

On the Equivalence of Maximum SNR and MMSE Estimation: Applications to Additive Non-Gaussian Channels and Quantized Observations

Luca Rugini and Paolo Banelli

Abstract

The minimum mean-squared error (MMSE) is one of the most popular criteria for Bayesian estimation. Conversely, the signal-to-noise ratio (SNR) is a typical performance criterion in communications, radar, and generally detection theory. In this paper we first formalize an SNR criterion to design an estimator, and then we prove that there exists an equivalence between MMSE and maximum-SNR estimators, for any statistics. We also extend this equivalence to specific classes of suboptimal estimators, which are expressed by a basis expansion model (BEM). Then, by exploiting an orthogonal BEM for the estimator, we derive the MMSE estimator constrained to a given quantization resolution of the noisy observations, and we prove that this suboptimal MMSE estimator tends to the optimal MMSE estimator that uses an infinite resolution of the observation. Besides, we derive closed-form expressions for the mean-squared error (MSE) and for the SNR of the proposed suboptimal estimators, and we show that these expressions constitute tight, asymptotically exact, bounds for the optimal MMSE and maximum SNR.

Index Terms

Bayesian estimation, maximum SNR, impulsive noise, Laplacian distributions, MMSE, non-Gaussian noise.

L. Rugini and P. Banelli are with the Department of Engineering, University of Perugia, 06125 Perugia, Italy (e-mail: luca.rugini@unipg.it, paolo.banelli@unipg.it)

I. INTRODUCTION

Bayesian estimation of a parameter, a source, or a signal, from noisy observations, is a general framework in statistical inference, with widespread applications in signal processing, communications, controls, machine learning, etc. [1]. The minimum mean-squared error (MMSE) is the most popular criterion in this framework, intuitively connected to the maximum signal-to-noise ratio (MSNR) criterion, mostly used for communication and detection applications [1], [2]. After the first seminal work in [3], the connections between the MMSE and the signal-to-noise ratio (SNR) have attracted several research interests, and there is a quite abundant literature to establish links among them and the mutual information (see [4]– [8] and the references therein). In the context of signal classification (i.e., detection), [9] has shown the interdependencies between the mean-squared error (MSE) and other second-order measures of quality, including many definitions of SNR. However, a thorough investigation of the links between MSE and SNR, in the context of estimation, is still lacking. Some connections between MMSE and SNR have been explored in [3], which proves that the MMSE in the additive noise channel is inversely proportional to the SNR. However, the SNR of [3] is defined at the input of the estimator, while we are interested in the SNR at the output of the estimator.

Motivated to further explore the links between SNR and MSE, in this paper we first define the SNR for the output of a generic estimator, and then we prove the equivalence between the MMSE and MSNR criteria in the context of estimation design. Actually, when the parameter to be estimated and the observations are jointly Gaussian, it is well known that the MMSE estimator, the maximum likelihood (ML) estimator, and the maximum a posteriori (MAP) estimator, are linear in the observation and are equivalent to the MSNR estimator (up to a scalar multiplicative coefficient) [11], [12]: indeed, in this simple Gaussian case, all these estimators produce the same output SNR, which is both maximum and identical to the input SNR. Differently, this paper considers a more general case, where the parameter to be estimated and the observations can be non-Gaussian. In this general case, to the best of our knowledge, the natural question if the MMSE and MSNR estimation criteria are equivalent or not, is still unanswered¹. While classical estimation typically deals with the MMSE criterion, some authors have been looking for an MSNR solution, such as [10], ignoring if this solution has anything to do with the MMSE solution. Specifically, this paper proves that the equivalence between MMSE and MSNR estimators always holds

¹We believe that this question has never been addressed in detail in the context of estimation problems: the investigation done in [9] for detection cannot be extended to estimation, since the SNR definitions used in [9] are quite different from the output SNR considered in this paper.

true, even when the parameter to be estimated and the observations are non-Gaussian: in this case, both the MMSE and the MSNR estimators are usually nonlinear in the observations. This equivalence establishes a strong theoretical link between MMSE and MSNR criteria, traditionally used in different contexts, i.e., estimation and detection, respectively.

Then, we prove that the equivalence between the MSNR and MMSE criteria holds true also for any suboptimal estimator that is expressed by a linear combination of fixed basis functions, according to a basis expansion model (BEM) [13]. Within this framework, we derive the suboptimal MMSE estimator, and other equivalent MSNR estimators, constrained to a given quantization resolution of the noisy observations. Notheworthy, each quantization-constrained estimator corresponds to a specific choice of the set of BEM functions. These quantization-constrained estimators may have practical interest in low-complexity applications that use analog-to-digital (A/D) converters with limited number of bits, such as low-power wireless sensor applications. Specifically, we prove that the suboptimal quantization-constrained MMSE (Q-MMSE) estimator tends to the optimal (unquantized) MMSE estimator that uses an infinite resolution of the observation. In addition, we derive closed-form expressions for the SNR and for the MSE of the proposed suboptimal estimators. Note that these closed-form expressions can be used as lower bounds on the SNR of the MSNR estimators, or as upper bounds on the MSE of the optimal MMSE estimator: indeed, in case of non-Gaussian statistics, analytical expressions for the MMSE value are difficult to obtain [14]; anyway, we also provide some analytical expressions for the MMSE and MSNR values.

To provide an example for practical applications, we apply the derived suboptimal estimators to an additive non-Gaussian noise model, where the noisy observation is simply a signal-plus-noise random variable. We include a numerical example where the signal has a Laplacian statistic, while the noise distribution is a Laplacian mixture, bearing in mind that the results in this paper are valid for any signal and noise statistics. The obtained results show that the proposed suboptimal Q-MMSE and quantization-constrained MSNR (Q-MSNR) estimators outperform other alternative estimators discussed in Section V. The numerical results also confirm and that, when the size of the quantization intervals tends to zero, the MSE (and SNR) of the Q-MMSE estimator tends to the optimal MMSE (and MSNR) value, as expected by design.

The rest of this paper is organized as follows. Section II proves the equivalence between the MSNR and MMSE criteria and discusses several theoretical links. In Section III, we derive the equivalence results for BEM-based estimators, such as the Q-MMSE. Section IV considers the special case of additive non-Gaussian noise channel, while Section V illustrates a numerical example. Section VI concludes the

paper.

II. MAXIMUM SNR AND MMSE ESTIMATORS

For real-valued scalar observation and parameters, Bayesian estimation deals with statistical inference of a random parameter of interest x from a possibly noisy observation y , assuming that the joint probability density function (pdf) $f_{XY}(x, y)$ is known. The estimator of the scalar parameter x is a function $g(\cdot)$ that produces the estimated parameter $\hat{x} = g(y)$. By a linear regression analysis, for any zero-mean x and y and any estimator $g(\cdot)$, it is possible to express the estimator output as

$$\hat{x} = g(y) = K_g x + w_g, \quad (1)$$

where

$$K_g = \frac{E_{XY}\{xg(y)\}}{\sigma_x^2}, \quad (2)$$

$\sigma_x^2 = E_X\{x^2\}$, and w_g is the zero-mean output noise, which is orthogonal to the parameter of interest x and characterized by $\sigma_{w_g}^2 = E_{W_g}\{w_g^2\}$. It is well known that the estimator $g_{\text{MMSE}}(\cdot)$ that minimizes the Bayesian MSE

$$J_g = E_{XY}\{(g(y) - x)^2\} \quad (3)$$

is expressed by [1], [2], [14], [15]

$$g_{\text{MMSE}}(y) = E_{X|Y}\{x|y\} = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx. \quad (4)$$

However, other Bayesian criteria are possible, such as the MAP, the minimum mean-absolute error, etc. [2]. Actually we may choose $g(\cdot)$ that maximizes the SNR at the estimator output in (1), as done for detection in [10], [16]. In this sense, the definition of K_g in (2) leads to the output SNR

$$\gamma_g = \frac{K_g^2 \sigma_x^2}{\sigma_{w_g}^2}, \quad (5)$$

defined as the power ratio of the noise-free signal and the uncorrelated noise in (1). Alternatively, we may maximize the gain K_g in (2) (instead of the SNR), under a power constraint.

Using the orthogonality in (1), the output power is

$$E_Y\{g^2(y)\} = K_g^2 \sigma_x^2 + \sigma_{w_g}^2, \quad (6)$$

and hence, using (2) and (5), we obtain

$$J_g = E_Y\{g^2(y)\} + (1 - 2K_g)\sigma_x^2 \quad (7)$$

$$= (1 - K_g)^2 \sigma_x^2 + \sigma_{w_g}^2. \quad (8)$$

From (5) and (8), it is straightforward that the MSE J_g and the SNR γ_g are linked by

$$J_g = (1 - K_g)^2 \sigma_x^2 + \frac{K_g^2 \sigma_x^2}{\gamma_g}. \quad (9)$$

A. Equivalence of MSNR and MMSE Estimators

While for jointly Gaussian statistics the equivalence between MSNR and MMSE is easy to establish (since the MMSE estimator is linear in y), herein we consider the most general case, without any assumption on the statistics of x and y .

Theorem 1: Among all the possible estimators $g(\cdot)$, the MMSE estimator (4) maximizes the SNR (5) at the estimator output, for any pdf $f_{XY}(x, y)$.

