Telecommun Syst (2015) 59:417–427
DOI 10.1007/s11235-014-9902-7

Bayesian compressive sensing for ultra-wideband channel estimation: algorithm and performance analysis

Mehmet Özgör · Serhat Erküçük · Hakan Ali Çırpan

Published online: 3 December 2014

Abstract Due to the sparse structure of ultra-wideband (UWB) channels, compressive sensing (CS) is suitable for UWB channel estimation. Among various implementations of CS, the inclusion of Bayesian framework has shown potential to improve signal recovery as statistical information related to signal parameters is considered. In this paper, we study the channel estimation performance of Bayesian CS (BCS) for various UWB channel models and noise conditions. Specifically, we investigate the effects of (i) sparse structure of standardized IEEE 802.15.4a channel models, (ii) signal-to-noise ratio (SNR) regions, and (iii) number of measurements on the BCS channel estimation performance, and compare them to the results of ℓ_1-*norm* minimization based estimation, which is widely used for sparse channel estimation. We also provide a lower bound on mean-square error (MSE) for the *biased* BCS estimator and compare it with the MSE performance of implemented BCS estimator. Moreover, we study the computation efficiencies of BCS and ℓ_1-*norm* minimization in terms of computation time by making use of the big-O notation. The study shows that BCS exhibits superior performance at higher SNR regions for adequate number of measurements and sparser channel models (e.g., CM-1 and CM-2). Based on the results of this study, the BCS method or the ℓ_1-*norm* minimization method can be preferred over the other one for different system implementation conditions.

Keywords Bayesian compressive sensing (BCS) · IEEE 802.15.4a channel models · ℓ_1-*norm* minimization · Mean-square error (MSE) lower bound · Ultra-wideband (UWB) channel estimation

1 Introduction

Ultra-wideband (UWB) impulse radio (IR) [1] is an emerging technology for wireless communications. Owing to distinguishing properties such as having low transmit power, low-cost simple structure, immunity to flat fading and capability of resolving multipath components individually with good time resolution, UWB-IR systems have received great interest from both academia and industry [2,3]. Considering these properties, UWB-IRs have been selected as the physical layer structure of wireless personal area network (WPAN) standard IEEE 802.15.4a for location and ranging, and low data rate applications [4,5]. In the implementation of UWB-IRs, one of the main challenges is the channel estimation [6]. Due to ultra-wide bandwidth of UWB-IRs, the main disadvantage of implementing the conventional maximum likelihood (ML) channel estimator is that very high sampling rates, i.e., very high speed A/D converters are required for precise channel estimation.

In order to overcome the high-rate sampling problem, compressive sensing (CS) theory proposed in [7,8] can be considered for UWB channel estimation. CS is a promising paradigm in signal processing, where a signal that is sparse in a known transform domain can be recovered with high probability from a set of random linear projections with much fewer measurements than usually required by the dimensions

M. Özgör · H. A. Çırpan
Department of Electronics and Communications Engineering, İstanbul Technical University, Maslak, 34469 Istanbul, Turkey
e-mail: ozgor@itu.edu.tr

H. A. Çırpan
e-mail: cirpanh@itu.edu.tr

S. Erküçük (✉)
Department of Electrical and Electronics Engineering, Kadir Has University, Fatih, 34083 Istanbul, Turkey
e-mail: serkucuk@khas.edu.tr

of this domain. As the received consecutive UWB pulses arrive with a considerable time delay and can be resolved individually at the receiver, sparse structure assumption is widely accepted for UWB multipath channels. Accordingly, CS has been exploited for UWB channel estimation [9,10], where the conventional ℓ_1-*norm* minimization method has been used to estimate UWB channel coefficients.

