

# CWCU LMMSE Estimation under Linear Model Assumptions

O. Lang, M. Huemer

Institute of Signal Processing, Johannes Kepler University Linz  
oliver.lang@jku.at, mario.huemer@jku.at

**Abstract.** The classical unbiasedness condition utilized e.g. by the best linear unbiased estimator (BLUE) is very stringent. By softening the "global" unbiasedness condition and introducing component-wise conditional unbiasedness conditions instead, the number of constraints limiting the estimator's performance can in many cases significantly be reduced. In this paper we extend the findings on the component-wise conditionally unbiased linear minimum mean square error (CWCU LMMSE) estimator under linear model assumptions. We discuss the CWCU LMMSE estimator for complex proper Gaussian parameter vectors, and for mutually independent (and otherwise arbitrarily distributed) parameters. Finally, the beneficial properties of the CWCU LMMSE estimator are demonstrated in two applications.

## 1 Introduction

Usually, when we talk about unbiased estimation of a parameter vector  $\mathbf{x} \in \mathbb{C}^{n \times 1}$  out of a measurement vector  $\mathbf{y} \in \mathbb{C}^{m \times 1}$ , then the estimation problem is treated in the classical framework [1]. Letting  $\hat{\mathbf{x}} = \mathbf{g}(\mathbf{y})$  be an estimator of  $\mathbf{x}$ , then the classical unbiased constraint asserts that

$$E_{\mathbf{y}}[\hat{\mathbf{x}}] = \int \mathbf{g}(\mathbf{y})p(\mathbf{y}; \mathbf{x})d\mathbf{y} = \mathbf{x} \quad \text{for all possible } \mathbf{x}, \quad (1)$$

where  $p(\mathbf{y}; \mathbf{x})$  is the probability density function (PDF) of vector  $\mathbf{y}$  parametrized by the unknown parameter vector  $\mathbf{x}$ . The index of the expectation operator shall indicate the PDF over which the averaging is performed. Eq. (1) can also be formulated in the Bayesian framework, where the parameter vector  $\mathbf{x}$  is treated as random, and whose realization is to be estimated. Here, the corresponding problem arises by demanding global conditional unbiasedness, i.e.

$$E_{\mathbf{y}|\mathbf{x}}[\hat{\mathbf{x}}|\mathbf{x}] = \int \mathbf{g}(\mathbf{y})p(\mathbf{y}|\mathbf{x})d\mathbf{y} = \mathbf{x} \quad \text{for all possible } \mathbf{x}. \quad (2)$$

The attribute *global* indicates that the condition is made on the whole parameter vector  $\mathbf{x}$ . However, the constricting requirement in (2) prevents the exploitation

---

This work was supported by the Austrian Science Fund (FWF): I683-N13.

of prior knowledge about the parameters, and hence leads to a significant reduction in the benefits brought about by the Bayesian framework.

In component-wise conditionally unbiased (CWCU) Bayesian parameter estimation [2]-[5], instead of constraining the estimator to be globally unbiased, we aim for achieving conditional unbiasedness on one parameter component at a time. Let  $x_i$  be the  $i^{\text{th}}$  element of  $\mathbf{x}$ , and  $\hat{x}_i = g_i(\mathbf{y})$  be an estimator of  $x_i$ . Then the CWCU constraints are

$$E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = \int g_i(\mathbf{y})p(\mathbf{y}|x_i)d\mathbf{y} = x_i, \quad (3)$$

for all possible  $x_i$  (and all  $i = 1, 2, \dots, n$ ). The CWCU constraints are less stringent than the global conditional unbiasedness condition in (2), and it turns out that a CWCU estimator in many cases allows the incorporation of prior knowledge about the statistical properties of the parameter vector.

The paper is organized as follows: In Section 2 we discuss the CWCU linear minimum mean square error (LMMSE) estimator under different linear model assumptions, and we extend the findings of [2]. We particularly distinguish between complex proper jointly Gaussian (cf. [6]), and mutually independent (and otherwise arbitrarily distributed) parameters. Then, in Section 3 the CWCU LMMSE estimator is compared against the best linear unbiased estimator (BLUE) and the LMMSE estimator in two different applications.

