac{\{\text{E}_x[\theta g(x+\eta)] - \text{E}_\theta(\theta)\text{E}_x[g(x+\eta)]\}^2}{\text{E}_x[g^2(x+\eta)] - \text{E}_x^2[g(x+\eta)]}. \tag{14}$$

The MSE $\mathcal{R}_{\text{L}}$ of $\hat{\theta}_{\text{L}}$ has the minimum

$$\mathcal{R}_{\text{L}}^{\min} = \text{var}(\theta) - \frac{\left\{\text{E}_x[\theta g(x+\eta^\dagger)] - \text{E}_\theta(\theta)\text{E}_x[g(x+\eta^\dagger)]\right\}^2}{\text{E}_x[g^2(x+\eta^\dagger)] - \text{E}_x^2[g(x+\eta^\dagger)]} \leq \mathcal{R}_{\text{NM}}^{\min}, \tag{15}$$

where $\text{var}(\theta) = \text{E}_\theta(\theta^2) - \text{E}_\theta^2(\theta)$ is the variance of θ.

Proof of *Theorem* 3 is given in Appendix B. It is shown in Table 1 that, without the added noise, the MSEs of the LMMSE estimator $\hat{\theta}_{\text{L}}(x)$ are 0.0701, 0.0833 and 0.0741 for three background noise types, respectively, which already are lower than that of the noise-modified estimator $\hat{\theta}_{\text{NM}}(x+\eta^*)$ with its optimal added noise η^* in (9). Utilizing the optimal added noise $\eta^\dagger$ in (14), the MSEs of $\hat{\theta}_{\text{L}}(x+\eta^\dagger)$ can be further reduced to 0.0547, 0.0396 and 0.0589. However, the improved MSEs of $\hat{\theta}_{\text{L}}(x+\eta^\dagger)$ still cannot approach the MSE achieved by the MMSE estimator $\hat{\theta}_{\text{ms}}(x)$. Moreover, the optimal "noise" $\eta = \eta^\dagger$ is still a constant bias, rather than a *bona fide* random noise.

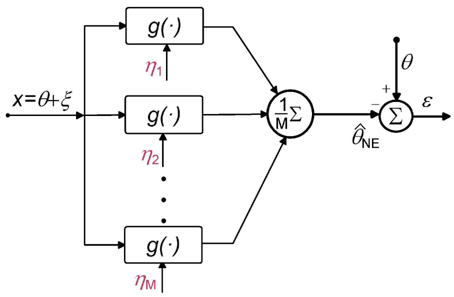

FIGURE 2. Block diagram representation of the noise-enhanced estimator $\hat{\theta}_{\text{NE}}$ in (16). Here, M mutually i.i.d. noise components η_m in sensors are optimized to minimize the MSE of $\hat{\theta}_{\text{NE}}$.

C. NOISE BENEFITS IN IDENTICAL SENSORS

The configurations of both noise-modified estimator $\hat{\theta}_{\text{NM}}$ and the LMMSE estimator $\hat{\theta}_{\text{L}}$ are only operated on one sensor. Next, consider an ensemble of identical sensors that receive the same input data x, and the mutually i.i.d. noise components η_m are also fed into each sensor, as shown in Fig. 2. Here, η_m are with the common PDF f_η and satisfy $\text{E}_\eta(\eta_m\eta_k) = 0$ for $m \neq k$ $(m, k = 1, 2, \cdots, M, M \geq 2)$. Then, the average value of all outputs of sensors forms the noise-enhanced estimator

$$\hat{\theta}_{\text{NE}} = \frac{1}{M}\sum_{m=1}^{M} g(x+\eta_m). \tag{16}$$

The MSE of $\hat{\theta}_{\text{NE}}$ can be computed as

$$\begin{aligned}\mathcal{R}_{\text{NE}} &= \mathrm{E}_{x,\eta}[(\theta - \hat{\theta}_{\text{NE}})^2] \\ &= \frac{\mathrm{E}_x\{\mathrm{E}_\eta[g^2(x+\eta)]\} + (M-1)\mathrm{E}_x\{\mathrm{E}^2_\eta[g(x+\eta)]\}}{M} \\ &\quad - 2\mathrm{E}_x\{\theta \mathrm{E}_\eta[g(x+\eta)]\} + \mathrm{E}_\theta(\theta^2), \qquad (17)\end{aligned}$$

where the correlations $\mathrm{E}_{x,\eta}[\theta g(x+\eta_m)] = \mathrm{E}_x\{\theta \mathrm{E}_\eta[g(x+\eta)]\}$ ($\forall m$) and $\mathrm{E}_x\{\mathrm{E}_\eta[g(x+\eta_m)]\mathrm{E}_\eta[\theta g(x+\eta_k)]\} = \mathrm{E}_x\{\mathrm{E}^2_\eta[g(x+\eta)]\}$ for $m \neq k$ ($m, k = 1, 2, \cdots, M$).

Theorem 4: For minimizing the MSE $\mathcal{R}_{\text{NE}}$ of the noise-enhanced estimator $\hat{\theta}_{\text{NE}}$ by mutually i.i.d. noise components η_m, the optimal noise is not with the PDF $f^{\text{o}}_\eta(\eta) = \delta(\eta - \eta^\ddagger)$ for a constant bias $\eta = \eta^\ddagger$. For the given background noise ξ and added noise components η_m, the MSE $\mathcal{R}_{\text{NE}}$ is a monotonically decreasing function of the sensor number M. Moreover, for a sufficiently large sensor number $M \to \infty$, the MSE satisfies

$$\lim_{M\to\infty} \min_{f_\eta} \mathcal{R}_{\text{NE}} \leq \mathcal{R}^{\min}_{\text{NM}}. \qquad (18)$$

Proof of *Theorem* 4 is given in Appendix C. Interestingly, as the sensor number $M \to \infty$, (16) can be asymptotically represented as

$$\hat{\theta}_{\text{NE}} = \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} g(x+\eta_m) = \mathrm{E}_\eta[g(x+\eta)], \qquad (19)$$

which is just the noise-enhanced estimator proposed by Uhlich [42]. *Theorem* 4 only tells us that the optimal noise is not a constant bias, but which type of noise is optimal to minimizing the MSE of the noise-enhanced estimator $\hat{\theta}_{\text{NE}}$? This non-convex problem is in general intractable, because the term $\mathrm{E}_x[\mathrm{E}^2_\eta(g(x+\eta))]$ in (17) is a nonlinear functional of the PDF f_η. Therefore, the minimization problem of the MSE $\min_{f_\eta} \mathcal{R}_{\text{NE}}$ usually employs the PDF approximation method [26], [42], [47], [48], [56] to obtain an approximate optimal solution form as

$$\tilde{f}^{\text{o}}_\eta(\eta) = \sum_{k=1}^{K} \lambda_k \phi\left(\frac{\eta - \mu_k}{\sigma_k}\right) \qquad (20)$$

with the normalization coefficients $\lambda_k \geq 0$ satisfying the constraint $\sum_{k=1}^{K} \lambda_k = 1$, and the Gaussian window function $\phi(u) = \exp(-u^2/2)/\sqrt{2\pi}$, means μ_k, standard deviations $\sigma_k \geq 0$. The approximate PDF $\tilde{f}^{\text{o}}_\eta$ will asymptotically converge to the existing optimal PDF f^{o}_η as the number K of the window function increases [26], [42], [47], [48], [56].

