

# INVESTIGATION OF SOME BIAS AND MSE ISSUES IN BLOCK-COMPONENT-WISE CONDITIONALLY UNBIASED LMMSE

*Mahdi Triki, and Dirk T.M. Slock\**

Eurecom Institute  
2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE  
Email: {triki,slock}@eurecom.fr

## ABSTRACT

Imposing a joint conditionally unbiasedness constraint on a vector of parameters reduces Bayesian estimation to deterministic parameter estimation. In Component-Wise Conditionally Unbiased (CWCU) Bayesian parameter estimation, every parameter in turn is treated as deterministic while the others are being treated as Bayesian. If the parameters are transmitted symbols, the CWCU approach corresponds to unbiased symbol detection whereas joint deterministic unbiasedness leads to a zero-forcing approach.

In this paper we consider an intermediate approach corresponding to Block-CWCU (BCWCU) LMMSE estimation. We investigate the interplay between block-size, joint bias and prior covariance rank. We also investigate two Sequential Interference Cancellation (SIC) implementations of BCWCU-LMMSE estimation. The results are illustrated with concrete examples.

## 1. INTRODUCTION

In most applications, estimator designs are subject to a tradeoff between bias and variance. Bias is due to 'mismatch' between the average value of the estimator and the true parameter (conditional bias); whereas variance arises from fluctuations in the estimator due to statistical sampling. The bias / variance tradeoff is typically fixed by minimizing the Mean Squared Error (MSE) under some constraints on the bias. If prior information on the parameter statistics is available, Bayesian estimation theory shows that under the Bayesian unbiasedness constraint, the MSE is bounded below by the Bayesian CRB. On the other hand, the MMSE estimator minimizes  $R_{\hat{\theta}\hat{\theta}}$ , the parameter estimation error correlation matrix, and not only the MSE, which is the trace of  $R_{\hat{\theta}\hat{\theta}}$ . Bayesian unbiasedness for random parameters corresponds to unbiasedness on the average, which is a very weak requirement. Indeed, in particular the MMSE estimator is unbiased. Hence, the MMSE estimator minimizes  $R_{\hat{\theta},\hat{\theta}}$  and the MSE, regardless of whether the Bayesian unbiasedness constraint is imposed or not.

In recent years, the Bayesian formulation of channel estimation has become popular, as it allows for instance the exploitation of the power delay profile. This allows to reduce the effective number of parameters to be estimated from an a priori delay spread range to the effective delay spread of the power delay profile. For SIMO, MISO or MIMO channels, the Bayesian formulation allows to exploit (spatial) correlation between antennas and reduces the number of parameters from the physical number

of antennas to a reduced effective number of uncorrelated antennas. When the channel is fading in time, the Doppler spectrum and hence correlation in time can be exploited via Wiener or Kalman filtering to further reduce the MSE. Bayesian estimation leads to biased channel estimates. This bias is detrimental for a number of applications: Maximum Likelihood Symbol Detection (MLSD) for e.g. a channel with delay spread using the Viterbi algorithm (the bias is as detrimental as in biased LMMSE symbol receivers), fitting a parametric (pathwise) model to the channel impulse response, or using the channel estimate for the design of the receiver or the transmitter. The type of unbiasedness that is required here is conditional unbiasedness (again, unbiasedness for Bayesian estimation corresponds to unbiasedness on the average). On the other hand, conditional unbiasedness for vectors of parameters is usually introduced globally, requiring all parameter components to be jointly unbiased. Such a stringent requirement, which corresponds to zero-forcing when the parameters are multiple symbols, prevents the exploitation of correlations between the parameters, and hence leads to a significant reduction in the benefits brought about by the Bayesian framework, the prior knowledge.