## 2 CWCU LMMSE estimation

We assume that a complex vector parameter  $\mathbf{x} \in \mathbb{C}^{n \times 1}$  is to be estimated based on a measurement vector  $\mathbf{y} \in \mathbb{C}^{m \times 1}$ . Additionally, we assume that  $\mathbf{x}$  and  $\mathbf{y}$  are connected via a linear model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (4)$$

where  $\mathbf{H} \in \mathbb{C}^{m \times n}$  is a known observation matrix,  $\mathbf{x}$  has mean  $E_{\mathbf{x}}[\mathbf{x}]$  and covariance matrix  $\mathbf{C}_{\mathbf{x}\mathbf{x}}$ , and  $\mathbf{n} \in \mathbb{C}^{m \times 1}$  is a zero mean noise vector with covariance matrix  $\mathbf{C}_{\mathbf{nn}}$  and independent of  $\mathbf{x}$ . Additional assumptions on  $\mathbf{x}$  will vary in the following. We note that the CWCU LMMSE estimator for the linear model under the assumption of complex proper Gaussian  $\mathbf{x}$  and complex and proper white Gaussian noise with covariance matrix  $\mathbf{C}_{\mathbf{nn}} = \sigma_n^2 \mathbf{I}$  has already been derived in [2].

As in LMMSE estimation we constrain the estimator to be linear (or actually affine), such that

$$\hat{\mathbf{x}} = \mathbf{E}\mathbf{y} + \mathbf{c}, \quad (5)$$

with  $\mathbf{E} \in \mathbb{C}^{n \times m}$  and  $\mathbf{c} \in \mathbb{C}^{n \times 1}$ . Note that in LMMSE estimation no assumptions on the specific form of the PDF  $p(\mathbf{x})$  have to be made. However, the situation is different in CWCU LMMSE estimation as will be shown shortly. Let us consider the  $i^{\text{th}}$  component of the estimator

$$\hat{x}_i = \mathbf{e}_i^H \mathbf{y} + c_i, \quad (6)$$

where  $\mathbf{e}_i^H$  denotes the  $i^{\text{th}}$  row of the estimator matrix  $\mathbf{E}$ . Furthermore, let  $\mathbf{h}_i \in \mathbb{C}^{m \times 1}$  be the  $i^{\text{th}}$  column of  $\mathbf{H}$ ,  $\bar{\mathbf{H}}_i \in \mathbb{C}^{m \times (n-1)}$  the matrix resulting from  $\mathbf{H}$  by deleting  $\mathbf{h}_i$ ,  $x_i$  be the  $i^{\text{th}}$  element of  $\mathbf{x}$ , and  $\bar{\mathbf{x}}_i \in \mathbb{C}^{(n-1) \times 1}$  the vector resulting from  $\mathbf{x}$  after deleting  $x_i$ . Then we can write  $\mathbf{y} = \mathbf{h}_i x_i + \bar{\mathbf{H}}_i \bar{\mathbf{x}}_i + \mathbf{n}$ , and (6) becomes

$$\hat{x}_i = \mathbf{e}_i^H (\mathbf{h}_i x_i + \bar{\mathbf{H}}_i \bar{\mathbf{x}}_i + \mathbf{n}) + c_i. \quad (7)$$

The conditional mean of  $\hat{x}_i$  therefore is

$$E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = \mathbf{e}_i^H \mathbf{h}_i x_i + \mathbf{e}_i^H \bar{\mathbf{H}}_i E_{\bar{\mathbf{x}}_i|x_i}[\bar{\mathbf{x}}_i|x_i] + c_i. \quad (8)$$

From (8) we can derive conditions that guarantee that the CWCU constraints (3) are fulfilled. There are at least the following possibilities:

1. (3) can be fulfilled for all possible  $x_i$  if the conditional mean  $E_{\bar{\mathbf{x}}_i|x_i}[\bar{\mathbf{x}}_i|x_i]$  is a linear function of  $x_i$ . For complex proper Gaussian  $\mathbf{x}$  this condition holds (for all  $i = 1, 2, \dots, n$ ).
2. (3) can be fulfilled for all possible  $x_i$  (and all  $i = 1, 2, \dots, n$ ) if  $E_{\bar{\mathbf{x}}_i|x_i}[\bar{\mathbf{x}}_i|x_i] = E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i]$  for all possible  $x_i$  (and all  $i = 1, 2, \dots, n$ ), which is true if the elements  $x_i$  of  $\mathbf{x}$  are mutually independent.
3. (3) is fulfilled for all possible  $x_i$  (and all  $i = 1, 2, \dots, n$ ) if  $\mathbf{e}_i^H \mathbf{h}_i = 1$ ,  $\mathbf{e}_i^H \bar{\mathbf{H}}_i = \mathbf{0}^T$ , and  $c_i = 0$  for  $i = 1, 2, \dots, n$ . These constraints and settings correspond to the ones of the BLUE.