Proof: Let us denote with $g_{\text{MMSE}}(y)$ the MMSE estimator (4), and with $K_{g_{\text{MMSE}}}$ its associated gain (2). In addition, let us denote with $g_{\text{MSNR}}(y)$ an estimator that maximizes the SNR (5), as expressed by

$$g_{\text{MSNR}}(y) = \arg \max_{g(\cdot)} \left[\frac{K_g^2 \sigma_x^2}{E_Y\{g^2(y)\} - K_g^2 \sigma_x^2} \right], \quad (10)$$

and by $K_{g_{\text{MSNR}}}$ its associated gain in (2). This MSNR estimator is not unique, since also any other estimator

$$g_{a,\text{MSNR}}(y) = a g_{\text{MSNR}}(y), \quad (11)$$

with $a \in \mathbb{R} \setminus \{0\}$, maximizes the SNR. Indeed, due to the scaling factor a , by means of (10) both the noise-free power $K_g^2 \sigma_x^2$ and the noise power $\sigma_{w_g}^2 = E_Y\{g^2(y)\} - K_g^2 \sigma_x^2$ are multiplied by the same quantity a^2 , hence the SNR in (5) is invariant with a . By (1) and (2), the gain $K_{g_{a,\text{MSNR}}}$ of $g_{a,\text{MSNR}}(y)$ is equal to

$$K_{g_{a,\text{MSNR}}} = a K_{g_{\text{MSNR}}}. \quad (12)$$

Conversely, the MMSE estimator is unique and has a unique gain $K_{g_{\text{MMSE}}}$. Thus, we have to prove the equivalence of the MMSE estimator $g_{\text{MMSE}}(y)$ with the specific $g_{a,\text{MSNR}}(y)$ characterized by $K_{g_{a,\text{MSNR}}} = K_{g_{\text{MMSE}}}$. Therefore, by (12), we have to choose the MSNR estimator with the specific value $a = \tilde{a}$ expressed by

$$\tilde{a} = \frac{K_{g_{\text{MMSE}}}}{K_{g_{\text{MSNR}}}}. \quad (13)$$

The MSNR estimator $g_{\tilde{a},\text{MSNR}}(y)$ is actually the MSNR estimator that corresponds to an optimization problem restricted to the subclass of all the estimators $g(\cdot)$ characterized by the same gain $K_g = K_{g_{\text{MMSE}}}$, as expressed by

$$g_{\tilde{a},\text{MSNR}}(y) = \arg \max_{g(\cdot), K_g = K_{g_{\text{MMSE}}}} \left[\frac{K_g^2 \sigma_x^2}{E_Y\{g^2(y)\} - K_g^2 \sigma_x^2} \right]. \quad (14)$$

Note that, despite the constraint $K_g = K_{g_{\text{MMSE}}}$, we still obtain the unconstrained MMSE estimator (4), which by definition belongs to the subclass of estimators being characterized by $K_g = K_{g_{\text{MMSE}}}$. Using the constraint $K_g = K_{g_{\text{MMSE}}}$, it is clear in (9) that the dependence of the MSE functional J_g on $g(\cdot)$ is only through γ_g , and no longer also through K_g as in the general case: consequently, the MMSE estimator is

$$\begin{aligned} g_{\text{MMSE}}(y) &= \arg \min_{g(\cdot), K_g = K_{g_{\text{MMSE}}}} [J_g] = \arg \min_{g(\cdot), K_g = K_{g_{\text{MMSE}}}} \begin{bmatrix} \frac{\sigma_x^2}{x} \\ \gamma_g \end{bmatrix} \\ &= \arg \max_{g(\cdot), K_g = K_{g_{\text{MMSE}}}} [\gamma_g] = g_{\bar{a}, \text{MSNR}}(y). \end{aligned} \quad (15)$$

Thus, (15) shows that the estimator that maximizes the SNR with a fixed $K_g = K_{g_{\text{MMSE}}}$ is equivalent to the estimator that minimizes the MSE, i.e., $g_{\bar{a}, \text{MSNR}}(y) = g_{\text{MMSE}}(y)$. ■

Basically, Theorem 1 explains that $\{g_{a, \text{MSNR}}(y)\}$ are all scaled versions of $g_{\text{MMSE}}(y)$. In other words, each scaled version of the MSNR produces the same SNR, but a different MSE: only a unique MSNR estimator is the MMSE estimator, and, in this sense, the two estimation criteria are equivalent.

B. Theoretical Properties of MSNR and MMSE Estimators

Property 1: The output power $E_Y\{g_{\text{MMSE}}^2(y)\}$ of the MMSE estimator (4) is equal to $K_{g_{\text{MMSE}}}\sigma_x^2$. Indeed, from (2) and (4), we obtain

$$\begin{aligned} K_{g_{\text{MMSE}}}\sigma_x^2 &= E_{XY}\{xg_{\text{MMSE}}(y)\} = E_{XY}\{xE_{X|Y}\{x|y\}\} \\ &= \int_{-\infty}^{\infty} E_{X|Y}\{xE_{X|Y}\{x|y\}|y\}f_Y(y)dy \\ &= \int_{-\infty}^{\infty} [E_{X|Y}\{x|y\}]^2 f_Y(y)dy \\ &= E_Y\{g_{\text{MMSE}}^2(y)\}. \end{aligned} \quad (16)$$

Property 2: The MMSE $J_{g_{\text{MMSE}}}$ is equal to $(1 - K_{g_{\text{MMSE}}})\sigma_x^2$. Indeed, from (7) and (16), we obtain

$$\begin{aligned} J_{g_{\text{MMSE}}} &= E_Y\{g_{\text{MMSE}}^2(y)\} + (1 - 2K_{g_{\text{MMSE}}})\sigma_x^2 \\ &= (1 - K_{g_{\text{MMSE}}})\sigma_x^2. \end{aligned} \quad (17)$$

Property 3: The power of the uncorrelated noise term w_g at the output of the MMSE estimator is equal to $K_{g_{\text{MMSE}}}(1 - K_{g_{\text{MMSE}}})\sigma_x^2$. Indeed, from (6), (16), and (17), we obtain

$$\begin{aligned} \sigma_{w_{g_{\text{MMSE}}}}^2 &= E_Y\{g_{\text{MMSE}}^2(y)\} - K_{g_{\text{MMSE}}}^2\sigma_x^2 \\ &= K_{g_{\text{MMSE}}}(1 - K_{g_{\text{MMSE}}})\sigma_x^2 \\ &= K_{g_{\text{MMSE}}}J_{g_{\text{MMSE}}}. \end{aligned} \quad (18)$$

Equation (18) confirms that $K_{g_{\text{MMSE}}} \in [0, 1]$.

Property 4: The MSNR $\gamma_{g_{\text{MSNR}}}$ is equal to $K_{g_{\text{MMSE}}}/(1 - K_{g_{\text{MMSE}}})$. Indeed, from (5) and (18), we obtain

$$\gamma_{g_{\text{MSNR}}} = \gamma_{g_{\text{MMSE}}} = \frac{K_{g_{\text{MMSE}}}^2 \sigma_x^2}{\sigma_{w_{g_{\text{MMSE}}}}^2} = \frac{K_{g_{\text{MMSE}}}}{1 - K_{g_{\text{MMSE}}}}. \quad (19)$$

By (16)–(19), the MSNR is related to the MMSE by

$$\gamma_{g_{\text{MSNR}}} = \gamma_{g_{\text{MMSE}}} = \frac{E_Y\{g_{\text{MMSE}}^2(y)\}}{J_{g_{\text{MMSE}}}} = \frac{\sigma_x^2 - J_{g_{\text{MMSE}}}}{J_{g_{\text{MMSE}}}}. \quad (20)$$

Property 5: The unbiased MMSE (UMMSE) estimator $g_{\text{UMMSE}}(y)$ maximizes the SNR: therefore, the UMMSE estimator is a scaled version of the MMSE estimator, i.e.,

$$g_{\text{UMMSE}}(y) = \frac{g_{\text{MMSE}}(y)}{K_{g_{\text{MMSE}}}}. \quad (21)$$

Indeed, for any estimator $g(y)$, we can make it unbiased by dividing $g(y)$ by K_g , as expressed by

$$\hat{x} = h(y) = \frac{g(y)}{K_g} = x + \frac{w_g}{K_g}. \quad (22)$$

By (1), $h(y) = K_h x + w_h$, therefore $K_h = 1$ and $w_h = w_g/K_g$. Hence, for unbiased estimators, the minimization over $h(\cdot)$ of the MSE $\sigma_{w_h}^2$ is equivalent to the minimization over $g(\cdot)$ of $\sigma_{w_g}^2/K_g^2$, which coincides with the maximization over $g(\cdot)$ of the SNR (5). As a consequence, the UMMSE estimator is the unique MSNR estimator characterized by $K_{g_{\text{MSNR}}} = 1$. Since all MSNR estimators are scaled versions of $g_{\text{MMSE}}(y)$, the unique UMMSE estimator coincides with (21).

Property 6: The MSE $J_{g_{\text{UMMSE}}}$ of the UMMSE estimator is equal to $J_{g_{\text{MMSE}}}/K_{g_{\text{MMSE}}}$. Indeed, from (21), (16), and (17), it is easy to show that

$$J_{g_{\text{UMMSE}}} = \frac{\sigma_{w_{g_{\text{UMMSE}}}}^2}{K_{g_{\text{UMMSE}}}^2} = \frac{1 - K_{g_{\text{MMSE}}}}{K_{g_{\text{MMSE}}}} \sigma_x^2 = \frac{J_{g_{\text{MMSE}}}}{K_{g_{\text{MMSE}}}}. \quad (23)$$

Since $K_{g_{\text{MMSE}}} \leq 1$, then $J_{g_{\text{UMMSE}}} \geq J_{g_{\text{MMSE}}}$.

The Properties 1-6, summarized in Table I, show that all the theoretical expressions for both MMSE and MSNR basically depend on $K_{g_{\text{MMSE}}}$. Since the definition of $K_{g_{\text{MMSE}}}$ in (2) involves a double integration over the joint pdf $f_{XY}(x, y)$, in general the exact value of $K_{g_{\text{MMSE}}}$ is difficult to obtain analytically. Hence, we introduce some suboptimal estimators that allow for an analytical evaluation of their MSE and SNR.