Example 2: For instance, consider $M = 1000$ identical quantizers of (12) and other parameters are the same as in *Example* 1. The sequential quadratic programming [56] is used to numerically solve the approximate PDF $\tilde{f}^{\text{o}}_\eta$ of (20). In Figs. 3 (a), (b) and (c), the approximate PDFs $\tilde{f}^{\text{o}}_\eta$ are plotted for estimating the uniform distributed parameter θ buried in three background noise types, respectively. It is seen in Fig. 3 that the approximate optimal added noise PDF $\tilde{f}^{\text{o}}_\eta$ exhibits non-trivial complicated shapes and varies with the background noise types. These approximate PDFs $\tilde{f}^{\text{o}}_\eta$ implies a *bona fide* random noise, rather than a constant bias. Moreover, substituting the obtained approximate optimal noise PDFs $\tilde{f}^{\text{o}}_\eta$ into (17), the corresponding MSE values of $\mathcal{R}_{\text{NE}}(\bar{x})$ are reduced to 0.0448, 0.0267 and 0.0536, as listed in Table 1, which are almost equal to the corresponding MSEs achieved by the MMSE estimator $\hat{\theta}_{\text{ms}}(x)$.

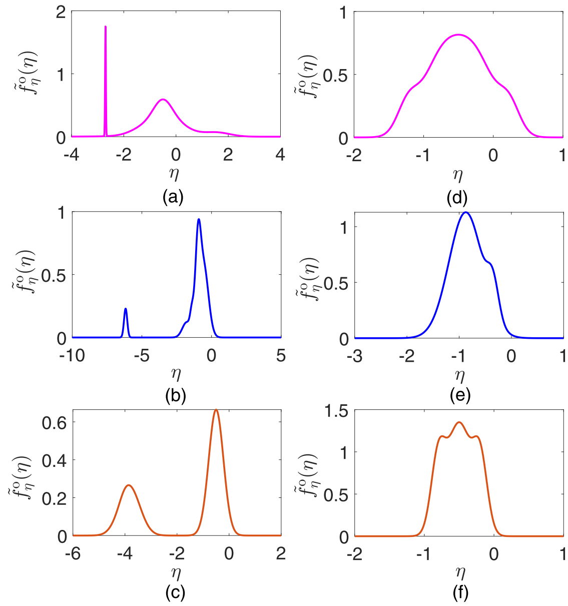

FIGURE 3. Approximate PDFs $\tilde{f}^{\text{o}}_\eta(\eta)$ for the noise enhanced estimator $\hat{\theta}_{\text{NE}}$ with background noise types of (a) Gaussian, (b) Rayleigh and (c) Laplace distributions. For the linear combination estimator $\hat{\theta}_{\text{LC}}$ in (21), approximate PDFs $\tilde{f}^{\text{o}}_\eta(\eta)$ are also presented for (d) Gaussian, (e) Rayleigh and (f) Laplace background noise types. Here, the windows number $K = 10$ in (20).

D. NOISE BENEFITS IN NONIDENTICAL SENSORS

From (16), the noise-enhanced estimator $\hat{\theta}_{\text{NE}}$ uniformly sets the same weight $1/M$ to the identical sensors g, which is inappropriate for an ensemble of nonidentical sensors g_m shown in Fig. 4. Carrying out a LMMSE transform on the sensor outputs $g_m(x+\eta_m)$, a linear combination estimator is established as

$$\hat{\theta}_{\text{LC}} = w_0 + \mathbf{w}^\top \mathbf{g}, \qquad (21)$$

where the sensor output vector $\mathbf{g} = [g_1(x+\eta_1), g_2(x+\eta_2), \cdots, g_M(x+\eta_M)]^\top$, the weight vector $\mathbf{w} = [w_1, w_2, \cdots, w_M]^\top$ and w_0 is the bias weight [48]. Then, the MSE of $\hat{\theta}_{\text{LC}}$ can be expressed as

$$\mathcal{R}_{\text{LC}} = \mathrm{E}_{x,\eta}[(\theta - \hat{\theta}_{\text{LC}})^2] = \mathrm{E}_{x,\eta}[(\theta - w_0 - \mathbf{w}^\top\mathbf{g})^2]. \qquad (22)$$

An interesting fact of $\mathcal{R}_{\text{LC}}$ in (22) is that the minimization of $\mathcal{R}_{\text{LC}}$ with respect to weights, uncoupled with the minimization of $\mathcal{R}_{\text{LC}}$ with respect to the added noise, can be first theoretically solved. Setting the derivative $\partial\mathcal{R}_{\text{LC}}/\partial w_0 = 0$ and the gradient $\partial\mathcal{R}_{\text{LC}}/\partial\mathbf{w} = \mathbf{0}$ produce $w_0 = \mathrm{E}_\theta(\theta) - \mathbf{w}^\top\mathrm{E}_{x,\eta}(\mathbf{g})$ and $\mathbf{w} = \mathbf{C}^{-1}\mathbf{p}$, where $\mathbf{p} = \mathrm{E}_{x,\eta}\{[\theta - \mathrm{E}_\theta(\theta)][\mathbf{g} - \mathrm{E}_{x,\eta}(\mathbf{g})]\}$ is the cross-correlation vector of the parameter θ and the sensor