This motivated us to introduce the Component-Wise Conditionally Unbiased (CWCU) Bayesian parameter estimation [2]. Instead of constraining the estimator to be globally unbiased, i.e.,  $E_{Y|\theta}(\hat{\theta} - \theta) = 0$ , we impose conditional unbiasedness on one parameter component at a time, i.e.,

$$E_{Y|\theta_k}(\hat{\theta}_k - \theta_k) = 0, \quad k = 1 : K \quad (1)$$

where  $E_{Y|x}[Z(Y, X)] = \int Z(Y, x)f_{Y|x}(y|x)dY$  denotes the expectation of  $Z(X, Y)$  on  $Y$  conditionally to  $X = x$ ; and  $\theta = [\theta_1 \dots \theta_K]^T$  is the parameter vector to be estimated. In this way, the parameter (component) of interest is constrained to be conditionally unbiased, while the other parameters are treated as nuisance parameters. Note that the component-wise concept can be defined at different levels. For example, consider multi-channel impulse response estimation. The component-wise concept can be defined at scalar level by considering conditional unbiasedness separately for different channels and time lags. It can also be defined at a block level by considering conditional unbiasedness jointly for different channels, and separately for different time lags, or even jointly for the different time lags and separately for different SISO channels.

This paper is organized as follows. In section 2, we introduce Block-CWCU-LMMSE estimation for a linear Gaussian Model. The interplay between block-size, joint bias and prior covariance rank is investigated in section 3. Application of the concept to channel estimation for mobile localization is presented in section 4.

---

\*Institut Eurécom's research is partially supported by its industrial members: BMW, Bouygues Télécom, Cisco Systems, France Télécom, Hitachi Europe, SFR, Sharp, STMicroelectronics, Swisscom, Thales.

## 2. BCWCU-LMMSE ESTIMATION FOR LINEAR GAUSSIAN MODEL

We consider a linear Gaussian model:

$$y = H\theta + v \quad (2)$$

where  $y$  is  $N \times 1$  vector containing the measured signal,  $\theta \sim \mathcal{N}(0, C_{\theta\theta})$  is a  $L \times 1$  vector containing the parameters to be estimated, and  $v \sim \mathcal{N}(0, \sigma_v^2 I_N)$  is an  $N \times 1$  additive white Gaussian noise independent from  $\theta$ .  $I_N$  represents the identity matrix of size  $N$ . As  $\theta$  and  $y$  are jointly Gaussian, minimizing the MSE leads to the LMMSE estimator:

$$\begin{aligned} \hat{\theta}_{lmmse} &= \arg \min_{\hat{\theta}=Fy} E \left\| \hat{\theta} - \theta \right\|^2 \\ &= \arg \min_{\hat{\theta}=Fy} \text{tr} \left\{ (FH - I) C_{\theta\theta} (FH - I)^H + \sigma_v^2 FF^H \right\} \\ &= \left( \sigma_v^2 I_L + C_{\theta\theta} H^H H \right)^{-1} C_{\theta\theta} H^H y \end{aligned} \quad (3)$$

where  $\text{tr}$  denotes the trace operator.

Under the joint unbiasedness constraint, minimizing the MSE leads to

$$\hat{\theta} = \begin{cases} \arg \min_{\hat{\theta}=Fy} E \left\| \hat{\theta} - \theta \right\|^2 \\ E_{Y|Y} (\hat{\theta} - \theta) = 0 \end{cases} = \begin{cases} \arg \min_{\hat{\theta}=Fy} \text{tr} \left\{ FF^H \right\} \\ FH = I_L \end{cases}$$

Then, joint unbiasedness prevents the exploitation of correlations between the parameters, and leads to a significant reduction in the benefits brought about by the Bayesian framework: the prior knowledge. Indeed, under global unbiasedness, the MMSE estimator corresponds to the BLUE, i.e.,

$$\hat{\theta}_{blue} = \left( H^H H \right)^{-1} H^H y \quad (4)$$

LMMSE and BLUE estimators are related by

$$\hat{\theta}_{lmmse} = \underbrace{C_{\theta\theta} \left( (H^H H) C_{\theta\theta} + \sigma_v^2 I_L \right)^{-1} (H^H H)}_{B_{lmmse}} \hat{\theta}_{blue} \quad (5)$$

where  $B_{lmmse}$  represents the bias of LMMSE estimation.