## 2.1 Complex proper Gaussian parameter vectors

We start with the first case from above, assume a complex proper Gaussian parameter vector, i.e.  $\mathbf{x} \sim \mathcal{CN}(E_{\mathbf{x}}[\mathbf{x}], \mathbf{C}_{\mathbf{x}\mathbf{x}})$ , and start with the derivation of the  $i^{\text{th}}$  component  $\hat{x}_i$  of the estimator. Because of the Gaussian assumption we have  $E_{\bar{\mathbf{x}}_i|x_i}[\bar{\mathbf{x}}_i|x_i] = E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i] + (\sigma_{x_i}^2)^{-1} \mathbf{C}_{\bar{\mathbf{x}}_i x_i} (x_i - E_{x_i}[x_i])$ , where  $\mathbf{C}_{\bar{\mathbf{x}}_i x_i} = E_{\mathbf{x}}[(\bar{\mathbf{x}}_i - E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i])(x_i - E_{x_i}[x_i])^H]$ , and  $\sigma_{x_i}^2$  is the variance of  $x_i$ . Consequently (8) becomes

$$E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = \mathbf{e}_i^H \mathbf{h}_i x_i + \mathbf{e}_i^H \bar{\mathbf{H}}_i (E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i] + (\sigma_{x_i}^2)^{-1} \mathbf{C}_{\bar{\mathbf{x}}_i x_i} (x_i - E_{x_i}[x_i])) + c_i. \quad (9)$$

Note that the only requirement on the noise vector so far was its independence on  $\mathbf{x}$ . From (9) we see that  $E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = x_i$  is fulfilled if

$$\mathbf{e}_i^H \mathbf{h}_i + \mathbf{e}_i^H \bar{\mathbf{H}}_i (\sigma_{x_i}^2)^{-1} \mathbf{C}_{\bar{\mathbf{x}}_i x_i} = 1 \quad (10)$$

and

$$c_i = \mathbf{e}_i^H \bar{\mathbf{H}}_i (\sigma_{x_i}^2)^{-1} \mathbf{C}_{\bar{\mathbf{x}}_i x_i} E_{x_i}[x_i] - \mathbf{e}_i^H \bar{\mathbf{H}}_i E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i]. \quad (11)$$

With (10), (11) can be reformulated according to

$$\begin{aligned} c_i &= E_{x_i}[x_i] - \mathbf{e}_i^H \mathbf{h}_i E_{x_i}[x_i] - \mathbf{e}_i^H \bar{\mathbf{H}}_i E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i] \\ &= E_{x_i}[x_i] - \mathbf{e}_i^H \mathbf{H} E_{\mathbf{x}}[\mathbf{x}]. \end{aligned} \quad (12)$$

Furthermore, (10) can be simplified to obtain the constraint

$$\mathbf{e}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}\mathbf{x}} = \sigma_{x_i}^2. \quad (13)$$

Inserting (6), (12) and (13) into the Bayesian MSE cost function  $E_{\mathbf{y},\mathbf{x}}[|\hat{x}_i - x_i|^2]$  immediately leads to the constrained optimization problem

$$\mathbf{e}_{\text{CL},i} = \arg \min_{\mathbf{e}_i} (\mathbf{e}_i^H (\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})\mathbf{e}_i - \sigma_{x_i}^2) \quad \text{s.t.} \quad \mathbf{e}_i^H \mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H \mathbf{e}_i = \sigma_{x_i}^2, \quad (14)$$

where "CL" shall stand for CWCU LMMSE. The solution can be found with the Lagrange multiplier method and is given by

$$\mathbf{e}_{\text{CL},i}^H = \frac{\sigma_{x_i}^2}{\mathbf{C}_{x_i\mathbf{x}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H} \mathbf{C}_{x_i\mathbf{x}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}. \quad (15)$$

Introducing the estimator matrix  $\mathbf{E}_{\text{CL}} = [\mathbf{e}_{\text{CL},1}, \mathbf{e}_{\text{CL},2}, \dots, \mathbf{e}_{\text{CL},n}]^H$  together with (12) and (15) immediately leads us to the first part of the