III. SUBOPTIMAL ESTIMATORS

Suboptimal MMSE and MSNR estimators for non-Gaussian statistics are interesting for several reasons. For instance, closed-form computation of the MMSE estimator $g_{\text{MMSE}}(y)$ in (4) may be cumbersome. Furthermore, the optimal MMSE nonlinear function $g_{\text{MMSE}}(y)$ may be too complicated to be implemented

TABLE I
SUMMARY OF THEORETICAL PROPERTIES

#	Meaning	Expression
1	MMSE output power	$E_Y\{g_{\text{MMSE}}^2(y)\} = K_{g_{\text{MMSE}}}\sigma_x^2$
2	MMSE value	$J_{g_{\text{MMSE}}} = (1 - K_{g_{\text{MMSE}}})\sigma_x^2$
3	MMSE output noise	$\sigma_{w_{g_{\text{MMSE}}}}^2 = K_{g_{\text{MMSE}}}(1 - K_{g_{\text{MMSE}}})\sigma_x^2$
4	MSNR value	$\gamma_{g_{\text{MSNR}}} = K_{g_{\text{MMSE}}}/(1 - K_{g_{\text{MMSE}}})$
5	UMMSE estimator	$g_{\text{UMMSE}}(y) = g_{\text{MMSE}}(y)/K_{g_{\text{MMSE}}}$
6	MSE of UMMSE estim.	$J_{g_{\text{UMMSE}}} = J_{g_{\text{MMSE}}}/K_{g_{\text{MMSE}}}$

by low-cost hardware, such as wireless sensors. Additionally, the MMSE $J_{g_{\text{MMSE}}}$ is difficult to compute in closed form. Consequently, a simpler analytical expression for a suboptimal estimator $g(\cdot)$ may permit to compute the associated MSE and SNR, which provide an upper bound on the MMSE and a lower bound on the MSNR, respectively.

Considering a wide class of suboptimal estimators, we assume that $g(\cdot)$ is expressed by a BEM of N known functions $u_i(\cdot)$ and N unknown coefficients g_i :

$$g(y) = \sum_{i=1}^N g_i u_i(y). \quad (24)$$

Each function $u_i(y)$ can be interpreted as a specific (possibly highly suboptimal) estimator, and $g(y)$ in (24) as a linear combination of simpler estimators. We are not interested in the optimization of the basis functions $\{u_i(\cdot)\}$: therefore, the design of $g(\cdot)$ becomes the design of the coefficients $\{g_i\}$. Actually, we have no constraints on the choice of $\{u_i(\cdot)\}$; for instance, saturating or blanking functions, or a mix of them, are typically beneficial to contrast impulsive noise [10], [16]. However, in Section III.C, we will show that an orthogonal design simplifies the computation of $\{g_i\}$, and that the proposed design is general enough for any context.

In the following two subsections, we show that, for BEM-constrained suboptimal estimators (24), the MSNR and MMSE design criteria still continue to be equivalent.

A. B-MSNR Estimators

Herein we derive the MSNR estimators constrained to the BEM (24), denoted as BEM-MSNR (B-MSNR) estimators. By (6) and (24), the SNR γ_g in (5) can be expressed by

$$\gamma_g = \frac{K_g^2 \sigma_x^2}{E_Y\{g^2(y)\} - K_g^2 \sigma_x^2} = \frac{\mathbf{g}^T \boldsymbol{\theta} \boldsymbol{\theta}^T \mathbf{g}}{\mathbf{g}^T (\sigma_x^2 \mathbf{R} - \boldsymbol{\theta} \boldsymbol{\theta}^T) \mathbf{g}}. \quad (25)$$

where

$$\mathbf{g} = [g_1, g_2, \dots, g_N]^T, \quad (26)$$

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_N]^T, \quad (27)$$

$$\theta_i = E_{XY}\{x u_i(y)\}, \quad (28)$$

$$\mathbf{R} = \begin{bmatrix} R_{11} & \cdots & R_{1N} \\ \vdots & \ddots & \vdots \\ R_{N1} & \cdots & R_{NN} \end{bmatrix}, \quad (29)$$

$$R_{ij} = E_Y\{u_i(y) u_j(y)\}. \quad (30)$$

In order to maximize (25), we take the eigenvalue decomposition of the symmetric matrix $\sigma_x^2 \mathbf{R} - \boldsymbol{\theta} \boldsymbol{\theta}^T = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T$, which is assumed to be full rank. Note that \mathbf{U} is orthogonal and $\boldsymbol{\Lambda}$ is diagonal. Then, we express the SNR in (25) as

$$\gamma_g = \frac{\mathbf{v}^T \mathbf{b} \mathbf{b}^T \mathbf{v}}{\mathbf{v}^T \mathbf{v}}, \quad (31)$$

where $\mathbf{v} = \boldsymbol{\Lambda}^{1/2} \mathbf{U}^T \mathbf{g}$ and $\mathbf{b} = \boldsymbol{\Lambda}^{-1/2} \mathbf{U}^T \boldsymbol{\theta}$. The ratio in (31) is maximum [17] when $\mathbf{v} = \mathbf{v}_{\text{B-MSNR}} = c \mathbf{b} = c \boldsymbol{\Lambda}^{-1/2} \mathbf{U}^T \boldsymbol{\theta}$, where $c \in \mathbb{R} \setminus \{0\}$ is an arbitrary constant, and therefore the SNR in (25) is maximum when the estimator is

$$\mathbf{g}_{\text{B-MSNR}} = \mathbf{U} \boldsymbol{\Lambda}^{-1/2} \mathbf{v}_{\text{B-MSNR}} = c (\sigma_x^2 \mathbf{R} - \boldsymbol{\theta} \boldsymbol{\theta}^T)^{-1} \boldsymbol{\theta}. \quad (32)$$

By (25) and (32), using the Sherman-Morrison formula [17], the SNR of B-MSNR estimators is expressed by

$$\gamma_{\text{B-MSNR}} = \boldsymbol{\theta}^T (\sigma_x^2 \mathbf{R} - \boldsymbol{\theta} \boldsymbol{\theta}^T)^{-1} \boldsymbol{\theta} = \frac{\boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}}{\sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}}. \quad (33)$$

B. B-MMSE Estimator

Now we derive the MMSE estimator constrained to the BEM (24), denoted as BEM-MMSE (B-MMSE) estimator. By (24) and (26)–(30), the MSE J_g in (8) becomes

$$J_g = \sigma_x^2 - 2 \mathbf{g}^T \boldsymbol{\theta} + \mathbf{g}^T \mathbf{R} \mathbf{g}. \quad (34)$$

By taking the derivative of (34) with respect to \mathbf{g} and setting it to zero, we obtain the B-MMSE estimator, expressed by

$$\mathbf{g}_{\text{B-MMSE}} = \mathbf{R}^{-1}\boldsymbol{\theta}. \quad (35)$$

By (34) and (35), the MSE of the B-MMSE estimator is

$$J_{g_{\text{B-MMSE}}} = \sigma_x^2 - \mathbf{g}_{\text{B-MMSE}}^T \mathbf{R} \mathbf{g}_{\text{B-MMSE}} = \sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}. \quad (36)$$

Using (36), the SNR (33) can be expressed by

$$\gamma_{\text{B-MSNR}} = \frac{\boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}}{J_{g_{\text{B-MMSE}}}} = \frac{\sigma_x^2 - J_{g_{\text{B-MMSE}}}}{J_{g_{\text{B-MMSE}}}}. \quad (37)$$

The similarity of (37) and (20) suggests a link between B-MMSE and B-MSNR estimators, as shown in Theorem 2.

Theorem 2: The B-MSNR estimator (32) coincides with the B-MMSE estimator (35), when $c = \sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}$.

Proof: Using the Sherman-Morrison formula [17], (32) becomes

$$\mathbf{g}_{\text{B-MSNR}} = \frac{c}{\sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}} \mathbf{R}^{-1} \boldsymbol{\theta}. \quad (38)$$

When $c = \sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}$, $\mathbf{g}_{\text{B-MSNR}}$ in (38) coincides with $\mathbf{g}_{\text{B-MMSE}}$ in (35). ■

Theorem 2 proves that the B-MMSE estimator maximizes the SNR (25) among all the BEM-based estimators: therefore, each B-MSNR estimator is a scaled version of the B-MMSE estimator. Also in this BEM-constrained case the equivalence between B-MMSE and B-MSNR estimators is valid for any statistic of the signal and of the noisy observation.

Note that in Theorem 2 the functions $\{u_i(\cdot)\}$ are arbitrary, but fixed. Differently, if we fix the coefficients $\{g_i\}$ in (24), and perform the optimization over a subset of functions, the equivalence between MMSE and MSNR solutions may not hold true. Indeed, in case of impulsive noise mitigation by means of a soft limiter (SL), expressed by $g_{\text{SL}}(y) = -\beta$ if $y \leq -\beta$, $g_{\text{SL}}(y) = y$ if $-\beta < y < \beta$, and $g_{\text{SL}}(y) = \beta$ if $y \geq \beta$, the optimization over $\beta > 0$ generally produces an MMSE solution [15] that is different from the MSNR solution [16]. Therefore, the equivalence between MMSE and MSNR estimators can be invalid for non-BEM-based suboptimal estimators.

In addition to MMSE, there exist other criteria that maximize the SNR: as shown in Appendix A, the BEM-based unbiased MMSE estimator and a BEM-based estimator that maximizes the gain (2) (subject to a power constraint) both produce the same SNR of B-MMSE and B-MSNR estimators.