Assume that  $\theta = [\theta_1^H \dots \theta_K^H]^H$  represents a decomposition in  $K$  sub-sets.  $\theta_k$   $k = 1 : K$  can be either a scalar or vector components. We denote by  $L_k$  the size of  $\theta_k$  ( $\sum_k L_k = L$ ). Imposing the BCWCU constraints leads to

$$\begin{aligned} \hat{\theta}_{bcwculmmse} &= \begin{cases} \arg \min_{\hat{\theta}=Fy} E \left\| \hat{\theta} - \theta \right\|^2 \\ E_{Y|\theta_k} (\hat{\theta}_k - \theta_k) = 0 \quad k = 1 : K \end{cases} \\ \iff &\begin{cases} \min_F \text{tr} \left\{ (FH - I_L) C_{\theta\theta} (FH - I_L)^H + \sigma_v^2 FF^H \right\} \\ e_k^H FH C_{\theta\theta} e_k = e_k^H C_{\theta\theta} e_k \quad k = 1 : K \end{cases} \end{aligned}$$

where  $e_k = [0 \ 0 \ I_{L_k} \ 0 \ 0]^H$  is the  $L \times L_k$  matrix such that  $e_k^H \theta = \theta_k$ .

If the  $\{\theta_k\}_k$  are decorrelated ( $C_{\theta\theta}$  is block-diagonal), the CWCU constraint becomes  $e_k^H FH e_k = I_k$ . In this case, the component

of interest  $\theta_k$  is treated as deterministic, whereas the other (correlated) parameter components  $\{\theta_p\}_{p \neq k}$  continue to be treated as Bayesian. Using Lagrange optimization, one can show that the BCWCU-LMMSE is given by:

$$\hat{\theta}_{bcwculmmse} = D_{bcw} \hat{\theta}_{lmmse} = D_{bcw} B_{lmmse} \hat{\theta}_{blue} \quad (6)$$

where  $D_{bcw} = \text{bdiag}(C_{\theta\theta}) (\text{bdiag}(B_{lmmse} C_{\theta\theta}))^\#$  is a block-diagonal matrix that ensures the component-wise unbiasedness constraint; and  $(\cdot)^\#$  denotes the Moore-Penrose pseudoinverse operator.  $\text{bdiag}(B) = \sum_{k=1}^K e_k \left( e_k^H B e_k \right) e_k^H$  is a  $L \times L$  block-diagonal matrix formed by the block-diagonal elements of the matrix  $B$ .

*Special cases:*

- If the parameters  $\theta_k$  are decorrelated ( $C_{\theta\theta} = \text{bdiag}(C_{\theta\theta})$ ),  $D_{bcw}$  can be simplified to:

$$D_{bcw} = (\text{bdiag}(B_{lmmse}))^{-1} \quad (7)$$

If furthermore all  $L_k = 1$  (and  $C_{\theta\theta} = \text{bdiag}(C_{\theta\theta})$ ), hence the CWCU-LMMSE corresponds to the Unbiased-LMMSE.

- If there is no-coupling between the parameters  $\theta_k$  neither through prior nor through data ( $B_{lmmse}$  is block-diagonal), the BCWCU-LMMSE corresponds to the BLUE estimation. The BCWCU-LMMSE estimation is of interest if there is a coupling through prior ( $C_{\theta\theta}$  is not block diagonal), and/or data ( $(H^H H)$  is not block diagonal).

Reciprocally, one can show that

- If  $D_{bcw}$  is block-diagonal, then the  $\{\theta_k\}_k$  are decorrelated.
- If  $D_{bcw} = B_{lmmse}^{-1}$ , then the  $\{\theta_k\}_k$  are decoupled.

Remark that in linear multi-user detection ( $C_{\theta\theta}$  is diagonal), CWCU-LMMSE estimation corresponds to Unbiased LMMSE; whereas, the (jointly) conditionally unbiased estimator (BLUE) corresponds to MMSE-ZF reception.

## 3. INTERPLAY BETWEEN GLOBAL BIAS AND PRIOR COVARIANCE RANK FOR BCWCU-LMMSE

We shall illustrate the concepts through an example. Consider a Base Station (BS) using  $M$ -element antenna array. The received signal over single-path propagation is an  $M \times 1$  vector given by:

$$y(k) = x(k) H A(\theta) + v(k) \quad (8)$$

where  $x(k)$  is a known scalar training sequence transmitted by the user,  $H$  in a  $M \times M$  known matrix describing the coupling between the antenna elements,  $v(k)$  is an additive white Gaussian noise, i.e.,  $\sim \mathcal{N}(0, \sigma_v^2 I_M)$ , and  $A(\theta)$  denotes the array response (function of the array geometry, and the direction of arrival  $\theta$ ).