**Proposition 1** *If the observed data  $\mathbf{y}$  follow the linear model in (4), where  $\mathbf{y} \in \mathbb{C}^{m \times 1}$  is the data vector,  $\mathbf{H} \in \mathbb{C}^{m \times n}$  is a known observation matrix,  $\mathbf{x} \in \mathbb{C}^{n \times 1}$  is a parameter vector with prior complex proper Gaussian PDF  $\mathcal{CN}(E_{\mathbf{x}}[\mathbf{x}], \mathbf{C}_{\mathbf{xx}})$ , and  $\mathbf{n} \in \mathbb{C}^{m \times 1}$  is a zero mean noise vector with covariance matrix  $\mathbf{C}_{\mathbf{nn}}$  and independent of  $\mathbf{x}$  (the PDF of  $\mathbf{n}$  is otherwise arbitrary), then the CWCU LMMSE estimator minimizing the Bayesian MSEs  $E_{\mathbf{y},\mathbf{x}}[|\hat{x}_i - x_i|^2]$  under the constraints  $E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = x_i$  for  $i = 1, 2, \dots, n$  is given by*

$$\hat{\mathbf{x}}_{\text{CL}} = E_{\mathbf{x}}[\mathbf{x}] + \mathbf{E}_{\text{CL}}(\mathbf{y} - \mathbf{H}E_{\mathbf{x}}[\mathbf{x}]), \quad (16)$$

with

$$\mathbf{E}_{\text{CL}} = \mathbf{D}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}, \quad (17)$$

where the elements of the real diagonal matrix  $\mathbf{D}$  are

$$[\mathbf{D}]_{i,i} = \frac{\sigma_{x_i}^2}{\mathbf{C}_{x_i\mathbf{x}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H}. \quad (18)$$

The mean of the error  $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}_{\text{CL}}$  (in the Bayesian sense) is zero, and the error covariance matrix  $\mathbf{C}_{\mathbf{ee},\text{CL}}$  which is also the minimum Bayesian MSE matrix  $\mathbf{M}_{\hat{\mathbf{x}}_{\text{CL}}}$  is

$$\mathbf{C}_{\mathbf{ee},\text{CL}} = \mathbf{M}_{\hat{\mathbf{x}}_{\text{CL}}} = \mathbf{C}_{\mathbf{xx}} - \mathbf{A}\mathbf{D} - \mathbf{D}\mathbf{A} + \mathbf{D}\mathbf{A}\mathbf{D}, \quad (19)$$

with  $\mathbf{A} = \mathbf{C}_{\mathbf{xx}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}\mathbf{H}\mathbf{C}_{\mathbf{xx}}$ . The minimum Bayesian MSEs are  $\text{Bmse}(\hat{x}_{\text{CL},i}) = [\mathbf{M}_{\hat{\mathbf{x}}_{\text{CL}}}]_{i,i}$ .

The part on the error performance can simply be proved by inserting in the definition of  $\mathbf{e}$  and  $\mathbf{C}_{\mathbf{ee}}$ , respectively. From (17) it can be seen that the CWCU LMMSE estimator matrix can be derived as the product of the diagonal matrix  $\mathbf{D}$  with the LMMSE estimator matrix  $\mathbf{E}_{\text{L}} = \mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1} = \mathbf{C}_{\mathbf{xx}}\mathbf{H}^H(\mathbf{H}\mathbf{C}_{\mathbf{xx}}\mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}$ . Furthermore, we have  $E_{\mathbf{y}|x_i}[\hat{x}_{\text{L},i}|x_i] = [\mathbf{D}]_{i,i}^{-1}x_i + (1 - [\mathbf{D}]_{i,i}^{-1})E_{x_i}[x_i]$  for the LMMSE estimator. From (18) it also follows that

$$\mathbf{D} = \text{diag}\{\mathbf{C}_{\mathbf{xx}}\}(\text{diag}\{\mathbf{A}\})^{-1}. \quad (20)$$

The CWCU LMMSE estimator will in general not commute over linear transformations, an exception is the transformation over a diagonal matrix as partly discussed in [5].