Among various implementations of CS, one approach has been to include the Bayesian model. Considering the sparse Bayesian model in [11], a Bayesian framework has been developed for CS in [12]. In [13], a hierarchical form of Laplace priors on signal coefficients is taken into consideration for Bayesian CS (BCS). Both of the frameworks have shown potential to improve signal recovery as the posterior density function over the associated sparse channel coefficients is considered. In [14], a Turbo BCS algorithm for sparse signal reconstruction through exploiting and integrating spatial and temporal redundancies in multiple sparse signal reconstruction is proposed. In [15], the Laplace prior based BCS algorithm in [13] has been modified for joint reconstruction of received sparse signals and channel parameters for multiuser UWB communications. In [16], the proposed approach in [12] is considered for UWB channel estimation, where BCS estimation results are compared to the ℓ_1-*norm* minimization results. However, the authors have not considered the effects of UWB channel models (i.e., sparsity condition) or additive noise level (i.e., Bayesian approach depends on the statistical information about channel parameters and additive noise) on the channel estimation performance.

In this paper, motivated by investigating the factors that affect the performance of BCS in realistic UWB channels, we study the effects of standardized IEEE 802.15.4a channel models, signal-to-noise ratio (SNR) regions, and number of measurements on the channel estimation performance. These factors are important to analyze as sparsity, noise level and measurements directly affect the BCS model. Accordingly, BCS channel estimation performance for various scenarios is compared to the ℓ_1-*norm* minimization based estimation [17], which is a method widely used for sparse channel estimation. Furthermore, it is important to specify a lower bound on the estimation error as a benchmark for the performance analysis of BCS estimators. Posterior Cramér-Rao lower bound (PCRLB), also referred to as the Bayesian CRLB, is a widely used bound that defines a lower bound on the mean-square error (MSE) of *unbiased* Bayesian estimators [18]. Indeed, CRLB is a lower bound *only* on the total variance of *unbiased* estimators [19], where MSE becomes equal to the variance for *unbiased* estimators. However, for *biased* estimators the bias term should be taken into account in addition to the variance of the estimator. By considering the bound in [20], we will present an MSE lower bound for *biased* Bayesian estimators with linear bias vectors to compare with the actual channel estimation performance of BCS. In addition, computation efficiency of BCS over the ℓ_1-*norm* minimization will be justified in terms of computation time by making use of the big-O notation. The comparison results provided are important in order to define the conditions where BCS may be preferred over the conventional ℓ_1-*norm* minimization method.

The rest of the paper is organized as follows. In Sect. 2, IEEE 802.15.4a channel models that are widely used in UWB communications are explained. In Sect. 3, the overview of CS theory, ℓ_1-*norm* minimization, Bayesian model and their applications to UWB channel estimation are presented. In Sect. 4, an MSE performance bound for a *biased* BCS estimator is provided. In Sect. 5, simulation results for performance comparison are presented. In Sect. 6, computation efficiencies of both BCS and ℓ_1-*norm* minimization are compared. Concluding remarks are given in Sect. 7.

2 UWB channel model

In this section, the discrete-time equivalent UWB channel model and the standardized IEEE 802.15.4a channel models are presented, respectively.

In order to obtain the discrete-time channel model, the general channel impulse response (CIR) should be presented first. Accordingly, the continuous-time channel $h(t)$ can be modeled as

$$h(t) = \sum_{k=1}^{L_r} h_k \delta(t - \tau_k), \tag{1}$$

where h_k represents the kth multipath gain coefficient, τ_k is the delay of the kth multipath component, $\delta(\cdot)$ is the Dirac delta function and L_r is the number of resolvable multipaths.

The continuous-time CIR given in (1) assumes that multipaths may arrive any time. This is referred to as the τ-spaced channel model [21]. If a pulse is T_s-seconds duration, then an approximate equivalent channel model can be obtained for practical purposes. Hence, the equivalent T_s-spaced channel model can be expressed as

$$h(t) = \sum_{n=1}^{N} c_n \delta(t - nT_s), \tag{2}$$

where $T_c = NT_s$ is the channel length and $\{c_n\}$'s are the resulting new channel coefficients [10]. Using (2), the discrete-time equivalent channel can be written as

$$\mathbf{h} = [c_1, c_2, \ldots, c_N]^T, \tag{3}$$

where the channel resolution is T_s. Assuming that $\mathbf{h}$ has K nonzero coefficients, the sparsity assumption of (3) is valid if $K \ll N$.