C. Q-MMSE Estimator

Herein we prove that, by choosing convenient basis functions $\{u_i(\cdot)\}$ in (24), the B-MMSE estimator (35) converges to the optimal MMSE estimator (4). Indeed, the rectangular disjoint (orthogonal) basis functions

$$u_i(y) = \begin{cases} 1, & \text{if } y_{i-1} < y \leq y_i, \\ 0, & \text{otherwise,} \end{cases} \quad (39)$$

for $i = 1, \dots, N$, with $y_0 = -\infty$ and $y_N = \infty$, greatly simplify the computation of the coefficients $\{g_i\}$. Basically, we are approximating the estimator $g(y)$ by a piecewise-constant function. Using (39), R_{ij} in (30) becomes

$$R_{ij} = \begin{cases} F_Y(y_i) - F_Y(y_{i-1}), & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \quad (40)$$

where $F_Y(y)$ is the cumulative distribution function (cdf) of the observation y . In this case, the matrix \mathbf{R} in (29) is diagonal. Therefore, the coefficients of this specific B-MMSE estimator (35), which we refer to as Q-MMSE estimator, simply become

$$g_{i,\text{Q-MMSE}} = \frac{\theta_i}{R_{ii}}, \quad (41)$$

while the associated MSE (36) is expressed by

$$J_{g_{\text{Q-MMSE}}} = \sigma_x^2 - \sum_{i=1}^N \frac{\theta_i^2}{R_{ii}}. \quad (42)$$

Note that the Q-MMSE estimator (41) can also be interpreted as the MMSE estimator when the observation y has been discretized using N quantization intervals $(y_{i-1}, y_i]$, for $i = 1, \dots, N$. Moreover, we should bear in mind that the number N of quantization levels, as well as the edges of the quantization intervals, are fixed but arbitrary. Thus, the proposed framework finds a natural application when the observed signal undergoes an A/D conversion stage.

However, it is important to prove that, in case of infinite number of quantization levels, the Q-MMSE estimator (41) tends to the optimal MMSE estimator (4) for unquantized observations: hence, the number N of quantization levels enables a tradeoff between performance and complexity.

Theorem 3: When the interval size $\Delta y_i = y_i - y_{i-1}$ tends to zero for $i = 2, \dots, N - 1$, and when y_1 and y_{N-1} tend to $y_0 = -\infty$ and $y_N = \infty$, respectively, then the Q-MMSE estimator (41) tends to the MMSE estimator (4).

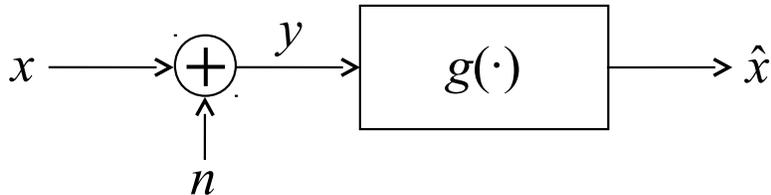


Fig. 1. Signal estimation in additive noise channels.

Proof: When $\Delta y_i \rightarrow 0$, for $i = 2, \dots, N - 1$, from (39) it is easy to show that $f_{X|Y}(x|y)u_i(y) \rightarrow f_{X|Y}(x|y_i)u_i(y)$; hence, for $i = 2, \dots, N - 1$, (28) gives

$$\begin{aligned} \theta_i &= \int_{-\infty}^{\infty} x \int_{y_{i-1}}^{y_i} f_{X|Y}(x|y) f_Y(y) dy dx \xrightarrow{\Delta y_i \rightarrow 0} \\ &\rightarrow f_Y(y_i) \Delta y_i \int_{-\infty}^{\infty} x f_{X|Y}(x|y_i) dx. \end{aligned} \quad (43)$$

In addition, from (39) and (40), we have

$$R_{ii} = \int_{y_{i-1}}^{y_i} f_Y(y) dy \xrightarrow{\Delta y_i \rightarrow 0} f_Y(y_i) \Delta y_i. \quad (44)$$

By taking the ratio between (43) and (44), $g_{i,Q\text{-MMSE}}$ in (41) tends to $g_{\text{MMSE}}(y_i) = E_{X|Y}\{x|y_i\}$ in (4), for $i = 2, \dots, N - 1$. This result can be extended in order to include $i = 1$ and $i = N$ by noting that, when $y_1 \rightarrow y_0 = -\infty$ and $y_{N-1} \rightarrow y_N = \infty$, then $f_{X|Y}(x|y) \rightarrow f_{X|Y}(x|y_1)$ for $y \in (y_0, y_1]$ and $f_{X|Y}(x|y) \rightarrow f_{X|Y}(x|y_{N-1})$ for $y \in (y_{N-1}, y_N)$. ■

Theorem 3 proves that, when the size of the quantization intervals tends to zero, the Q-MMSE estimator converges to the MMSE estimator, regardless of the statistics of the signal of interest x and of the noisy observation y . In particular, the SNR of the Q-MMSE estimator converges to the SNR of the MMSE estimator. Moreover, since a Q-MMSE estimator is a particular B-MMSE estimator, by Theorem 2, the Q-MMSE estimator is also a Q-MSNR estimator, for the same set of quantization thresholds. Noteworthy, if we would optimize the quantization intervals $\{(y_{i-1}, y_i]\}$ [i.e., the functions $\{u_i(\cdot)\}$ in (39)] by keeping the coefficients g_i as fixed, we could end up with different quantization thresholds in an MMSE and MSNR sense.

IV. Q-MMSE IN ADDITIVE NOISE CHANNELS

Herein we provide further insights on the coefficients (41) of the Q-MMSE estimator, when the observations are impaired by an additive noise n , independent from x , as expressed by

$$y = x + n \quad (45)$$

and depicted in Fig. 1. The additive noise model (45) occurs in several applications, especially if the data are obtained by quantized measurements. Indeed, Q-MMSE estimators are particularly useful in realistic scenarios where either the source, or the noise, or both, depart from the standard Gaussian assumption. These scenarios include: (a) additive noise with a high level of impulsiveness [18]– [24]; (b) additive noise whose pdf is a mixture of statistics caused by the random occurrence of different noise sources [25]– [29]; (c) source represented by a pdf mixture, such as in applications (e.g., audio, medical, etc.) that involve effective denoising of sounds or images [30], [31]. The optimal coefficients $\{g_i\}$ obviously depend on the specific pdfs of source and noise, and the numerical results reported in Section V give some evidence of the usefulness of Q-MMSE estimation in an additive non-Gaussian observation model.

According to the BEM model, we assume that the quantization thresholds have been fixed by some criterion. Despite possible criteria for threshold optimization are beyond the scope of this work, in Section V we give some insights about this issue and consider some heuristic solutions.

To specialize the results of Section III to the additive noise model in (45), we observe that the pdf $f_Y(y)$ is the convolution between $f_X(x)$ and $f_N(n)$. Thus, the coefficients θ_i and R_{ii} defined in (28) and (30) can be calculated from the first-order statistics of x and n . Using (45), (28) and (39), we obtain

$$\theta_i = \int_{-\infty}^{\infty} x f_X(x) \int_{y_{i-1}-x}^{y_i-x} f_N(n) dn dx = D(y_i) - D(y_{i-1}), \quad (46)$$

where

$$D(y) = \int_{-\infty}^{\infty} x f_X(x) F_N(y - x) dx. \quad (47)$$

An alternative expression can be obtained by exchanging the integration order, which leads to

$$\theta_i = \int_{-\infty}^{\infty} f_N(n) \int_{y_{i-1}-n}^{y_i-n} x f_X(x) dx dn = D(y_i) - D(y_{i-1}), \quad (48)$$

where

$$D(y) = \int_{-\infty}^{\infty} f_N(n) [(y - n) F_X(y - n) - I_X(y - n)] dn, \quad (49)$$

$$I_X(y) = \int_{-\infty}^y F_X(x) dx. \quad (50)$$

Which expression is preferable, between (47) and (49), depends on the expressions of $f_X(x)$ and $f_N(n)$.

Using (40) and (45), we obtain

$$R_{ii} = F_Y(y_i) - F_Y(y_{i-1}), \quad (51)$$

$$F_Y(y) = \int_{-\infty}^{\infty} f_X(x) F_N(y-x) dx \quad (52)$$

$$= \int_{-\infty}^{\infty} f_N(n) F_X(y-x) dn. \quad (53)$$

Thus, using (41), either (46) or (48), and (51), the Q-MMSE estimator for the additive noise model (45) is expressed by

$$g_{i,\text{Q-MMSE}} = \frac{D(y_i) - D(y_{i-1})}{F(y_i) - F(y_{i-1})}. \quad (54)$$

V. A NUMERICAL EXAMPLE

In this section, we want to numerically compare the MSE and the SNR performances of the Q-MMSE estimator with those of the optimal MMSE estimator, in order to show the usefulness of Q-MMSE estimators with a limited number of quantization levels. Therefore, first we derive the mathematical expressions of the optimal MMSE estimator and of the Q-MMSE estimator, assuming a non-trivial additive noise model (45) where both the signal and the noise are non-Gaussian. Specifically, we model the signal x with a Laplace pdf

$$f_X(x) = \frac{\alpha}{2} e^{-\alpha|x|}, \quad (55)$$

with $\alpha = \sqrt{2}/\sigma_x$, and the noise n with a Laplace mixture pdf

$$f_N(n) = p_0 \frac{\beta_0}{2} e^{-\beta_0|n|} + p_1 \frac{\beta_1}{2} e^{-\beta_1|n|}, \quad (56)$$

with $\{\beta_m = \sqrt{2}/\sigma_{n,m}\}_{m=0,1}$, $R = \sigma_{n,0}^2/\sigma_{n,1}^2$, $\sigma_n^2 = p_0\sigma_{n,0}^2 + p_1\sigma_{n,1}^2$, $p_0 + p_1 = 1$ and $0 \leq p_0 \leq 1$. Basically, (56) models a noise generated by two independent sources: each noise source, characterized by a Laplace pdf with average power $\sigma_{n,m}^2$, occurs with probability p_m . Similar results can be obtained by modeling either the noise, or the signal, or both, as a Gaussian mixture, thus covering a wide range of practical applications of non-Gaussian denoising.