The Direction of Arrival (DoA)  $\theta$  is generally estimated using a two step approach:

1. Estimate the array response vector  $A(\theta)$ .
2. Compute the DoA based on the array manifold  $A(\cdot)$ .

In the literature, the Least-Squares (LS) technique (which corresponds to BLUE in this problem) is proposed for the estimating of the array response vector [3, 4, 5]

$$\hat{A}_{blue} = \sigma_x^{-2} \left( H^H H \right)^{-1} H^H \sum_k x^*(k) y(k) \quad (9)$$

where  $\sigma_x^2 = \sum_k |x(k)|^2$  represents the energy of the training sequence  $\{x(k)\}_k$ .

As we have seen in the previous section, BLUE provides an unbiased, but noisy estimate, i.e.

$$E_{Y|A} \left\{ \hat{A}_{blue} \right\} = A(\theta) \quad (10)$$

On the other hand, if prior information is available, it can be used to reduce the estimation SNR. In the following, we will investigate the effect of the use of a Bayesian prior on the estimation bias. We assume that the direction of arrival is varying around a unknown nominal DoA  $\theta_0$ , i.e.,

$$\theta = \theta_0 + \delta\theta \quad (11)$$

And, we will have to estimate multiple instances of  $\theta$ . Using a first order approximation, we have

$$A(\theta) \approx A(\theta_0) + \delta\theta A'(\theta_0) \quad (12)$$

where  $A'(\theta_0) = \left. \frac{\partial A(\theta)}{\partial \theta} \right|_{\theta=\theta_0}$  denotes the gradient of  $A(\theta)$  at  $\theta = \theta_0$ .  $A(\theta)$  is random due to  $\delta\theta$ . Assuming  $\delta\theta$  to have zero mean and variance  $\sigma_\delta^2$ , the covariance matrix of  $A(\theta)$  becomes:

$$\begin{aligned} C_A &= E \left\{ A(\theta) A^H(\theta) \right\} \\ &= A(\theta_0) A^H(\theta_0) + \sigma_\delta^2 A'(\theta_0) A'^H(\theta_0). \end{aligned}$$

As  $\text{rank}(C_A) = 2$ , using the eigen decomposition, the prior covariance matrix can be written as:

$$C_A = U \begin{bmatrix} \Lambda_C & 0 \\ 0 & 0 \end{bmatrix} U^H \quad (13)$$

where  $\Lambda_C$  is a  $2 \times 2$  diagonal matrix, and  $U$  is a unitary matrix.

By introducing  $U_A = U \begin{bmatrix} I_2 & 0 & \cdots & 0 \end{bmatrix}^H$ ,  $C_A$  can be simplified to

$$C_A = U_A \Lambda_C U_A^H \quad (14)$$

Note that  $A(\theta_0)$ ,  $A'(\theta_0)$  are unknown. Only the covariance  $C_A$  (and  $U_A$ ) is known. Remark also that  $A(\theta)$  lives in the subspace spanned by  $U_A$ . Then, we can introduce zero-mean random variables  $\eta$ , and  $\gamma$  such that

$$A(\theta) = U_A \begin{bmatrix} \eta \\ \gamma \end{bmatrix}. \quad (15)$$

From (5), one can show that

$$B_{lmmse} U_A = U_A \underbrace{(\Lambda_C \Lambda_H + \sigma_v^2 I_2)^{-1} \Lambda_C \Lambda_H}_{\Lambda_B} \quad (16)$$

where  $\Lambda_H = U_A (H^H H) U_A^H$  is a  $2 \times 2$  matrix, not necessarily diagonal. Then, the expected value of the LMMSE estimate is

$$\begin{aligned} E_{Y|A} \left\{ \hat{A}_{lmmse} \right\} &= B_{lmmse} A(\theta) \\ &= U_A \Lambda_B \begin{bmatrix} \eta \\ \gamma \end{bmatrix} = U_A \begin{bmatrix} \eta' \\ \gamma' \end{bmatrix}. \end{aligned} \quad (17)$$

In summary, the LMMSE estimates the array response in the right subspace, but with biased weighing. If  $\Lambda_B$  is not a multiple of

identity, the LMMSE estimate of  $A(\theta)$  leads to erroneous DoA estimation.