## 2.2 Complex parameter vectors with mutually independent elements

In case the elements of the parameter vector are mutually independent (8) becomes

$$E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = \mathbf{e}_i^H \mathbf{h}_i x_i + \mathbf{e}_i^H \bar{\mathbf{H}}_i E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i] + c_i. \quad (21)$$

$E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = x_i$  is fulfilled if  $\mathbf{e}_i^H \mathbf{h}_i = 1$  and  $c_i = -\mathbf{e}_i^H \bar{\mathbf{H}}_i E_{\bar{\mathbf{x}}_i}[\bar{\mathbf{x}}_i]$ . No further assumptions on the PDF of  $\mathbf{x}$  are required. Following similar arguments as above again leads to a constrained optimization problem [5]. Solving it leads to

**Proposition 2** *If the observed data  $\mathbf{y}$  follow the linear model in (4), where  $\mathbf{y} \in \mathbb{C}^{m \times 1}$  is the data vector,  $\mathbf{H} \in \mathbb{C}^{m \times n}$  is a known observation matrix,  $\mathbf{x} \in \mathbb{C}^{n \times 1}$  is a parameter vector with mean  $E_{\mathbf{x}}[\mathbf{x}]$ , mutually independent elements and covariance matrix  $\mathbf{C}_{\mathbf{xx}} = \text{diag}\{\sigma_{x_1}^2, \sigma_{x_2}^2, \dots, \sigma_{x_n}^2\}$ ,  $\mathbf{n} \in \mathbb{C}^{m \times 1}$  is a zero mean noise vector with covariance matrix  $\mathbf{C}_{\mathbf{nn}}$  and independent of  $\mathbf{x}$  (the PDF of  $\mathbf{n}$  is otherwise arbitrary), then the CWCU LMMSE estimator minimizing the Bayesian MSEs  $E_{\mathbf{y},\mathbf{x}}[|\hat{x}_i - x_i|^2]$  under the constraints  $E_{\mathbf{y}|x_i}[\hat{x}_i|x_i] = x_i$  for  $i = 1, 2, \dots, n$  is given by (16) and (17), where the elements of the real diagonal matrix  $\mathbf{D}$  are*

$$[\mathbf{D}]_{i,i} = \frac{1}{\sigma_{x_i}^2 \mathbf{h}_i^H (\mathbf{H} \mathbf{C}_{\mathbf{xx}} \mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1} \mathbf{h}_i}. \quad (22)$$

Since for mutually independent parameters the  $i^{\text{th}}$  row of the LMMSE estimator is  $\mathbf{e}_{L,i}^H = \sigma_{x_i}^2 \mathbf{h}_i^H (\mathbf{H} \mathbf{C}_{\mathbf{xx}} \mathbf{H}^H + \mathbf{C}_{\mathbf{nn}})^{-1}$  it follows from (22) that

$$[\mathbf{D}]_{i,i} = (\mathbf{e}_{L,i}^H \mathbf{h}_i)^{-1}. \quad (23)$$

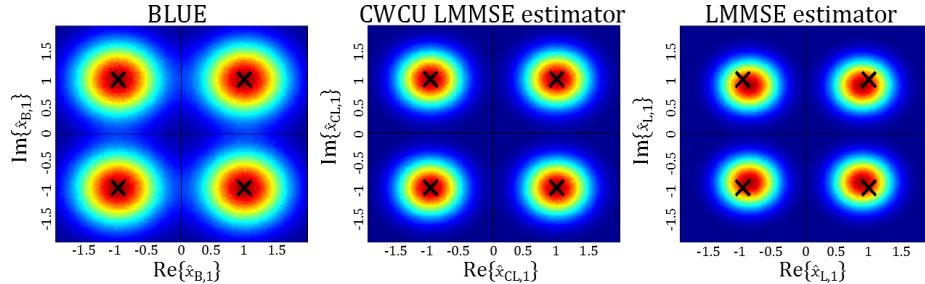
It therefore holds that  $\text{diag}\{\mathbf{E}_{\text{CL}} \mathbf{H}\} = \mathbf{1}$ . Furthermore, in [5] we showed that for mutually independent parameters  $\mathbf{e}_{\text{CL},i}$  does not depend on  $\sigma_{x_i}^2$  and is also given by  $\mathbf{e}_{\text{CL},i} = (\mathbf{h}_i^H \mathbf{C}_i^{-1} \mathbf{h}_i)^{-1} \mathbf{C}_i^{-1} \mathbf{h}_i$ , where  $\mathbf{C}_i = \bar{\mathbf{H}}_i \mathbf{C}_{\bar{\mathbf{x}}_i \bar{\mathbf{x}}_i} \bar{\mathbf{H}}_i^H + \mathbf{C}_{\mathbf{nn}}$ .