As detailed in Appendix B, direct computation of (4) with (55) and (56) yields the optimal MMSE

estimator

$$g_{\text{MMSE}}(y) = \text{sgn}(y) \times \frac{\sum_{m=0}^1 p_m [C_{1,m}(e^{-\beta_m|y|} - e^{-\alpha|y|}) - C_{2,m}\beta_m|y|e^{-\alpha|y|}]}{\sum_{m=0}^1 p_m C_{2,m} (\alpha e^{-\beta_m|y|} - \beta_m e^{-\alpha|y|})}, \quad (57)$$

$$C_{1,m} = \frac{\alpha^2 \beta_m^2}{(\alpha^2 - \beta_m^2)^2}, \quad C_{2,m} = \frac{\alpha \beta_m}{2(\alpha^2 - \beta_m^2)}. \quad (58)$$

The Q-MMSE estimator can be calculated by solving (47) and (52) using the pdf in (55) and (56): as detailed in Appendix C, when $y > 0$, this calculation leads to

$$D(y) = \sum_{m=0}^1 p_m \frac{\beta_m^2 (3\alpha^2 - \beta_m^2) e^{-\alpha y}}{2\alpha(\alpha^2 - \beta_m^2)^2} - \sum_{m=0}^1 p_m \frac{\alpha^2 \beta_m e^{-\beta_m y}}{(\alpha^2 - \beta_m^2)^2} + \sum_{m=0}^1 p_m \frac{\beta_m^2 y e^{-\alpha y}}{2(\alpha^2 - \beta_m^2)}, \quad (59)$$

$$F_Y(y) = 1 - \sum_{m=0}^1 p_m \frac{\alpha^2 e^{-\beta_m y} - \beta_m^2 e^{-\alpha y}}{2(\alpha^2 - \beta_m^2)}, \quad (60)$$

which inserted into (54) give the final result.

In addition to MMSE and Q-MMSE, other two alternative estimators are included in this comparison: (a) the sampled MMSE (S-MMSE) estimator $g_{i,\text{S-MMSE}}$, obtained by sampling the optimal MMSE estimator $g_{\text{MMSE}}(\cdot)$ at the midpoint of each quantization interval, e.g., $g_{i,\text{S-MMSE}} = g_{\text{MMSE}}((y_{i-1} + y_i)/2)$; and (b) the optimal quantizer (OQ) obtained by applying the Lloyd-Max algorithm [32] to the signal pdf $f_X(x)$. Note that the Lloyd-Max OQ exploits the statistical knowledge of the parameter of interest x only, and neglects the noise, while the Q-MMSE estimator-quantizer also exploits the knowledge of the pdf of noise n : hence, the Q-MMSE estimator is expected to give better performance.

With reference to the choice of the $N - 1$ thresholds $\{y_i\}$ of the Q-MMSE estimators, a heuristic approach chooses all the $N - 1$ thresholds equispaced, such that the overload probability $P_{\text{ol}} = P\{y \in [-\infty, y_1) \cup [y_{N-1}, \infty)\}$ of the quantizer is fixed: this limits the amount of saturating distortion. Another option is to choose the non-uniform thresholds $\{y_i\}$ given by the Lloyd-Max algorithm [32] applied to the signal pdf $f_X(x)$ in (55). For all the quantized estimators, we use the acronym NU for non-uniform quantization and U for uniform quantization.

Fig. 2 compares the shape of the Q-MMSE estimator $g_{\text{Q-MMSE}}(\cdot)$ with the shape of the optimal (unquantized) MMSE estimator $g_{\text{MMSE}}(\cdot)$, when $\sigma_x = 1$, $\sigma_n = 4$, $R = 0.001$, $p_0 = 0.9$, and the $N - 1$ thresholds are equispaced between $y_1 = -10$ and $y_{N-1} = 10$, which induce an overload probability

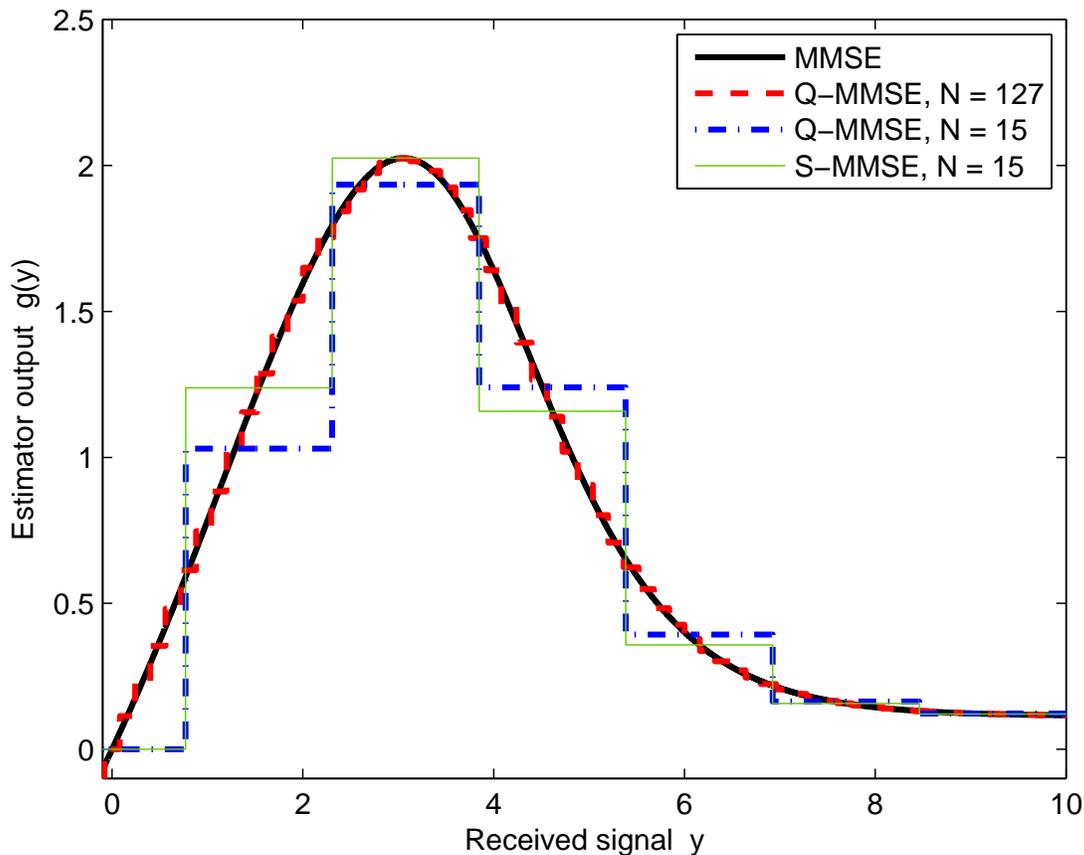


Fig. 2. Comparison between the optimal (unquantized) MMSE estimator and Q-MMSE estimators with uniform quantization (N is the number of intervals).

$P_{o1} \approx 0.0327$. Since all the considered MMSE estimators are odd functions of the input y , Fig. 2 only displays the positive half. Fig. 2 confirms that, when the number N of quantization levels increases, the Q-MMSE estimator tends to the optimal MMSE estimator. Note also that the Q-MMSE estimator $g_{i,Q-MMSE}$ is different from the staircase curve of the S-MMSE estimator $g_{i,S-MMSE}$.

Fig. 3 shows the SNR gain G_g provided by different estimators $g(\cdot)$. The SNR gain G_g is defined as

$$G_g = \frac{\gamma_g}{\sigma_x^2/\sigma_n^2}, \quad (61)$$

where γ_g is the SNR at the output of the estimator, and σ_x^2/σ_n^2 is the SNR at the input of the estimator. The signal and noise parameters are the same of Fig. 2, except for the variable σ_n . Fig. 3 compares the SNR performance of Q-MMSE, S-MMSE, and OQ estimators, assuming uniform and non-uniform