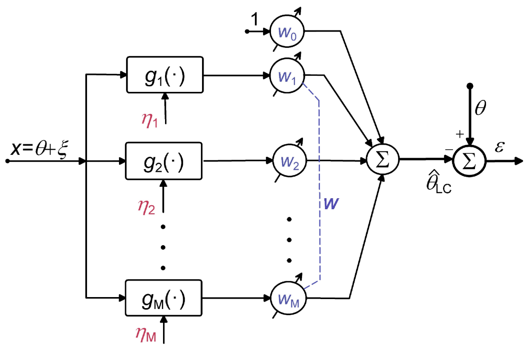

FIGURE 4. **Block diagram representation of the linear combination estimator $\hat{\theta}_{LC}$ in (21). Besides the injection of mutually independent added noise components η_m, each sensor is also endowed with an adjustable weight w_m.**

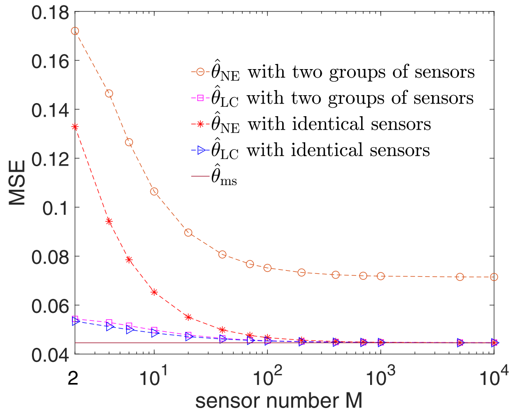

FIGURE 5. **MSEs of the linear combination estimator $\hat{\theta}_{LC}$ in (21) and the noise-enhanced estimator $\hat{\theta}_{NE}$ in (16) versus the sensor number M.**

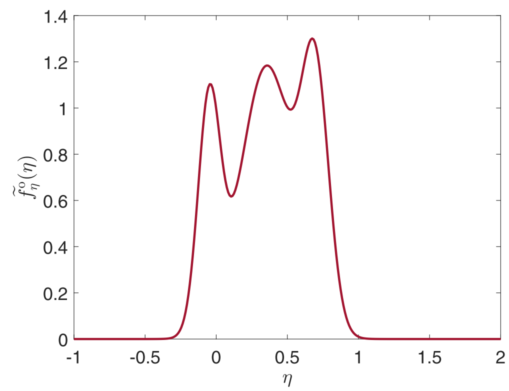

FIGURE 6. **Approximate PDF $\tilde{f}_\eta^o(\eta)$ for the linear combination estimator $\hat{\theta}_{LC}$ in (21) with two groups of nonidentical sensors g_m. Here, the number of sensors is $M = 200$, one group of $M/2$ quantizers has the threshold $\gamma_1 = 0$, and the other group of $M/2$ quantizers is with the threshold $\gamma_2 = 1$. The background Gaussian noise ξ is with the zero-mean and standard deviation $\sigma_\xi = \sqrt{0.1}$. The windows number $K = 10$ in (20).**

vector $\mathbf{g}$ and $\mathbf{C} = \mathrm{E}_{x,\eta}\{[\mathbf{g} - \mathrm{E}_{x,\eta}(\mathbf{g})][\mathbf{g} - \mathrm{E}_{x,\eta}(\mathbf{g})]^\top\}$ is the covariance matrix of $\mathbf{g}$ [48], [57]. Then, using the optimal weights w_0 and $\mathbf{w}$, the linear combination estimator $\hat{\theta}_{LC}$ in the LMMSE sense can be rewritten as

$$\hat{\theta}_{LC} = \mathrm{E}_\theta(\theta) + \mathbf{p}^\top\mathbf{C}^{-1}[\mathbf{g} - \mathrm{E}_{x,\eta}(\mathbf{g})] \tag{23}$$

and the minimized MSE $\mathcal{R}_{LC}$ with respect to weights is given by

$$\mathcal{R}_{LC} = \mathrm{var}(\theta) - \mathbf{p}^\top\mathbf{C}^{-1}\mathbf{p}. \tag{24}$$

Theorem 5: The linear combination estimator $\hat{\theta}_{LC}$ is never worse than the noise-enhanced estimator $\hat{\theta}_{NE}$ in (16) in the same environment, and the MSEs satisfy $\mathcal{R}_{LC} \leq \mathcal{R}_{NE}$.

Theorem 5 is proved in Appendix D, and is illustrated in the following examples.

Example 3: Consider again *Example* 1 and minimize the MSE $\mathcal{R}_{LC}$ in (24) by optimizing the added noise. For $M = 1000$ identical quantizers $g_m = g$ of (12) with the same threshold $\gamma = 0$, the approximate noise PDFs $\tilde{f}_\eta^o$ of the optimization problem of $\min_{f_\eta} \mathcal{R}_{LC}$ are also numerically solved for three background noise types, as shown in Fig. 3 (d), (e) and (f), respectively. The corresponding MSE values of $\mathcal{R}_{LC}$ are 0.0447, 0.0266 and 0.0533, as listed in Table 1, which also approach to the MSE achieved by the MMSE estimator $\hat{\theta}_{ms}$. However, compared to the noise-enhanced estimator $\hat{\theta}_{NE}$, $\hat{\theta}_{LC}$ improves the MSE slightly in estimating parameters from the observations of a large number of identical sensors.