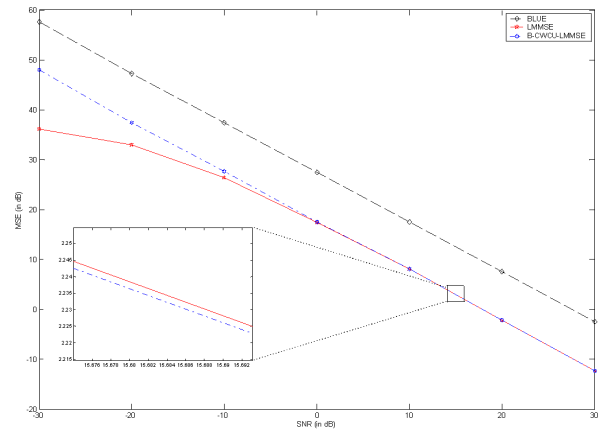
If we impose Block-CWCU constraints, under some regularity assumptions, one can show that using a block-size 2 ( $L_k \geq 2 \forall k$ ):

$$D_{bcw} B_{lmmse} U_A = D_{bcw} U_A \Lambda_B = U_A$$

Thus, the BCWCU-LMMSE (with a block-size 2), guarantees **joint** unbiasedness, i.e., the expected value of the BCWCU-LMMSE estimate is

$$\begin{aligned} E_{Y|A} \left\{ \hat{A}_{bcwulmmse} \right\} &= D_{bcw} B_{lmmse} A(\theta) \\ &= U_A \begin{bmatrix} \eta \\ \gamma \end{bmatrix} = A(\theta) \end{aligned} \quad (18)$$

In figure 1, we plot the estimation MSE =  $\text{tr} \{ C_{\tilde{A}\tilde{A}} \}$  of the BLUE, LMMSE, and BCWCU-LMMSE estimates. The MSE is averaged over 500 Montecarlo runs. The matrices  $C_{\theta\theta}$  (having rank 2), and  $H$  are generated randomly.  $M$  was chosen equal to 10.



**Fig. 1.** Estimation MSE of the BLUE, LMMSE, and BCWCU-LMMSE estimators as a function of SNR

Thus, for the limited rank prior covariance matrix case, BCWCU-LMMSE reduces the estimation noise, while guaranteeing joint unbiasedness.

The result can be easily generalized to an arbitrary prior covariance rank. This leads to the following theorem.

**Theorem:** Let  $m$  denotes the rank of the prior covariance matrix  $C_{\theta\theta}$ . Then BCWCU-LMMSE, with block sizes of at least  $m$ , guarantees the joint unbiasedness.

#### 4. BCWCU-LMMSE FOR MULTIPLE CHANNEL ESTIMATION

In this section, we will focus on one particular problem setting, in which the channels from different Base Stations (BSs) to a Mobile Station (MS) need to be estimated jointly. The estimation of the transmission channel plays a crucial role in communication systems (for mobile positioning applications, multi-user detection...). Channel parameters are observed indirectly by the received data : convolved with a known training sequence and embedded in a (white Gaussian) noise.

$$y = \sum_{k=1}^K X_k h_k + v \quad (19)$$

where

- $y = [y_1^H \cdots y_N^H]^H$  denotes received data.  $N$  is the data length.
- $v = [v_1^H \cdots v_N^H]^H$  represents the additive white gaussian noise.
- $K$  is the number of base stations.
- $h_k = [h_{k,1}^H \cdots h_{k,L_k}^H]^H$  denotes the Channel Impulse Response (CIR) between the MS and the  $k^{th}$  BS.  $L_k$  is  $k^{th}$  CIR length.
- $X_k = \begin{bmatrix} x_1 & \cdots & x_{L_k} \\ \vdots & & \vdots \\ x_N & \cdots & x_{N+L_k-1} \end{bmatrix}$  is an  $N \times L_k$  Hankel matrix characterizing the training sequence of the  $k^{th}$  BS.

Using a compact notation, the received data can be written as:

$$y = Xh + v \quad (20)$$

where  $X = [X_1 \cdots X_K]$ , and  $h = [h_1^H \cdots h_K^H]^H$ .