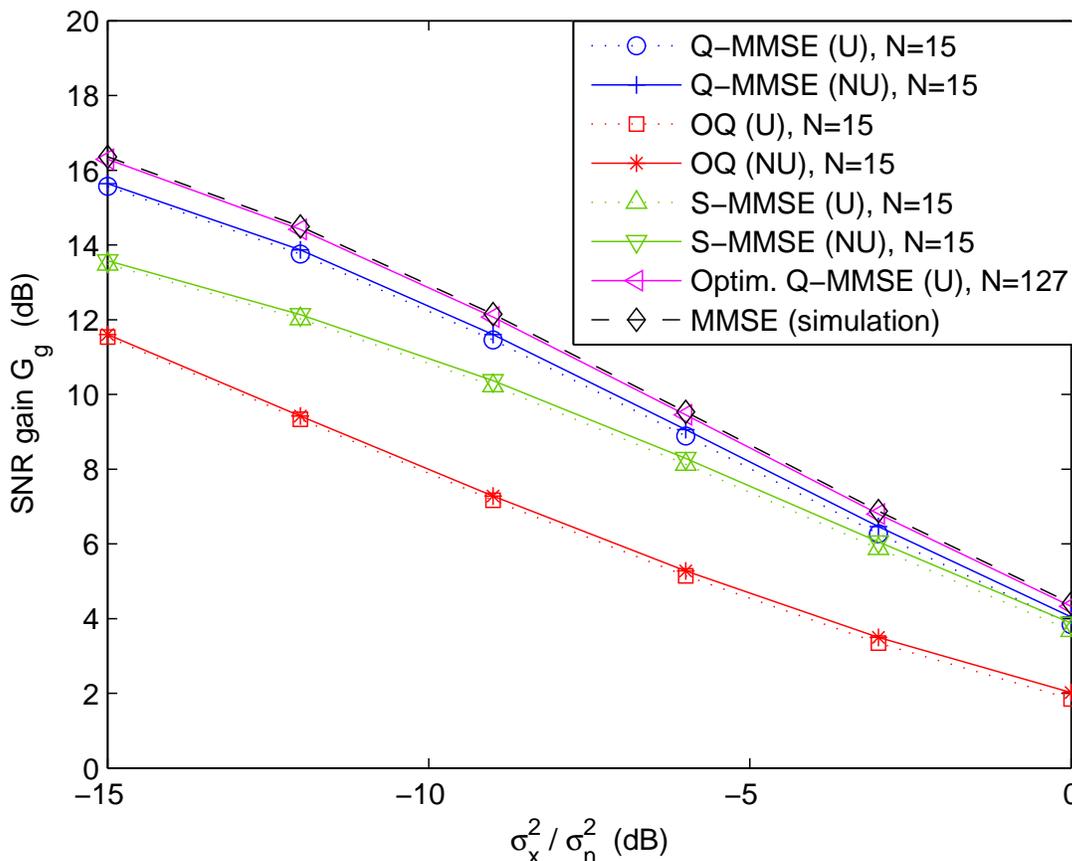


Fig. 3. SNR gain G_g of different estimators as a function of the input SNR σ_x^2/σ_n^2 .

quantization versions (with labels U and NU in the legend of Fig. 3): the overload regions are the same for both versions and have been selected by the Lloyd-Max algorithm, which ends up with an overload probability $P_{ol} \approx 0.0093$ when $\sigma_x^2/\sigma_n^2 = 0$ dB and $P_{ol} \approx 0.0805$ when $\sigma_x^2/\sigma_n^2 = -12$ dB. As a reference, Fig. 3 also includes an optimal Q-MMSE (with $N = 127$) with uniform quantization obtained by an exhaustive maximization of the SNR gain over all the possible choices for the overload regions (i.e., for all the possible choices of $y_1 = -y_{N-1}$): this is equivalent to an optimization of the interval size $\Delta y = (y_{N-1} - y_1)/(N - 2)$ of the uniform quantization intervals. When the number of quantization intervals N is sufficiently high, the SNR of this optimal Q-MMSE estimator basically coincides with the SNR of the optimal (unquantized) MMSE, whose simulated SNR gain is included in Fig. 3 as well.

Fig. 3 confirms that the SNR gain of the Q-MMSE estimator is larger than for the other quantized estimators, provided that the quantization intervals are the same. The SNR of the Q-MMSE estimator

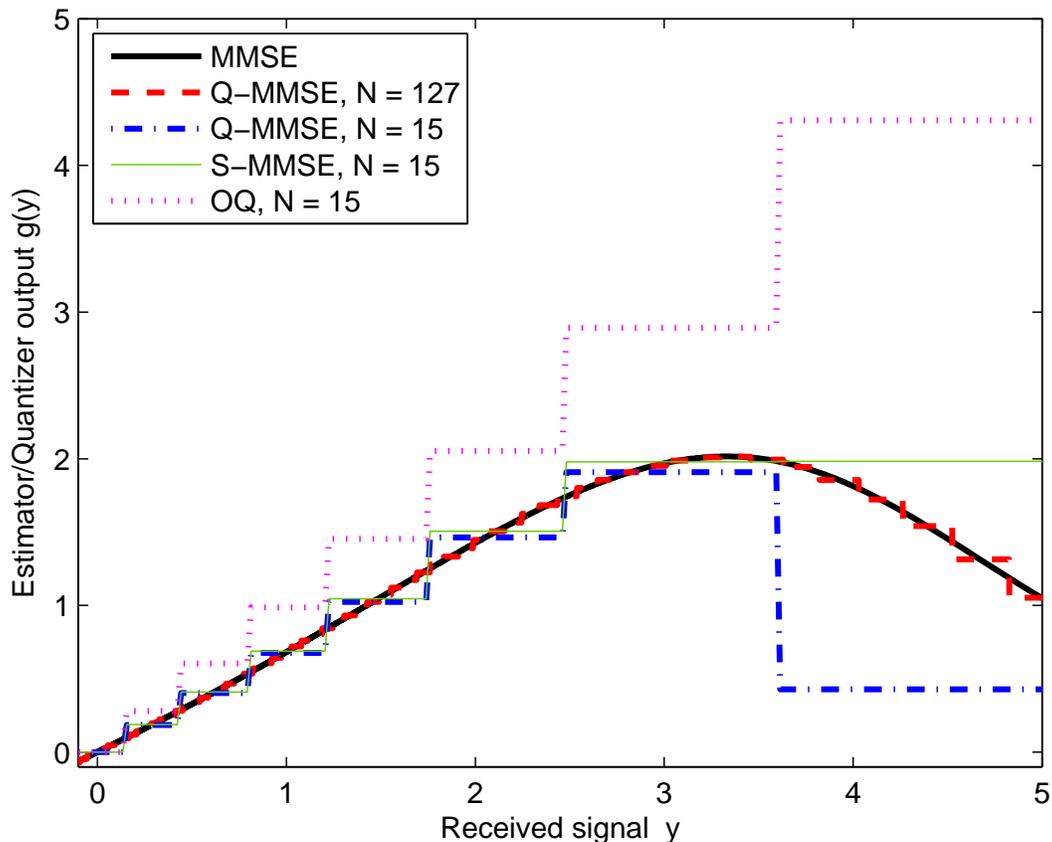


Fig. 4. Comparison between different estimators with non-uniform quantization (N is the number of intervals).

can be further improved by increasing the number of intervals and by optimizing the (uniform) interval sizes, as shown in Fig. 3 by the curve with $N = 127$ with optimized overload regions. In addition, Fig. 3 shows that the SNR of the optimal Q-MMSE estimator is very close to the simulated SNR of the optimal (unquantized) MMSE estimator. Therefore, the proposed Q-MMSE approach permits to obtain analytical tight lower bounds on the SNR of the optimal (unquantized) MMSE estimator.

Fig. 4 compares the function $g(y)$ for the estimators of Fig. 3 with non-uniform quantization, when $\sigma_x^2/\sigma_n^2 = -15$ dB. Fig. 4 highlights that the function $g(y)$ of the Lloyd-Max OQ is nondecreasing, because the noise is neglected; differently, the function $g(y)$ of the (Q-) MMSE estimators can be non-monotonic, like in this specific example.

Fig. 5 displays the MSE of different estimators, in the same scenario of Fig. 3. It is evident that the Q-MMSE estimator provides the lowest MSE among all the quantized estimators that use the same

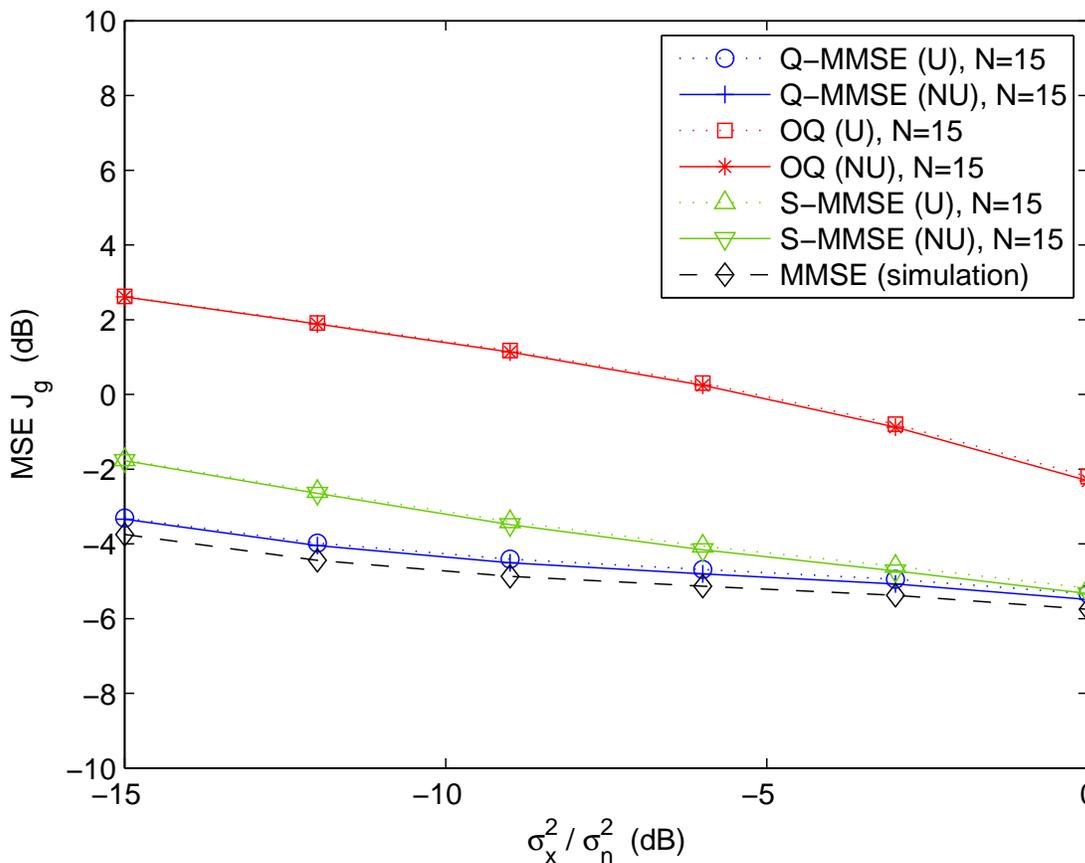


Fig. 5. MSE J_g of different estimators as a function of the input SNR σ_x^2/σ_n^2 .

quantization intervals. Note that the analytical MSE of the Q-MMSE estimator can be used as an upper bound of the minimum value J_{MMSE} (obtained in Fig. 5 by simulation). Similarly to the SNR analysis of Fig. 3, tighter upper bounds on the MMSE $J_{g_{\text{MMSE}}}$ can be obtained by increasing the number of intervals N and by further optimization over all the possible overload regions.