Example 4: Consider M (even number) quantizers g_m with two groups: one group of $M/2$ quantizers has the same threshold $\gamma_1 = 0$, and the other group of $M/2$ quantizers is with the same threshold $\gamma_2 = 1$. The background noise ξ is selected as Gaussian distributed with the zero-mean and standard deviation $\sigma_\xi = \sqrt{0.1}$, and other conditions are the same as in *Example* 1. The MSEs of estimators $\hat{\theta}_{NE}$ and $\hat{\theta}_{LC}$ are plotted as a function of the sensor number M in Fig. 5, wherein the MSE is minimized with respect to the added noise by the approximate PDF in (20). It is seen in Fig. 5 that, upon increasing the sensor number M and dividing these sensors into two groups, the MSE of the linear combination estimator $\hat{\theta}_{LC}$ (□) still approaches 0.0446 achieved by the MMSE estimator $\hat{\theta}_{ms}$ (dashed line) asymptotically. While, under the same condition, the MSE of the noise enhanced estimator $\hat{\theta}_{NE}$ (○) approaches 0.0715 asymptotically at large M, rather than 0.0446. The reason is that the linear combination estimator $\hat{\theta}_{LC}$ allocates different weights to sensors g_m with different thresholds. For instance, for $M = 200$, the weights $w_m = 3.0976 \times 10^{-3}$ $(m = 1, 2, \cdots, 100)$ for the group of quantizers with threshold γ_1, and the weights $w_k = 5.1738 \times 10^{-3}$ $(k = 1, 2, \cdots, 100)$ for the group of quantizers with threshold γ_2. But the estimator $\hat{\theta}_{NE}$ in (16) always endows all sensors with a fixed weight $1/M$, regardless of the distinction of thresholds for two groups of sensors. For comparison, the MSEs of both $\hat{\theta}_{LC}$ (▷) and $\hat{\theta}_{NE}$ (∗) are also plotted in Fig. 5 for M identical sensors g with threshold $\gamma_1 = 0$, which also all approach the MMSE value of 0.0446 as the number M increases. For a moderate sensor number (e.g. $2 \leq M \leq 10^2$), the MSE $\mathcal{R}_{LC}$ (▷) of $\hat{\theta}_{LC}$ also clearly outperforms the MSE $\mathcal{R}_{NE}$ (∗) of $\hat{\theta}_{NE}$. These points indicate the superiority of the linear combination estimator $\hat{\theta}_{\mathrm{LC}}$ over the noise enhanced estimator $\hat{\theta}_{\mathrm{NE}}$ for Bayesian parameter estimation. In addition, for the total number $M = 200$ of sensors with two groups, the approximate noise PDFs $\tilde{f}_{\eta}^{\mathrm{o}}$ that minimizes the MSE $\mathcal{R}_{\mathrm{LC}}$ to 0.0451 is also numerically solved and shown in Fig. 6, which also has quite a complicated structure. For other threshold settings of γ (not shown here), the superiority of the linear combination estimator is also confirmed.

IV. CONCLUSION

In this paper, added noise is intentionally injected into the observed data, we first theoretically address a crucial question of the increase of the BCRB of the updated observations, and then prove that the updated MMSE estimator yet cannot provide a lower MSE than that the original MMSE estimator. We mainly investigate the noise benefits in certain types of suboptimal Bayesian estimators that are widely employed due to their ease of implementation and low cost. For the noise-modified estimator in (7) or the LMMSE estimator in (13) based on a single sensor, the optimal added noise that minimizes the estimator MSE is just a constant bias, but not *bona fide* random noise. However, for the noise-enhanced estimator in (16) and the linear combination estimator in (21) established on an ensemble of sensors, it is observed that the optimal noise that improves the estimator and makes it as efficient as the original MMSE estimator is a *bona fide* random signal, rather than a constant bias. Especially, for an ensemble of two groups sensors with different settings, the linear combination estimator in (21), benefiting from the optimal added noise, can still approach the MSE of the original MMSE estimator when the sensor number is sufficiently large.

Some open questions remain. For instance, it is seen in (23) that the construction of the linear combination estimator $\hat{\theta}_{\mathrm{LC}}$ requires the theoretical two-moments of the cross-correlation between the parameter and the sensor outputs and the covariance of the sensor outputs. To find these statistical quantities we need the joint PDF of the observation data and the added noise, of which, most of time, we have no knowledge of the probabilistic structure. Thus, under this circumstance, how do we establish a practical estimator? If the observation data is stationary and ergodic, can we approximately compute these desired statistical characteristics from one sampling realization of data? Or on the minimization of the least squares error criterion, how do we establish an easily implementable least-square estimator and which kind of added noise is optimal for improving the MSE of the least-square estimator? In addition, in many signal estimation problems, the observations are obtained in a sequential order as time processes. So, can we present a sequential linear combination estimator in the LMMSE sense that continuously update weights and the added noise according to the new incoming data? These interesting questions deserve the further study.

APPENDIXES

APPENDIX A

PROOF OF THEOREM 2

Consider the case of $s(\theta) = \theta$ without loss of generality. For the updated observation data $\bar{x}_n = \theta + \xi_n + \eta_n = \theta + z_n$, the joint PDF of the random parameter θ and the data vector $\bar{\mathbf{x}}$ is described as $f_{\bar{\mathbf{x}},\theta}(\bar{\mathbf{x}}, \theta) = f_{\bar{\mathbf{x}}|\theta}(\bar{\mathbf{x}}|\theta) f_\theta(\theta)$, where the conditional PDF can be expressed as $f_{\bar{\mathbf{x}}|\theta}(\bar{\mathbf{x}}|\theta) = \prod_{n=1}^{N} f_z(\bar{x}_n - \theta) = \prod_{n=1}^{N} \int f_\xi(\bar{x}_n - \theta - \eta_n) f_\eta(\eta_n) d\eta_n = \int f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta}) f_{\boldsymbol{\eta}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$. Then, the updated MMSE estimator is given by $\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}}) = \mathrm{E}_{\theta|\bar{\mathbf{x}}}[\theta|\bar{\mathbf{x}}] = \int \theta f_{\theta|\bar{\mathbf{x}}}(\theta|\bar{\mathbf{x}}) d\theta$ with the conditional posterior PDF $f_{\theta|\bar{\mathbf{x}}}(\theta|\bar{\mathbf{x}}) = f_{\bar{\mathbf{x}},\theta}(\bar{\mathbf{x}}, \theta)/f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}})$ and the PDF $f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}}) = \int f_{\bar{\mathbf{x}},\theta}(\bar{\mathbf{x}}, \theta) d\theta$. The MSE of $\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})$ is given by $\bar{\mathcal{R}}_{\mathrm{ms}} = \mathrm{E}_{\bar{\mathbf{x}},\theta}[(\theta - \hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}}))^2] = \mathrm{E}_\theta[\theta^2] - \mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2]$. Since the two-moment $\mathrm{E}_\theta[\theta^2]$ is given, then we now find the optimal added noise vector $\boldsymbol{\eta}$ to maximize the term $\mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2]$.