## 2.3 Other cases

If  $\mathbf{x}$  is whether complex proper Gaussian nor a vector with mutually independent parameters, then we have the following possibilities: If  $E_{\mathbf{y}|x_i}[\hat{x}_i|x_i]$  is a linear function of  $x_i$  for all  $i = 1, 2, \dots, n$  then we can derive the CWCU LMMSE estimator similar as in Section 2.1. In the remaining cases still an estimator can be found that fulfills the CWCU constraints. As discussed above the choice  $\mathbf{e}_i^H \mathbf{h}_i = 1$ ,  $\mathbf{e}_i^H \bar{\mathbf{H}}_i = \mathbf{0}^T$  together with  $c_i = 0$  for all  $i = 1, 2, \dots, n$  ensures that (3) holds. Inserting these constraints into the Bayesian MSE cost functions and solving the constrained optimization problems leads to

$$\hat{\mathbf{x}}_{\text{B}} = \mathbf{E}_{\text{B}} \mathbf{y} = (\mathbf{H}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{y}, \quad (24)$$

with  $\mathbf{C}_{\text{ee,B}} = (\mathbf{H}^H \mathbf{C}_{\mathbf{nn}}^{-1} \mathbf{H})^{-1}$  as the Bayesian error covariance matrix.  $\mathbf{e}_i^H \mathbf{h}_i = 1$  &  $\mathbf{e}_i^H \bar{\mathbf{H}}_i = \mathbf{0}$  for all  $i = 1, 2, \dots, n$  is equivalent to  $\mathbf{E} \mathbf{H} = \mathbf{I}$ . This implies  $\hat{\mathbf{x}}_{\text{B}} = \mathbf{E}_{\text{B}} \mathbf{y} = \mathbf{x} + \mathbf{E}_{\text{B}} \mathbf{n}$ . It follows that the estimator in (24) also fulfills the global unbiasedness condition  $E_{\mathbf{y}|\mathbf{x}}[\hat{\mathbf{x}}_{\text{B}}|\mathbf{x}] = \mathbf{x}$  for every  $\mathbf{x} \in \mathbb{C}^{n \times 1}$ . This estimator which is the BLUE is not able to exploit any prior knowledge about  $\mathbf{x}$ . Usually the BLUE is treated in the classical instead of the Bayesian framework.



**Fig. 1.** Visualization of the relative frequencies of the estimates  $\hat{x}_{B,1}$ ,  $\hat{x}_{CL,1}$ , and  $\hat{x}_{L,1}$ , respectively. The black crosses mark the ideal 4-QAM constellation points.

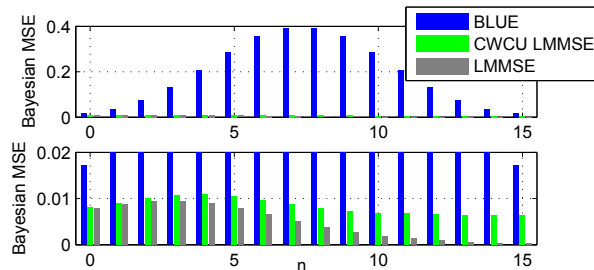
### 3 Applications

#### 3.1 QPSK data estimation

An example that exhibits the properties of the CWCU LMMSE concept most demonstrative is the estimation of channel distorted and noisy received quadrature amplitude modulated (QAM) data symbols. We assume an underlying linear model as in (4), with a parameter vector  $\mathbf{x}$  consisting of 4 mutually independent 4-QAM symbols, each out of  $\{\pm 1 \pm j\}$ , complex proper additive white Gaussian noise (AWGN) with variance  $\sigma_n^2$ , and a 4x4 channel matrix  $\mathbf{H}$ . Due to the mutual independence of the 4-QAM data symbols we use the CWCU LMMSE estimator from Proposition 2. The experiment is repeated a large number of times for a fixed  $\sigma_n^2$  and for a particularly chosen channel matrix  $\mathbf{H}$ . Fig. 1 visualizes the relative frequencies of the estimates  $\hat{x}_{B,1} = [\hat{\mathbf{x}}_B]_1$  (BLUE),  $\hat{x}_{CL,1} = [\hat{\mathbf{x}}_{CL}]_1$  (CWCU LMMSE),  $\hat{x}_{L,1} = [\hat{\mathbf{x}}_L]_1$  (LMMSE) in the complex plane. The estimates of both  $\hat{x}_{B,1}$  and  $\hat{x}_{CL,1}$  are centered around the true constellation points since these estimators fulfill the CWCU constraints. The Bayesian MSE of  $\hat{x}_{CL,1}$  is clearly below the one of  $\hat{x}_{B,1}$  since the former is able to incorporate the prior knowledge inherent in  $\mathbf{C}_{\mathbf{x}\mathbf{x}} = \sigma_{\mathbf{x}}^2 \mathbf{I}$ .  $\hat{x}_{L,1}$  is conditionally biased towards the prior mean which is 0. While the LMMSE and CWCU LMMSE estimator lead to the same bit error ratio (BER) when the decision boundaries are adapted properly. However, the BLUE shows a worse performance in the MSE and in the BER.