VI. CONCLUSION

In this paper, we have studied a meaningful definition of the MSNR estimator, and we established its equivalence with the MMSE estimator, regardless of the statistics of the noise and of the parameter of interest. We have also extended this equivalence to a specific class of suboptimal estimators expressed as a linear combination of arbitrary (fixed) functions; conversely, we have explained that the same equivalence does not hold true in general for non-BEM suboptimal estimators.

The developed theoretical framework has been instrumental to study Bayesian estimators whose input is a quantized observation of a parameter of interest corrupted by an additive noise. We have shown that, when the size of the quantization intervals goes to zero, the Q-MMSE (Q-MSNR) estimator exactly tends to the MMSE (MSNR) estimator for unquantized observations. Furthermore, by a practical example, we have shown that, using a fairly limited number of quantization levels, the Q-MMSE estimator can easily approach the performance of the optimal (unquantized) MMSE estimator: the designed Q-MMSE estimator, clearly, outperforms in SNR (and in MSE) other suboptimal estimators.

APPENDIX A - OTHER BEM-BASED ESTIMATORS

We detail BEM-based estimators that produce the maximum SNR, similarly to B-MMSE and B-MSNR estimators: unbiased estimators and a maximum-gain estimator.

Unbiased estimators are defined by $E_{Y|X}\{g(y)|x\} = x$ and hence are characterized by $K_g = 1$ in (1). By (2), (24), (26)–(28), for the BEM-based estimators we have

$$K_g = \frac{\mathbf{g}^T \boldsymbol{\theta}}{\sigma_x^2}. \quad (62)$$

Therefore, the BEM-based unbiased MSNR (B-UMSNR) estimator is obtained by maximizing (25) subject to the constraint $\mathbf{g}^T \boldsymbol{\theta} = \sigma_x^2$, while the BEM-based unbiased MMSE (B-UMMSE) estimator is obtained by minimizing (34) subject to the same constraint. By inserting the constraint $\mathbf{g}^T \boldsymbol{\theta} = \sigma_x^2$ into (25) and (34), both optimizations are equivalent to the minimization of the output power $E_Y\{g^2(y)\} = \mathbf{g}^T \mathbf{R} \mathbf{g}$ subject to $\mathbf{g}^T \boldsymbol{\theta} = \sigma_x^2$, which leads to

$$\mathfrak{g}_{\text{B-UMSNR}} = \mathfrak{g}_{\text{B-UMMSE}} = \frac{\sigma_x^2}{\boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}} \mathbf{R}^{-1} \boldsymbol{\theta}. \quad (63)$$

The solution (63) is equivalent to (38) with

$$c = \frac{\sigma_x^2(\sigma_x^2 - \boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta})}{\boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}}. \quad (64)$$

Hence, the B-UMMSE estimator gives the maximum SNR achievable by BEM-based estimators, and is a scaled version of the B-MMSE estimator (35).

An alternative Bayesian criterion is the maximization of the gain K_g (2) or (62), subject to a power constraint. Using the output power constraint $E_Y\{g^2(y)\} = \mathbf{g}^T \mathbf{R} \mathbf{g} = P$, the BEM-based maximum-gain (B-MG) estimator is expressed by

$$\mathfrak{g}_{\text{B-MG}} = \sqrt{\frac{P}{\boldsymbol{\theta}^T \mathbf{R}^{-1} \boldsymbol{\theta}}} \mathbf{R}^{-1} \boldsymbol{\theta}, \quad (65)$$

which is a scaled version of the B-MMSE estimator and hence an MSNR estimator among the BEM-based estimators.

APPENDIX B - DERIVATION OF (57)

Here we show that the computation of (4), when the signal pdf is (55) and the noise pdf is (56), leads to (57). First, using the Bayes' theorem, the MMSE estimator (4) is rewritten as

$$g_{\text{MMSE}}(y) = \frac{\int_{-\infty}^{\infty} x f_{Y|X}(y|x) f_X(x) dx}{f_Y(y)}; \quad (66)$$

in addition, the noise pdf (56) can be rewritten as

$$f_N(n) = \sum_{m=0}^1 p_m f_{N,m}(n) \quad (67)$$

$$f_{N,m}(n) = \frac{\beta_m}{2} e^{-\beta_m |n|}. \quad (68)$$

Using (45), (55), (56), (67), and (68), the numerator of (66), for $y > 0$, can be rewritten as

$$\int_{-\infty}^{\infty} x f_{Y|X}(y|x) f_X(x) dx \quad (69)$$

$$= \int_{-\infty}^{\infty} x f_N(y-x) f_X(x) dx \quad (70)$$

$$= \sum_{m=0}^1 p_m \int_{-\infty}^{\infty} x f_{N,m}(y-x) f_X(x) dx \quad (71)$$

$$= I_1(y) + I_2(y) + I_3(y), \quad (72)$$

where

$$I_1(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m}{4} e^{-\beta_m y} \int_{-\infty}^0 x e^{(\alpha + \beta_m)x} dx, \quad (73)$$

$$I_2(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m}{4} e^{-\beta_m y} \int_0^y x e^{(-\alpha + \beta_m)x} dx, \quad (74)$$

$$I_3(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m}{4} e^{\beta_m y} \int_y^{\infty} x e^{(-\alpha - \beta_m)x} dx. \quad (75)$$

The three integrals (73), (74), and (75), can be solved using

$$\int x e^{ax} dx = \begin{cases} \frac{1}{a^2} [(ax - 1)e^{ax}] + C, & \text{if } a \neq 0, \\ \frac{1}{2} x^2 + C, & \text{if } a = 0, \end{cases} \quad (76)$$

where C is an arbitrary constant. If we assume that $\alpha \neq \beta_m$, for $m = 0, 1$, then (73)–(75) become

$$I_1(y) = - \sum_{m=0}^1 p_m \frac{\alpha \beta_m e^{-\beta_m y}}{4(\alpha + \beta_m)^2}, \quad (77)$$

$$I_2(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m \{[(\beta_m - \alpha)y - 1]e^{-\alpha y} + e^{-\beta_m y}\}}{4(\beta_m - \alpha)^2}, \quad (78)$$

$$I_3(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m [(\alpha + \beta_m)y + 1]e^{-\alpha y}}{4(\alpha + \beta_m)^2}. \quad (79)$$

Hence, the numerator of (66), for $y > 0$, is equal to

$$I_1(y) + I_2(y) + I_3(y) = \quad (80)$$

$$= \sum_{m=0}^1 p_m \left[C_{1,m}(e^{-\beta_m y} - e^{-\alpha y}) - C_{2,m}\beta_m y e^{-\alpha y} \right], \quad (81)$$

where $C_{1,m}$ and $C_{2,m}$ are expressed by (58). If we repeat the same procedure for $y < 0$, we obtain a similar equation.

On the other hand, using (45), (55), (67) and (68), the denominator of (66) is equal to

$$f_Y(y) = f_X(y) * f_N(y) = f_X(y) * \sum_{m=0}^1 p_m f_{N,m}(y) \quad (82)$$

$$= \sum_{m=0}^1 p_m [f_X(y) * f_{N,m}(y)] = \sum_{m=0}^1 p_m f_{Y,m}(y), \quad (83)$$

where $*$ denotes convolution and

$$f_{Y,m}(y) = f_X(y) * f_{N,m}(y). \quad (84)$$

By denoting with $C_X(u)$ the characteristic function associated with the pdf $f_X(x)$, (84) translates into

$$C_{Y,m}(u) = C_X(u)C_{N,m}(u) = \frac{\alpha^2}{\alpha^2 + 4\pi^2 u^2} \frac{\beta_m^2}{\beta_m^2 + 4\pi^2 u^2}. \quad (85)$$

If we assume that $\alpha \neq \beta_m$, for $m = 0, 1$, then (85) can be decomposed in partial fractions as

$$C_{Y,m}(u) = \frac{\beta_m^2}{\beta_m^2 - \alpha^2} \frac{\alpha^2}{\alpha^2 + 4\pi^2 u^2} + \frac{\alpha^2}{\alpha^2 - \beta_m^2} \frac{\beta_m^2}{\beta_m^2 + 4\pi^2 u^2} \quad (86)$$

$$= \frac{\beta_m^2}{\beta_m^2 - \alpha^2} C_X(u) + \frac{\alpha^2}{\alpha^2 - \beta_m^2} C_{N,m}(u), \quad (87)$$

which, by means of (82) and (83), leads to

$$f_{Y,m}(y) = \frac{\beta_m^2}{\beta_m^2 - \alpha^2} f_X(y) + \frac{\alpha^2}{\alpha^2 - \beta_m^2} f_{N,m}(y), \quad (88)$$

$$f_Y(y) = \sum_{m=0}^1 p_m \left[\frac{\beta_m^2}{\beta_m^2 - \alpha^2} f_X(y) + \frac{\alpha^2}{\alpha^2 - \beta_m^2} f_{N,m}(y) \right]. \quad (89)$$