Consider a real-value vector function $f(\mathbf{Z}) = Z_1^2/Z_2$ with $Z = [Z_1, Z_2]^T$ and $Z_2 > 0$. Since its Hessian matrix $\nabla^2 f(\mathbf{Z}) = 2[Z_2, -Z_1]^T[Z_2, -Z_1]/Z_2^3$ is positive semidefinite, then $f(\mathbf{Z})$ is convex [33], [46], [58]. Thus, Jensen's inequality $\mathrm{E}_\eta^2[Z_1]/\mathrm{E}_\eta[Z_2] \leq \mathrm{E}_\eta[Z_1^2/Z_2]$ indicates that $\mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2]$ satisfies

$$\begin{aligned}
\mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2] &= \int \frac{\left[\int \theta f_{\bar{\mathbf{x}}|\theta}(\bar{\mathbf{x}}|\theta) f_\theta(\theta) d\theta\right]^2}{f_{\bar{\mathbf{x}}}(\bar{\mathbf{x}})} d\bar{\mathbf{x}} \\
&= \int \frac{\mathrm{E}_{\boldsymbol{\eta}}^2\{\mathrm{E}_\theta[\theta f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]\}}{\mathrm{E}_{\boldsymbol{\eta}}\{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]\}} d\bar{\mathbf{x}} \\
&\leq \int \mathrm{E}_{\boldsymbol{\eta}} \left\{ \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]}{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]} \right\} d\bar{\mathbf{x}} \qquad (25) \\
&= \mathrm{E}_{\boldsymbol{\eta}} \left[\int \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]}{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]} d\bar{\mathbf{x}} \right] \\
&\leq \max_{\boldsymbol{\eta}} \int \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]}{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta})]} d\bar{\mathbf{x}} \\
&= \max_{\boldsymbol{\eta}} p(\boldsymbol{\eta}), \qquad (26)
\end{aligned}$$

where the inequality (25) accords to the Jensen's inequality [33], [46], [58] and (26) holds for $f_{\boldsymbol{\eta}}^{\mathrm{o}}(\boldsymbol{\eta}) = \delta(\boldsymbol{\eta} - \boldsymbol{\eta}^*)$ with the constant vector $\boldsymbol{\eta}^* = \arg\max_{\boldsymbol{\eta}} p(\boldsymbol{\eta})$. Furthermore, from (26) and noting $d\bar{\mathbf{x}} = d(\bar{\mathbf{x}} - \boldsymbol{\eta}^*)$, we find

$$\begin{aligned}
\mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2] &\leq \int \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta}^*)]}{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\bar{\mathbf{x}} - \theta - \boldsymbol{\eta}^*)]} d\bar{\mathbf{x}} \\
&= \int \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\mathbf{x} - \theta)]}{\mathrm{E}_\theta[f_{\boldsymbol{\xi}}(\mathbf{x} - \theta)]} d\mathbf{x} \\
&= \int \frac{\mathrm{E}_\theta^2[\theta f_{\boldsymbol{\xi}}(\mathbf{x} - \theta)]}{f_{\mathbf{x}}^2(\mathbf{x})} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \\
&= \int \left(\int \theta f_{\mathbf{x}|\theta}(\mathbf{x}|\theta) \right)^2 f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \\
&= \mathrm{E}_{\mathbf{x}}[\hat{\theta}_{\mathrm{ms}}(\mathbf{x})^2]. \qquad (27)
\end{aligned}$$

Thus, we find $\bar{\mathcal{R}}_{\mathrm{ms}} = \mathrm{E}_\theta[\theta^2] - \mathrm{E}_{\bar{\mathbf{x}}}[\hat{\vartheta}_{\mathrm{ms}}(\bar{\mathbf{x}})^2] \geq \mathrm{E}_\theta[\theta^2] - \mathrm{E}_{\mathbf{x}}[\hat{\theta}_{\mathrm{ms}}(\mathbf{x})^2] = \mathcal{R}_{\mathrm{ms}}$. *Theorem* 2 holds.

APPENDIX B
PROOF OF THEOREM 3

The MSE of $\hat{\theta}_{\mathrm{L}}(x+\eta)$ in (13) can be written as $\mathcal{R}_{\mathrm{L}} = \mathrm{E}_{x,\eta}\left[(\theta - wg(x+\eta) - w_0)^2\right]$. Setting the derivative $\partial R_{\mathrm{L}}/\partial w_0 = 0$, the optimum biasing weight w_0 is solved as $w_0 = \mathrm{E}_\theta(\theta) - w\mathrm{E}_{x,\eta}[g(x+\eta)]$, which happens to make $\hat{\theta}_{\mathrm{L}}$ unbiased as $\mathrm{E}_{x,\eta}(\hat{\theta}_{\mathrm{L}}) = \mathrm{E}_\theta(\theta)$. Substituting w_0 into $\mathcal{R}_{\mathrm{L}}$ and setting $\partial R_{\mathrm{L}}/\partial w = 0$, we have

$$w = \frac{\mathrm{E}_{x,\eta}[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_{x,\eta}\left[\tilde{g}^2(x+\eta)\right]} \tag{28}$$

with $\tilde{g}(x+\eta) = g(x+\eta) - \mathrm{E}_x\{\mathrm{E}_\eta[g(x+\eta)]\}$ and $\tilde{\theta} = \theta - \mathrm{E}_\theta(\theta)$. Subsequently, with the solved weight w in (28), the MSE of $\hat{\theta}_{\mathrm{L}}$ can be rewritten as

$$\mathcal{R}_{\mathrm{L}} = \mathrm{var}(\theta) - \frac{\mathrm{E}^2_{x,\eta}[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_{x,\eta}\left[\tilde{g}^2(x+\eta)\right]}. \tag{29}$$

Since the variance $\mathrm{var}(\theta)$ of θ is given, then we maximize the second term of (29) as

$$\begin{aligned}\frac{\mathrm{E}^2_{x,\eta}[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_{x,\eta}\left[\tilde{g}^2(x+\eta)\right]} &= \frac{\mathrm{E}^2_{x,\eta}[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_{x,\eta}\left[g^2(x+\eta)\right] - \mathrm{E}^2_{x,\eta}[g(x+\eta)]} \\ &\le \frac{\mathrm{E}^2_\eta\{\mathrm{E}_x[\tilde{\theta}\tilde{g}(x+\eta)]\}}{\mathrm{E}_\eta\{\mathrm{E}_x[g^2(x+\eta)]\} - \mathrm{E}_\eta\{\mathrm{E}^2_x[g(x+\eta)]\}} \qquad (30)\\ &= \frac{\mathrm{E}^2_\eta\{\mathrm{E}_x[\tilde{\theta}\tilde{g}(x+\eta)]\}}{\mathrm{E}_\eta\{\mathrm{E}_x\left[\tilde{g}^2(x+\eta)\right]\}} \\ &\le \mathrm{E}_\eta\left\{\frac{\mathrm{E}^2_x[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_x[\tilde{g}^2(x+\eta)]}\right\} \qquad (31)\\ &\le \max_\eta \frac{\mathrm{E}^2_x[\tilde{\theta}\tilde{g}(x+\eta)]}{\mathrm{E}_x[\tilde{g}^2(x+\eta)]}.\end{aligned}$$