#### 3.2 Channel estimation

As a second example to demonstrate the properties of the CWCU LMMSE estimator we choose the well-known channel estimation problem for IEEE 802.11a/n WLAN standards [7], which extends our investigations in [5]. The standards define two identical length  $N = 64$  preamble symbols designed such that their discrete Fourier transformed (DFT) versions show  $\pm 1$  at 52 subcarrier positions (indexes  $\{1, \dots, 26, 38, \dots, 63\}$ ) and zeros at the remaining ones (indexes  $\{0, 27, \dots, 37\}$ ). The channel impulse response (CIR) is modeled as a zero mean



**Fig. 2.** Top: Bayesian MSEs of the estimated CIR coefficients; Bottom: zoomed version.

complex proper Gaussian vector, i.e.  $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_{\mathbf{h}\mathbf{h}})$ , with  $\mathbf{C}_{\mathbf{h}\mathbf{h}} = \text{diag}\{\sigma_0^2, \sigma_1^2, \dots, \sigma_{l_h-1}^2\}$  and exponentially decaying power delay profile with  $\sigma_i^2 = (1 - e^{-T_s/\tau_{rms}}) e^{-iT_s/\tau_{rms}}$  for  $i = 0, 1, \dots, l_h - 1$ .  $T_s$  and  $\tau_{rms}$  are the sampling time and the channel delay spread, respectively, and  $l_h$  is the channel length which can be assumed to be considerably smaller than  $N$ . In our setup we chose  $T_s = 50\text{ns}$ ,  $\tau_{rms} = 100\text{ns}$ , and  $l_h = 16$ . The transmission of the training symbols over the channel can again be written as a linear model with complex proper AWGN noise, and with  $\mathbf{h}$  as the vector parameter whose realization is to be estimated, cf. [5]. Fig. 2 shows the Bayesian MSEs of the BLUE ( $\hat{\mathbf{h}}_B$ ), the LMMSE estimator ( $\hat{\mathbf{h}}_L$ ), and the CWCU LMMSE estimator ( $\hat{\mathbf{h}}_{CL}$ ) for a time domain noise variance of  $\sigma_n^2 = 0.01$ . Proposition 2 has been used to derive  $\hat{\mathbf{h}}_{CL}$  since the elements of  $\mathbf{h}$  are mutually independent.  $\hat{\mathbf{h}}_{CL}$  almost reaches the performance of  $\hat{\mathbf{h}}_L$ , and in contrast to the latter it additionally shows the property of conditional unbiasedness. Both estimators incorporate the prior knowledge inherent in  $\mathbf{C}_{\mathbf{h}\mathbf{h}}$  which results in a huge performance gain over  $\hat{\mathbf{h}}_B$ . We now turn to frequency response estimators and note that the vector of frequency response coefficients  $\tilde{\mathbf{h}} \in \mathbb{C}^{64 \times 1}$  (which corresponds to the DFT of the zero-padded impulse response  $[\mathbf{h}^T \mathbf{0}^T]^T$ ) consists of proper Gaussian elements, but the PDF of  $\tilde{\mathbf{h}}$  cannot be written in the form of a multivariate proper Gaussian PDF. The LMMSE estimator  $\hat{\tilde{\mathbf{h}}}_L$  is simply obtained by the DFT of  $[\hat{\mathbf{h}}_L^T \mathbf{0}^T]^T$  (since it commutes over linear transformations). As discussed in [5], the BLUE  $\hat{\tilde{\mathbf{h}}}_B$  can be derived correspondingly. The CWCU LMMSE estimator  $\hat{\tilde{\mathbf{h}}}_{CL}$  cannot be derived in this way since it does not commute over general linear transformations. However, although  $\tilde{\mathbf{h}}$  is not a proper Gaussian vector,  $E_{\tilde{\mathbf{h}}_i|\tilde{h}_i}[\tilde{\mathbf{h}}_i|\tilde{h}_i]$  is linear in  $\tilde{h}_i$  (for all  $i = 0, 1, \dots, N - 1$ ), and one can easily show that (16)-(18) can be applied to determine the CWCU LMMSE estimator. The frequency domain version of the prior covariance matrix  $\mathbf{C}_{\mathbf{h}\mathbf{h}}$  is required for its derivation. Fig. 3 shows the Bayesian MSEs of  $\hat{\tilde{\mathbf{h}}}_B$ ,  $\hat{\tilde{\mathbf{h}}}_L$ , and  $\hat{\tilde{\mathbf{h}}}_{CL}$ , respectively.  $\hat{\tilde{\mathbf{h}}}_B$  is outperformed by  $\hat{\tilde{\mathbf{h}}}_L$  and  $\hat{\tilde{\mathbf{h}}}_{CL}$  at all frequencies, but the performance loss is significant at the large gap from subcarrier 27 to 37, where no training information is available. In contrast,  $\hat{\tilde{\mathbf{h}}}_L$  and  $\hat{\tilde{\mathbf{h}}}_{CL}$  show excellent interpolation properties along this gap. Large estimation errors of  $\hat{\tilde{\mathbf{h}}}_B$  in this spectral region are spread over all time domain samples