Therefore, by (89), (55), (56), and (67), the denominator of (66) is equal to

$$f_Y(y) = \sum_{m=0}^1 p_m \frac{\alpha \beta_m^2 e^{-\alpha|y|}}{2(\beta_m^2 - \alpha^2)} + \sum_{m=0}^1 p_m \frac{\alpha^2 \beta_m e^{-\beta_m|y|}}{2(\alpha^2 - \beta_m^2)} \quad (90)$$

$$= \sum_{m=0}^1 p_m C_{2,m} (\alpha e^{-\beta_m|y|} - \beta_m e^{-\alpha|y|}), \quad (91)$$

where $C_{2,m}$ is expressed by (58). By inserting (69)–(72), (81), and (90)–(91) into (66), we obtain the mathematical expression of $g_{\text{MMSE}}(y)$ for $y > 0$, and, by repeating the same procedure for negative values of y , we obtain the final expression of $g_{\text{MMSE}}(y)$ reported in (57)–(58), which is valid for all values of y . Note that, since the signal pdf and the noise pdf are both symmetric, the MMSE estimator is an odd function of y , and therefore $g_{\text{MMSE}}(-y) = -g_{\text{MMSE}}(y)$.

APPENDIX C - DERIVATION OF (59) AND (60)

Herein we detail the computation of $D(y)$ in (59) and of $F_Y(y)$ in (60): these two quantities are derived by calculating (47) and (52), respectively, for the additive noise model (45), when the signal pdf is expressed by (55) and the noise pdf is expressed by (56). Indeed, (47) and (52) are necessary in order to compute the Q-MMSE estimator, expressed by (41), via (46)–(47) and (51)–(53). The derivations of $D(y)$ and $F_Y(y)$ are performed only for $y > 0$ (those for $y < 0$ are similar).

By (67) and (68), the noise cdf can be expressed as

$$F_N(n) = \sum_{m=0}^1 p_m F_{N,m}(n) \quad (92)$$

$$F_{N,m}(n) = \begin{cases} \frac{1}{2} e^{\beta_m n}, & \text{if } n < 0, \\ 1 - \frac{1}{2} e^{-\beta_m n}, & \text{if } n \geq 0, \end{cases} \quad (93)$$

and therefore, by (55), $D(y)$ in (47) becomes

$$D(y) = \sum_{m=0}^1 p_m \int_{-\infty}^{\infty} x f_X(x) F_{N,m}(y-x) dx \quad (94)$$

$$= I_4(y) + I_5(y) + I_6(y) + I_7(y) + I_8(y), \quad (95)$$

where

$$I_4(y) = \sum_{m=0}^1 p_m \frac{\alpha}{2} \int_{-\infty}^0 x e^{\alpha x} dx, \quad (96)$$

$$I_5(y) = - \sum_{m=0}^1 p_m \frac{\alpha}{4} e^{-\beta_m y} \int_{-\infty}^0 x e^{(\alpha+\beta_m)x} dx, \quad (97)$$

$$I_6(y) = \sum_{m=0}^1 p_m \frac{\alpha}{2} \int_0^y x e^{-\alpha x} dx, \quad (98)$$

$$I_7(y) = - \sum_{m=0}^1 p_m \frac{\alpha}{4} e^{-\beta_m y} \int_0^y x e^{(-\alpha+\beta_m)x} dx, \quad (99)$$

$$I_8(y) = \sum_{m=0}^1 p_m \frac{\alpha}{4} e^{\beta_m y} \int_y^{\infty} x e^{-(\alpha+\beta_m)x} dx. \quad (100)$$

By assuming $\alpha \neq \beta_m$, for $m = 0, 1$, and by solving the five integrals in (96)–(100) using (76), it is easy to show that $D(y)$ in (94) becomes equal to (59).

The cdf $F_Y(y)$ can be easily calculated from (88) and (93), which lead to

$$F_{Y,m}(y) = \frac{\beta_m^2}{\beta_m^2 - \alpha^2} F_X(y) + \frac{\alpha^2}{\alpha^2 - \beta_m^2} F_{N,m}(y) \quad (101)$$

$$= 1 + \frac{\beta_m^2 e^{-\alpha y} - \alpha^2 e^{-\beta_m y}}{2(\alpha^2 - \beta_m^2)}, \quad (102)$$

where we have used $F_X(y) = 1 - \frac{1}{2}e^{-\alpha y}$ for $y > 0$. Using (102) with (82)–(83), we obtain the final expression (60).

REFERENCES

- [1] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [2] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY: Springer-Verlag, 1994.
- [3] D. Guo, S. Shamai, and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1283, Apr. 2005.
- [4] D. Guo, S. Shamai, and S. Verdú, “Mutual information and conditional mean estimation in Poisson channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1837–1849, May 2008.
- [5] D. Guo, Y. Wu, S. Shamai, and S. Verdú, “Estimation in Gaussian noise: Properties of the minimum mean-square error,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.
- [6] K. Venkat and T. Weissman, “Pointwise relations between information and estimation in Gaussian noise,” *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6264–6281, Oct. 2012.
- [7] Y. Wu and S. Verdú, “Functional properties of minimum mean-square error and mutual information,” *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1289–1301, Mar. 2012.
- [8] Y. Yang and R. S. Blum, “MIMO radar waveform design based on mutual information and minimum mean-square error estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, Jan. 2007.

- [9] W. Gardner, "A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. 28, no. 6, pp. 807–816, June 1980.
- [10] S. V. Zhidkov, "Performance analysis and optimization of OFDM receiver with blanking nonlinearity in impulsive noise environment," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 234–242, Jan. 2006.
- [11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [12] E. Akyol, K. B. Viswanatha, and K. Rose, "On conditions for linearity of optimal estimation," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3497–3508, June 2012.
- [13] G. B. Giannakis and C. Tepedelenlioğlu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. of the IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.
- [14] J. T. Flâm, S. Chatterjee, K. Kansanen, and T. Ekman, "On MMSE estimation: A linear model under Gaussian mixture statistics," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3840–3845, July 2012.
- [15] P. Banelli, "Bayesian estimation of a Gaussian source in Middleton's class-A impulsive noise," *IEEE Signal Process. Lett.*, vol. 20, no. 10, pp. 956–959, Oct. 2013.
- [16] S. V. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 5–9, Jan. 2008.
- [17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD: Johns Hopkins Univ. Press, 2013.
- [18] J. Friedmann, H. Messer, and J. Cardoso, "Robust parameter estimation of a deterministic signal in impulsive noise," *IEEE Trans. Signal Process.*, vol. 48, no. 4, pp. 935–942, Apr. 2000.
- [19] G. Ndo, P. Siohan, and M. H. Hamon, "Adaptive noise mitigation in impulsive environment: application to power-line communications," *IEEE Trans. Power Del.*, vol. 25, no. 2, pp. 647–656, Apr. 2010.
- [20] C. H. Yih, "Iterative interference cancellation for OFDM signals with blanking nonlinearity in impulsive noise channels," *IEEE Signal Process. Lett.*, vol. 19, no. 3, pp. 147–150, Mar. 2012.
- [21] D. F. Tseng, Y. S. Han, W. H. Mow, and L. C. Chang, "Robust clipping for OFDM transmissions over memoryless impulsive noise channels," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1110–1113, July 2012.
- [22] E. Alsusa and K. M. Rabie, "Dynamic peak-based threshold estimation method for mitigating impulsive noise in power-line communication systems," *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2201–2208, Oct. 2013.
- [23] F. H. Juwono, Q. Guo, D. Huang, and K. P. Wong, "Deep clipping for impulsive noise mitigation in OFDM-based power-line communications," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1335–1343, June 2014.
- [24] F. H. Juwono, Q. Guo, Y. Chen, L. Xu, D. Huang, and K. P. Wong, "Linear combining of nonlinear preprocessors for OFDM-based power-line communications," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 253–260, Jan. 2016.
- [25] T. C. Chuah, B. S. Sharif, and O. R. Hinton, "Nonlinear decorrelator for multiuser detection in non-Gaussian impulsive environments," *IET Electron. Lett.*, vol. 36, no. 10, pp. 920–922, 11 May 2000.
- [26] H. Deliç and A. Hocanin, "Robust detection in DS-CDMA," *IEEE Trans. Veh. Technol.*, vol. 51, no. 1, pp. 155–170, Jan. 2002.
- [27] N. Güney, H. Deliç and M. Koca, "Robust detection of ultra-wideband signals in non-Gaussian noise," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 4, pp. 1724–1730, June 2006.
- [28] N. C. Beaulieu and B. Hu, "Soft-limiting receiver structures for time-hopping UWB in multiple-access interference," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 810–818, Mar. 2008.
- [29] N. C. Beaulieu and S. Niranjayan, "UWB receiver designs based on a Gaussian-Laplacian noise-plus-MAI model," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 997–1006, Mar. 2010.

- [30] H. Rabbani, R. Nezafat, and S. Gazor, "Wavelet-domain medical image denoising using bivariate Laplacian mixture model," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2826–2837, Dec. 2009.
- [31] H. Rabbani and S. Gazor, "Image denoising employing local mixture models in sparse domains," *IET Image Process.*, vol. 4, no. 5, pp. 413–428, 2010.
- [32] A. Gersho, "Quantization," *IEEE Commun. Mag.*, vol. 15, no. 5, pp. 16–29, Sept. 1977.