Using Jensen's inequality [33], [46], [55], [58], the inequality of (30) holds due to $\mathrm{E}^2_{x,\eta}[g(x+\eta)] = \mathrm{E}^2_\eta\{\mathrm{E}_x[g(x+\eta)]\} \le \mathrm{E}_\eta\{\mathrm{E}^2_x[g(x+\eta)]\}$ based on the convex function x^2, and the inequality of (31) is valid for the convex $f(\mathbf{Z}) = Z_1^2/Z_2$ given in Appendix A. Then, we find the optimal noise PDF $f^{\mathrm{o}}_\eta(\eta) = \delta(\eta - \eta^\dagger)$ in (14) and the minimum MSE in (15). Furthermore, the minimum MSE $\mathcal{R}^{\min}_{\mathrm{L}}$ of the estimator $\hat{\theta}_{\mathrm{L}}$ also satisfies

$$\begin{aligned}&\mathcal{R}^{\min}_{\mathrm{L}} \\ &= \min_{w,f_\eta} \mathrm{E}_{\bar{x},\theta}\left[\left(\theta - \mathrm{E}_\theta(\theta) - wg(x+\eta) + w\mathrm{E}_{x,\eta}[g(x+\eta)]\right)^2\right] \\ &\le \min_{f_\eta} \mathrm{E}_{\bar{x},\theta}\left[\left(\theta - \mathrm{E}_\theta(\theta) - g(x+\eta) + \mathrm{E}_{x,\eta}[g(x+\eta)]\right)^2\right] \\ &\le \mathrm{E}_{x,\theta}\left[\left(\theta - \mathrm{E}_\theta(\theta) - g(x+\eta^*) + \mathrm{E}_x[g(x+\eta^*)]\right)^2\right] \\ &= \mathcal{R}^{\min}_{\mathrm{NM}} - (\mathrm{E}_\theta(\theta) - \mathrm{E}_x[g(x+\eta^*)])^2 \\ &\le \mathcal{R}^{\min}_{\mathrm{NM}},\end{aligned} \tag{32}$$

where the constant η^* is given in (9). Then, *Theorem* 3 holds.

APPENDIX C
PROOF OF THEOREM 4

If the optimal noise η has the PDF $f^{\mathrm{o}}_\eta(\eta) = \delta(\eta - \eta^\ddagger)$, then M equivalent constants $\eta_m = \eta^\ddagger$ are added to M identical sensors g with the same outputs $g(x+\eta^\ddagger)$. Thus, the noise enhanced estimator $\hat{\theta}_{\mathrm{NE}} = \sum_{m=1}^M g(x+\eta^\ddagger)/M = g(x+\eta^\ddagger)$ reduces to the output of a single sensor. Moreover, $\mathrm{E}_\eta(\eta_m\eta_k) = (\eta^\ddagger)^2 \neq 0$ does not satisfy the the mutually independent assumption of η_m. Therefore, the optimal noise must not be a constant bias. Using the Jensen inequality and $\mathrm{E}_\eta[g(x+\eta)] \neq g(x+\eta)$ for the given observation data x, we obtain the inequality $\mathrm{E}_\eta[g^2(x+\eta)] > \mathrm{E}^2_\eta[g(x+\eta)]$. Then, we have

$$\mathrm{E}_x\left\{\mathrm{E}_\eta[g^2(w+\eta)]\right\} > \mathrm{E}_x\left\{\mathrm{E}^2_\eta[g(w+\eta)]\right\}. \tag{33}$$

Immediately, we find

$$\begin{aligned}&\frac{\mathrm{E}_x[\mathrm{E}_\eta(g^2(x+\eta))] + (M-1)\mathrm{E}_x[\mathrm{E}^2_\eta(g(x+\eta))]}{M} \\ &> \frac{\mathrm{E}_x[\mathrm{E}_\eta(g^2(x+\eta))] + M\mathrm{E}_x[\mathrm{E}^2_\eta(g(x+\eta))]}{M+1}.\end{aligned} \tag{34}$$

From (34), we deduce that $\mathcal{R}_{\mathrm{NE}}$ in (17) is a monotonically decreasing function of the sensor number M, when the observation data x and the added noise are given. Furthermore, for a sufficiently large number $M \to \infty$, the MSE $\mathcal{R}_{\mathrm{NE}}$ in (17) can be simplified as

$$\begin{aligned}\lim_{M\to\infty}\min_{f_\eta} \mathcal{R}_{\mathrm{NE}} &= \min_{f_\eta}\Big(\mathrm{E}_x[\mathrm{E}^2_\eta(g(x+\eta))] \\ &\quad - 2\mathrm{E}_x\{\theta\mathrm{E}_\eta[g(x+\eta)]\} + \mathrm{E}_\theta(\theta^2)\Big) \\ &= \min_{f_\eta} \mathrm{E}_x\{[\theta - \mathrm{E}_\eta(g(x+\eta))]^2\} \\ &\le \mathrm{E}_x\{[\theta - g(x+\eta^*)]^2\} = \mathcal{R}^{\min}_{\mathrm{NM}},\end{aligned} \tag{35}$$

where the constant η^* is given in (9). Then, *Theorem* 4 is proved.

APPENDIX D
PROOF OF THEOREM 5

For an ensemble of sensors g_m, the noise-enhanced estimator can be rewritten as

$$\hat{\theta}_{\mathrm{NE}} = \frac{1}{M}\sum_{m=1}^M g_m(x+\eta_m) \tag{36}$$

and its minimum MSE can be expressed as

$$\begin{aligned}\mathcal{R}^{\min}_{\mathrm{NE}} &= \min_{f_\eta} \mathrm{E}_{x,\eta}\left[\left(\theta - \hat{\theta}_{\mathrm{NE}}\right)^2\right] \\ &= \mathrm{E}_{x,\eta}\left\{\left[\theta - \frac{1}{M}\sum_{m=1}^M g_m(x+\eta_m)\right]^2\right\}\Bigg|_{f_\eta = f^{\mathrm{o}*}_\eta}\end{aligned} \tag{37}$$

with the optimal added noise PDF $f^{\mathrm{o}*}_\eta$. However, for any PDF f_η of the added noise and the $M \times 1$ dimensional vector $\mathbf{1}$ of all ones, the MSE of the designed linear combination estimator in (23) can be expressed as

$$\begin{aligned}\mathcal{R}_{\mathrm{LC}} &= \min_{w_0,\mathbf{w}} \mathrm{E}_{x,\eta}\left[(\theta - w_0 - \mathbf{w}^\top\mathbf{g})^2\right] \\ &\le \mathrm{E}_{x,\eta}\left\{\left[\theta - \frac{1}{M}\sum_{m=1}^M g_m(x+\eta_m)\right]^2\right\}\Bigg|_{w_0=0,\mathbf{w}=\mathbf{1}/M} \\ &= \mathcal{R}_{\mathrm{NE}}.\end{aligned} \tag{38}$$

Of course, even when the added noise PDF f_η receives the expression f_η^{o*} that is optimal for the noise-enhanced estimator $\hat{\theta}_{\text{NE}}$, the inequality (38) also holds, resulting in $\mathcal{R}_{\text{LC}}|_{f_\eta=f_\eta^{o*}} \le \mathcal{R}_{\text{NE}}^{\min}$. Thus, *Theorem* 5 holds.