Whereas the direct use of a Bayesian channel estimate for an interfering signal allows to better suppress the interference, its use for the user of interest may lead to a bias problem. This bias is detrimental for a number of applications. For example, the estimation bias is undesirable for mobile localization applications (e.g. Time of Arrival (ToA) is estimated by fitting a parametric model to the channel impulse response) [1]. That is why the Bayesian prior is rarely taking into account for such applications. Channel estimation is done typically based on Least-Squares (LS) or Matching Pursuit (MP) approaches [3, 5, 7, 6]. On the other hand, imposing joint unbiasedness between the CIRs coming from different base-stations is not required: we can allow for interference (contribution of other base-stations), if this can be motivated by a noise reduction.

Note that the problem has a special structure. In fact, the channel impulse responses and their individual coefficients are decorrelated ( $C_{hh}$  is diagonal). On the other hand, the data covariance

matrix  $X^H X = \begin{bmatrix} X_1^H X_1 & \cdots & X_1^H X_K \\ \vdots & & \vdots \\ X_K^H X_1 & \cdots & X_K^H X_K \end{bmatrix}$  can not be assumed

to be block-diagonal due to:

- The limited length of the training sequence (for example in 3G, the channel estimation is done only in the Idle Period Down-Link (IPDL) [1]). Thus, despite the input being white, the training sequence is not long enough to lead to a spherical estimate of the input covariance matrix  $X^H X$ .
- The range of the CIR powers. In fact, despite the quantities  $X_k^H X_k$  being approximately white ( $\approx \sigma_k^2 I_{L_k}$ ), e.g.  $X_1^H X_K$  can not be neglected with respect to  $X_K^H X_K$ .

Thus, the channel impulses are coupled through the data covariance matrix  $X^H X$  and the BCWCU-LMMSE is of interest.

The LMMSE estimate is given by:

$$\begin{aligned} \hat{h}_{lmmse} &= \underbrace{\left( \sigma_v^2 \left( X^H X \right)^{-1} C_{hh}^{-1} + I \right)^{-1} \left( X^H X \right)^{-1} X^H y}_{B_{lmmse}} \\ &= B_{lmmse} \hat{h}_{blue} \end{aligned}$$

The bloc component-wise unbiasedness constraint is formulated as

$$E \left[ \hat{h}_k / h_k \right] = h_k \quad k = 1 : K \quad (21)$$

As the prior covariance matrix is diagonal, minimizing the MSE (under BCWCU constraints) leads to

$$\hat{h}_{bcwculmmse} = (\text{bdiag}(B_{lmmse}))^{-1} B_{lmmse} \hat{h}_{blue} \quad (22)$$

#### 4.1. SIC implementation of the BCWCU-LMMSE estimator

The inherent complexity of the BCWCU-LMMSE scheme is cubic in  $L = \sum_k L_k$  (the same as for the LMMSE and the BLUE estimators). For practical implementation, the Successive Interference Cancellation (SIC) approach can be used to approximate the BCWCU-LMMSE estimator, with a complexity linear in  $L$ .

Successive interference cancellation multi-channel estimation is a scheme in which CIR's are estimated successively. The approach successively cancels the interference from the next strongest channel. Assume that channels have been ordered in order of decreasing  $SNR_i = \frac{\sigma_x^2 \|h_i\|^2}{\sigma_v^2} = \frac{\sigma_i^2}{\sigma_v^2}$  at the channel estimator input.

First, we compute an unbiased estimate of the first (strongest) CIR (the BLUE is proportional to the matched filter). The contribution of weaker CIRs is ignored, i.e.,

$$\hat{h}_1 = \frac{1}{N\sigma_x^2} X_1^H y \quad (23)$$

Then, the LMMSE estimator is derived, the interfering signal is recreated at the receiver, and subtracted from the received waveform.

$$\begin{aligned} \hat{h}_1^{SIC} &= \left( \sigma_v^2 C_{hh,1}^{-1} + X_1^H X_1 \right)^{-1} (N\sigma_x^2) \hat{h}_1 \\ \hat{y}_1 &= X_1 \hat{h}_1^{SIC} \\ y &\leftarrow y - \hat{y}_1 \end{aligned}$$

Remark that even if  $X^H X$  is not approximately diagonal, the non-diagonal elements of  $X_1^H X_1$  can be neglected (as the number of unknowns is  $M$  times less). One recursion of the SIC implementation of the BCWCU-LMMSE is described in the table below:

SIC implementation of the BCWCU-LMMSE		
#	Computation	cost
channel estimation		
1	$\hat{h}_k = \frac{1}{N\sigma_k^2} X_k^H y$	$O(NL_k)$
Interference cancellation		
2	$\hat{h}_k^{lmmse} = \left( \frac{\sigma_v^2}{N\sigma_k^2} C_{hh,k}^{-1} + I_{L_k} \right)^{-1} \hat{h}_k$	$O(L_k)$
3	$\hat{y}_k = X_k \hat{h}_k^{lmmse}$	$O(NL_k)$
4	$y = y - \hat{y}_k$	

In this manner successive BS CIRs does not have to encounter interference caused by initial BS CIRs. SIC leads to good performance for all channel estimates: initial CIR estimates improve because the later channels have less power which means less interference for the initial channels, and later CIR estimates improve because early BS's interference has been cancelled out.

#### 4.2. Modified SIC implementation of the BCWCU-LMMSE estimator

In the linear SIC approach above there are two sources of error:

- Ignoring the contribution of channels with lower powers.
- Non-perfect cancellation of estimated channels.

In the following, we will try to alleviate the propagation of the estimation error (due to non-perfect interference cancellation). We suggest taking, at each step  $k$ , the estimate  $\hat{h}_k$  computed from the joint LMMSE estimation of  $h^{(k)} = [h_1^H \cdots h_k^H]^H$ . As in the classic SIC approach, the contribution of channels with lower powers  $\bar{h}^{(k)} = [h_{k+1}^H \cdots h_K^H]^H$  is ignored.

The LMMSE solution is given by

$$\hat{h}^{LMMSE,(k)} = \underbrace{\left( C_{hh}^{(k)-1} + \frac{1}{\sigma_v^2} X^{(k)H} X^{(k)} \right)^{-1}}_{B_k^{-1}} \underbrace{\frac{1}{\sigma_v^2} X^{(k)H} y}_{y_{MF}^{(k)}} y$$

where  $C_{hh}^{(k)} = E \{ h^{(k)} h^{(k)H} \}$ , and  $X^{(k)} = [X_1 \cdots X_k]$ .

By denoting  $b_k = \left( C_{hh}^{(k)-1} + \frac{1}{\sigma_v^2} X_k^H X_k \right)$ , and

$R_{k-1,k} = \frac{1}{\sigma_v^2} X^{(k-1)H} X_k$ ,  $B_{k+1}$  can be decomposed as

$$B_{k+1} = \begin{bmatrix} B_k & R_{k,k+1} \\ R_{k,k+1}^H & b_{k+1} \end{bmatrix} \quad (24)$$

$$= \begin{bmatrix} I & 0 \\ R_{k,k+1}^H B_k^{-1} & I \end{bmatrix} \begin{bmatrix} B_k & 0 \\ 0 & b_{k+1} - R_{k,k+1}^H B_k^{-1} R_{k,k+1} \end{bmatrix} \begin{bmatrix} I & B_k^{-1} R_{k,k+1} \\ 0 & I \end{bmatrix}$$

Then, the component of interest is given by:

$$\hat{h}_{k+1}^{mod-SIC} = [0 \cdots 0 I] B_{k+1}^{-1} y_{MF}^{(k+1)}$$

$$= \frac{1}{\sigma_v^2} \left( b_{k+1} - R_{k,k+1}^H B_k^{-1} R_{k,k+1} \right)^{-1} X_{k+1}^H \left( y - X^{(k)} B_k^{-1} y_{MF}^{(k)} \right)$$

We recognize the same structure as in the classic SIC. The modified SIC algorithm is described in the table below

Modified SIC implementation of the BCWCU-LMMSE		
#	Computation	cost
channel estimation		
1	$\hat{h}_k = \frac{1}{N\sigma_v^2} X_k^H y^{(k)}$	$O(NL_k)$
Interference cancellation		
2	$P_k = (b_k - R_{k-1,k}^H B_{k-1}^{-1} R_{k-1,k})$	$O(L_k (\sum_{p=1}^{k-1} L_p)^2)$
3	$\hat{h}_k^{mod-SIC} = \frac{N\sigma_v^2}{\sigma_v^2} P_k^{-1} \hat{h}_k$	$O(L_k^3)$
4	$\hat{y}_k = X_k \hat{h}_k^{mod-SIC}$	$O(NL_k)$
5	$y = y - \hat{y}_k$	
6	$B_k^{-1}$ (updated using MIL on (24))	$O(L_k (\sum_{p=1}^{k-1} L_p)^2)$

Remark that the complexity of the scheme is  $O(L^3)$  (as BCWCU-LMMSE). From this point of view, it presents no advantage. However, the performance of the proposed scheme can be interpreted as a bound on the performance of the SIC approach in section 4.1 (there is no propagation of the estimation error).