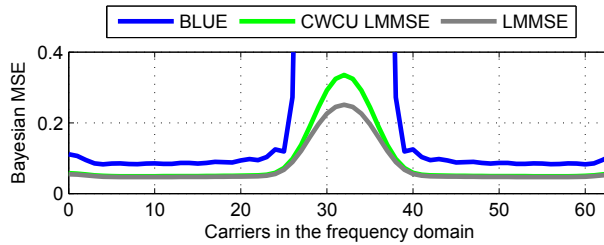


Fig. 3. Bayesian MSEs for the elements of  $\hat{\mathbf{h}}_{\mathbf{B}}$ ,  $\hat{\mathbf{h}}_{\mathbf{L}}$ , and  $\hat{\mathbf{h}}_{\mathbf{CL}}$ , respectively.

which explains the poor performance of  $\hat{\mathbf{h}}_{\mathbf{B}}$ . Note that in practice this is only critical if  $\hat{\mathbf{h}}_{\mathbf{B}}$  is incorporated in the receiver processing, however, pure frequency domain receivers only require estimates at the occupied 52 subcarrier positions.

## 4 Conclusion

In this work we investigated the CWCU LMMSE estimator for the linear model. First, we derived the estimator for complex proper Gaussian parameter vectors, and for the case of mutually independent (and otherwise arbitrarily distributed) parameters. For the remaining cases the CWCU LMMSE estimator may correspond to a globally unbiased estimator. The implications of the CWCU constraints have been demonstrated in a data estimation example using a discrete alphabet, and in a channel estimation application. In both applications the CWCU LMMSE estimator considerably outperforms the globally unbiased BLUE.

## References

1. S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall PTR, 1st edition, Apr. 2010.
2. M. Triki, D.T.M. Slock, "Component-Wise Conditionally Unbiased Bayesian Parameter Estimation: General Concept and Applications to Kalman Filtering and LMMSE Channel Estimation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, Nov. 2005, pp. 670-674.
3. M. Triki, A. Salah, D.T.M. Slock, "Interference cancellation with Bayesian channel models and application to TDOA/IPDL mobile positioning," in *Proc. International Symposium on Signal Processing and its Applications*, Aug. 2005, pp. 299-302.
4. M. Triki, D.T.M. Slock, "Investigation of Some Bias and MSE Issues in Block-Component-Wise Conditionally Unbiased LMMSE," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, Nov. 2006, pp.1420-1424.
5. M. Huemer, O. Lang, "On Component-Wise Conditionally Unbiased Linear Bayesian Estimation," in *Proc. Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Nov. 2014, pp. 879-885.
6. T. Adali, P.J. Schreier, L.L. Scharf, "Complex-Valued Signal Processing: The Proper Way to Deal With Improperity," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101-5125, Nov. 2011.
7. IEEE Std 802.11a-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-Speed Physical Layer in the 5 GHz Band, 1999.