REFERENCES

[1] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 52, no. 5, pp. 4691–4698, Jul. 2002.

[2] B. Ando and S. Graziani, "Adding noise to improve measurement," *IEEE Instrum. Meas. Mag.*, vol. 4, no. 1, pp. 24–31, Mar. 2001.

[3] S. Engelberg, "An introduction to dither," *IEEE Instrum. Meas. Mag.*, vol. 15, no. 6, pp. 50–54, Dec. 2012.

[4] L. Roberts, "Picture coding using pseudo-random noise," *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 145–154, Feb. 1962

[5] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Trans. Commun.*, vol. 12, no. 4, pp. 162–165, Dec. 1964.

[6] R. Benzi, A. Sutera, and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A, Math. Gen.*, vol. 14, no. 11, pp. L453–L457, 1981.

[7] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, "Stochastic resonance," *Rev. Mod. Phys.*, vol. 70, no. 1, pp. 233–287, 1998.

[8] V. S. Anishchenko, A. B. Neiman, F. Moss, and L. Shimansky-Geier, "Stochastic resonance: Noise-enhanced order," *Uspekhi Fizicheskikh Nauk*, vol. 42, no. 1, pp. 7–36, 1999.

[9] G. Harmer, B. Davis, and D. Abbott, "A review of stochastic resonance: Circuits and measurement," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 2, pp. 299–309, Apr. 2002.

[10] F. Moss, "Stochastic resonance and sensory information processing: A tutorial and review of application," *Clin. Neurophysiol.*, vol. 115, no. 2, pp. 267–281, Feb. 2004.

[11] L. M. Ward, "Physics of neural synchronisation mediated by stochastic resonance," *Contemp. Phys.*, vol. 50, no. 5, pp. 563–574, Sep. 2009.

[12] M. D. McDonnell, N. G. Stocks, C. E. M. Pearce, and D. Abbott, *Stochastic Resonance: From Suprathreshold Stochastic Resonance To Stochastic Signal Quantization*. New York, NY, USA: Cambridge Univ. Press, 2008.

[13] M. D. Mcdonnell and D. Abbott, "What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology," *PLoS Comput. Biol.*, vol. 5, no. 5, May 2009, Art. no. e1000348.

[14] M. D. McDonnell, "Is electrical noise useful?" *Proc. IEEE*, vol. 99, no. 2, pp. 242–246, Feb. 2011.

[15] Z. Qiao, Y. Lei, and N. Li, "Applications of stochastic resonance to machinery fault detection: A review and tutorial," *Mech. Syst. Signal Process.*, vol. 122, pp. 502–536, May 2019.

[16] S. Lu, Q. He, and J. Wang, "A review of stochastic resonance in rotating machine fault detection," *Mech. Syst. Signal Process.*, vol. 116, pp. 230–260, Feb. 2019.

[17] O. White, J. Babič, C. Trenado, L. Johannsen, and N. Goswami, "The promise of stochastic resonance in falls prevention," *Frontiers Physiol.*, vol. 9, p. 1865, Jan. 2019.

[18] J. J. Collins, C. C. Chow, and T. T. Imhoff, "Stochastic resonance without tuning," *Nature*, vol. 376, no. 6537, pp. 236–238, Jul. 1995.

[19] J. J. Collins, C. C. Chow, and T. T. Imhoff, "Aperiodic stochastic resonance in excitable systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 52, no. 4, pp. R3321–R3324, Jul. 2002.

[20] N. G. Stocks, "Suprathreshold stochastic resonance in multilevel threshold systems," *Phys. Rev. Lett.*, vol. 84, no. 11, pp. 2310–2313, Jul. 2002.

[21] F. Chapeau-Blondeau and X. Godivier, "Theory of stochastic resonance in signal transmission by static nonlinear systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 55, no. 2, pp. 1478–1495, Jul. 2002.

[22] A. B. Gershman and J. F. Böhme, "A pseudo-noise approach to direction finding," *Signal Process.*, vol. 71, no. 1, pp. 1–13, Nov. 1998.

[23] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[24] F. Chapeau-Blondeau, "Noise-aided nonlinear Bayesian estimation," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 66, no. 3, 2002, Art. no. 032101.

[25] F. Chapeau-Blondeau and D. Rousseau, "Noise-enhanced performance for an optimal Bayesian estimator," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1327–1334, May 2004.

[26] S. Bayram, S. Gezici, and H. V. Poor, "Noise enhanced hypothesis-testing in the restricted Bayesian framework," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3972–3989, Aug. 2010.

[27] S. Zozor and P.-O. Amblard, "Noise-aided processing: Revisting dithering in a ΣΔ quantizer," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3202–3210, Aug. 2005.

[28] H. Chen, P. K. Varshney, S. M. Kay, and J. H. Michels, "Theory of the stochastic resonance effect in signal detection: Part I–fixed detectors," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3172–3184, Jul. 2007.

[29] H. Chen and P. Varshney, "Theory of the stochastic resonance effect in signal detection—Part II: Variable detectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5031–5041, Oct. 2008.

[30] D. Rousseau and F. Chapeau-Blondeau, "Noise-improved Bayesian estimation with arrays of one-bit quantizers," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 6, pp. 2658–2662, Dec. 2007.

[31] F. Chapeau-Blondeau and D. Rousseau, "Raising the noise to improve performance in optimal processing," *J. Stat. Mech. Theory Exp.*, vol. 53, no. 1, pp. 215–238, 2009.

[32] H. Chen, P. Varshney, and J. Michels, "Noise enhanced parameter estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5074–5081, Oct. 2008.

[33] G. O. Balkan and S. Gezici, "CRLB based optimal noise enhanced parameter estimation using quantized observations," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 477–480, May 2010.