Motivated by the fact that channels are ordered by decreasing power and the observation that  $X_k^H X_k$  is approximately proportional to the identity matrix, we approximate  $B_k$  and  $b_{k+1}$  in (24) by diagonal matrices. We call the resulting scheme Modified & Simplified SIC, It has a computational complexity of  $O(L^2)$ .

We analyze the performance of the proposed algorithms by comparing their estimation MSE (computed by Monte Carlo simulations). The received signal is assumed to be the superposition of the contribution of 5 base stations, and embedded in a white Gaussian noise. The relative received signal powers are respectively 0, -5, -10, -15, -20 dB. The power delay profile is generated according to the channel model "Vehicular B". Figure 2 plots the curves of the estimation MSE of the 5<sup>th</sup> (the weakest) BS. The curves show that the SIC implementations well approximate the BCWCU-LMMSE estimator at low SNR. We remark also that the simplifications introduced to the modified scheme do not affect the estimation accuracy, and that the modified SIC outperforms the classic one (at the expense of additional complexity).

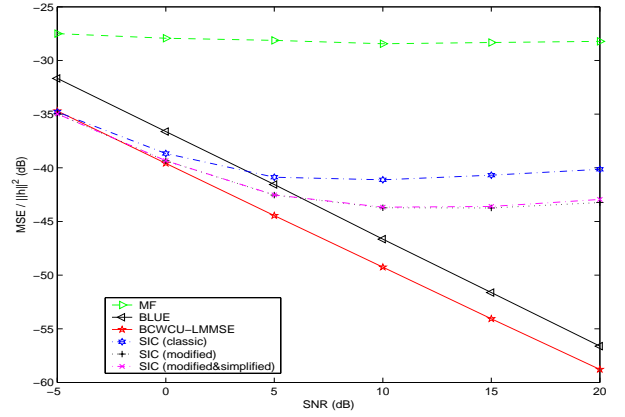


Fig. 2. CIR estimation accuracy using MF, BLUE, BCWCU-LMMSE, SIC, Modified SIC estimators

#### 5. REFERENCES

- [1] Mahdi Triki, Salah Abdellatif, and Dirk T.M. Slock. "Interference Cancellation with Bayesian Channel Models and Application to TDOA/IPDL," *In proc. of ISSPA*, August 2005.
- [2] Mahdi Triki, and Dirk T.M. Slock. "Component-Wise Conditionally Unbiased Bayesian Parameter Estimation: General Concept and Applications to Kalman Filtering and LMMSE Channel Estimation," *In proc. of Asilomar Conference*, November 2005.
- [3] S. Ozen, M.D. Zoltowski, and M. Fimoff. "A Novel Channel Estimation Method: Blending Correlation and Least-Squares Based Approaches," *In proc. of IEEE ICASSP*, Vol.3, pp. 2281-2284, May 2002.
- [4] S. Ozen, and M.D. Zoltowski. "Time-of-Arrival (TOA) Estimation Based Structured Sparse Channel Estimation Algorithm, with Applications to Digital TV Receivers," *In proc. of IEEE ICASSP*, Vol.4, pp. 481-484, April 2003.
- [5] A. Tarighat, N. Khajehnouri, and A.H. Sayed. "CDMA Location Using Multiple Antennas and Interference Cancellation," *In proc. of IEEE VTC*, Vol.4, pp. 2711-2715, April 2003.
- [6] S.F. Cotter, B.D. Rao. "Sparse Channel Estimation via Matching Pursuit with Application to Equalization," *In IEEE Trans. on Communications*, Vol.50, Issues:3, pp. 374-377, March 2002.
- [7] S. Kim, R.A. Iltis. "A Matching-Pursuit/GSIC-based Algorithm for DS-CDMA Sparse-Channel Estimation," *In IEEE Signal Processing Letters*, Vol.11, Issues:1, pp. 12-15, January 2004.