[34] S. Kay, "Can detectability be improved by adding noise?" *IEEE Signal Process. Lett.*, vol. 7, no. 1, pp. 8–10, Jan. 2000.

[35] S. Zozor and P. Amblard, "Stochastic resonance in locally optimal detectors," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3177–3181, Dec. 2003.

[36] A. Patel and B. Kosko, "Optimal noise benefits in Neyman–Pearson and inequality-constrained statistical signal detection," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1655–1669, May 2009.

[37] A. Patel and B. Kosko, "Optimal mean-square noise benefits in quantizer-array linear estimation," *IEEE Signal Process. Lett.*, vol. 17, no. 12, pp. 1005–1009, Dec. 2010.

[38] H. Chen and P. Varshney, "Performance limit for distributed estimation systems with identical one-bit quantizers," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 466–471, Jan. 2010.

[39] H. Chen, L. R. Varshney, and P. K. Varshney, "Noise-enhanced information systems," *Proc. IEEE*, vol. 102, no. 10, pp. 1607–1621, Oct. 2014.

[40] J. Zhu, X. Lin, R. S. Blum, and Y. Gu, "Parameter estimation from quantized observations in multiplicative noise environments," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 4037–4050, Aug. 2015.

[41] H. Soganci, S. Gezici, and O. Arikan, "Optimal signal design for multi-parameter estimation problems," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6074–6085, Nov. 2015.

[42] S. Uhlich, "Bayes risk reduction of estimators using artificial observation noise," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5535–5545, Oct. 2015.

[43] A. B. Akbay and S. Gezici, "Noise benefits in joint detection and estimation problems," *Signal Process.*, vol. 118, pp. 235–247, Jan. 2016.

[44] H. Dong, H. Wang, X. Shen, and Z. Jiang, "Effects of second-order matched stochastic resonance for weak signal detection," *IEEE Access*, vol. 6, pp. 46505–46515, 2018.

[45] Q. Zhai and Y. Wang, "Noise effect on signal quantization in an array of binary quantizers," *Signal Process.*, vol. 152, pp. 265–272, Nov. 2018.

[46] Y. Pan, F. Duan, F. Chapeau-Blondeau, and D. Abbott, "Noise enhancement in robust estimation of location," *IEEE Trans. Signal Process.*, vol. 66, no. 8, pp. 1953–1966, Apr. 2018.

[47] T. Yang, S. Liu, H. Liu, M. Tang, X. Tan, and X. Zhou, "Noise benefits parameter estimation in LMMSE sense," *Digit. Signal Process.*, vol. 73, pp. 153–163, Feb. 2018.

[48] F. Duan, Y. Pan, F. Chapeau-Blondeau, and D. Abbott, "Noise benefits in combined nonlinear Bayesian estimators," *IEEE Trans. Signal Process.*, vol. 67, no. 17, pp. 4611–4623, Sep. 2019.

[49] D. Simon, J. Sulam, Y. Romano, Y. M. Lu, and M. Elad, "MMSE approximation for sparse coding algorithms using stochastic resonance," *IEEE Trans. Signal Process.*, vol. 67, no. 17, pp. 4597–4610, Sep. 2019.

[50] X. Liu, H. Wang, Y. Meng, and M. Fu, "Classification of hyperspectral image by CNN based on shadow area enhancement through dynamic stochastic resonance," *IEEE Access*, vol. 7, pp. 134862–134870, 2019.

[51] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[52] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, "Distributed compression-estimation using wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 27–41, Jul. 2006.
[53] M. Shirazi and A. Vosoughi, "On Bayesian fisher information maximization for distributed vector estimation," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 4, pp. 628–645, Dec. 2019.
[54] H. L. von Trees, K. L. Bell, and Z. Tan, *Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*. Hoboken, NJ, USA: Wiley, 2013.
[55] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.
[56] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer-Verlag, 2006.
[57] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1985.
[58] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

JIA LIU was born in China, in 1995. She received the B.Eng. degree in communication engineering from Qingdao University, in 2018, where she is currently pursuing the master's degree in system science. Her research interests include signal estimation and machine learning.

FABING DUAN was born in China, in 1974. He received the master's degree in engineering mechanics from the China University of Mining and Technology, Beijing, in 1999, and the Ph.D. degree in solid mechanics from Zhejiang University, China, in 2002. From 2002 to 2003, he was a Postdoctoral Fellow of the University of Angers, France. Since 2004, he has been with Qingdao University, China, where he is currently a Professor of system science. His research interests include nonlinear systems and signal processing.

FRANÇOIS CHAPEAU-BLONDEAU was born in France, in 1959. He received the Diploma degree in engineering 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, Angers, in 1994. In 1988, he was a Research Associate with the Department of Biophysics, Mayo Clinic, Rochester, MN, USA, working on biomedical ultrasonics. Since 1990, he has been with the University of Angers, where he is currently a Professor of electrical and electronic engineering. His research interests include information theory, signal processing and imaging, and interactions between physics and information sciences.

DEREK ABBOTT (Fellow, IEEE) was born in 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 1997, under the supervision of K. Eshraghian and B. R. Davis.

From 1978 to 1986, he was a Research Engineer with the GEC Hirst Research Centre, London. 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 coauthored *Stochastic Resonance* (Cambridge University Press, 2008), and *Terahertz Imaging for Biomedical Applications* (Springer–Verlag, 2012). His research interests include multidisciplinary physics and electronic engineering applied to complex systems, networks, game theory, energy policy, stochastics, and biophotonics.

Prof. Abbott is a Fellow of the Institute of Physics (IoP), U.K., and an Honorary Fellow of Engineers Australia. He currently serves on the IEEE Publication Services and Products Board (PSPB). He has received a number of awards, including the Tall Poppy Award for Science in 2004, the Australian Research Council Future Fellowship, in 2012, the David Dewhurst Medal, in 2015, the Barry Inglis Medal, in 2018, and the M. A. Sargent Medal for eminence in engineering, in 2019. He has been an Editor and/or Guest Editor of 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 has served on the Editorial Board of the PROCEEDINGS OF THE IEEE, from 2009 to 2014. He has been on the Editorial Board of IEEE ACCESS, since 2015. He currently serves on the Editorial Board of IEEE ACCESS, *Scientific Reports* (Nature), *Royal Society Open Science*, and the *Frontiers of Physics*. He is the Co-Editor of *Quantum Aspects of Life* (Imperial College Press, 2008).