Based on the discrete-time equivalent channel model above, the UWB channels are widely accepted as having a sparse structure. This assumption for UWB channels plays an important role in CS based UWB channel estimation. However, the channel environment should be inspected to prove this assumption. In [22], a comprehensive model for UWB propagation channels, which was accepted as the standardized channel model for IEEE 802.15.4a, has been developed considering various channel environments and conducting different measurement campaigns. These environments include indoor residential, indoor office, outdoor, industrial environments, agricultural areas and body area networks with having either a line-of-sight (LOS) or a non-LOS (NLOS) transmitter-receiver connection. In [10], the sparsity assumption of UWB channels has been discussed over the widely used channel models CM-1 (LOS residential indoor), CM-2 (NLOS residential indoor), CM-5 (LOS outdoor) and CM-8 (NLOS industrial). In order to investigate the effects of channel sparsity on the BCS channel estimation performance, we will consider the same channel models in the current study. More details on the channel models CM-1, CM-2, CM-5 and CM-8 can be found in [10] and [22].

3 CS for UWB channel estimation

Assuming that the UWB channels are sparse, CS can be employed for UWB channel estimation in order to overcome the high-rate sampling problem. In the following, we will present the overview of CS theory and its application to UWB channel estimation, and the Bayesian CS model, respectively.

3.1 Overview of compressive sensing

Consider the problem of reconstructing a discrete-time signal $\mathbf{x} \in \Re^N$ which can be represented in an arbitrary basis $\boldsymbol{\Psi} \in \Re^{N \times N}$ with the weighting coefficients $\boldsymbol{\theta} \in \Re^N$ as

$$\mathbf{x} = \sum_{n=1}^{N} \psi_n \theta_n = \boldsymbol{\Psi}\boldsymbol{\theta}. \tag{4}$$

Suppose that $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_N]^T$ has only K nonzero coefficients, where $K \ll N$ and $\boldsymbol{\Psi} = [\psi_1, \psi_2, \ldots, \psi_N]$, $\psi_n \in \Re^N$. As $\mathbf{x}$ is a linear combination of only K basis vectors, it can be called a K-sparse signal and can be expressed as

$$\mathbf{x} = \sum_{i=1}^{K} \psi_{n_i} \theta_{n_i}, \tag{5}$$

where $\{n_i\}$'s are the indices that correspond to nonzero coefficients. By projecting $\mathbf{x}$ onto a random measurement matrix $\boldsymbol{\Phi} \in \Re^{M \times N}$, a set of measurements $\mathbf{y} \in \Re^M$ can be obtained as

$$\mathbf{y} = \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\theta}, \tag{6}$$

where $M \ll N$. Here, the measurement matrix should be incoherent with the basis in addition to the sparsity condition for accurately estimating the weighting coefficients. The incoherency is usually achieved by random matrices with independent identically distributed (i.i.d) elements from Gaussian or Bernoulli distributions [23]. Instead of using the N-sample $\mathbf{x}$ to estimate the weighting coefficients $\boldsymbol{\theta}$, the M-sample measurement vector $\mathbf{y}$ can be used. Accordingly, $\boldsymbol{\theta}$ can be estimated as

$$\hat{\boldsymbol{\theta}} = \text{argmin} \, \|\boldsymbol{\theta}\|_1 \quad \text{subject to} \quad \mathbf{y} = \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\theta} \,, \tag{7}$$

where ℓ_p-*norm* is denoted as $\|\boldsymbol{\theta}\|_p = \left(\sum_{n=1}^{N} |\theta_n|^p\right)^{\frac{1}{p}}$. The reconstruction problem hence becomes an ℓ_1-norm optimization problem, and estimating $\boldsymbol{\theta}$ from the vector $\mathbf{y}$ instead of $\mathbf{x}$ corresponds to a lower sampling rate at the receiver.

The CS theory explained in (4)–(7) can be employed to UWB channel estimation. Suppose that $\mathbf{g} \in \Re^N$ is the discrete-time representation of the received signal given as

$$\mathbf{g} = \mathbf{P}\mathbf{h} + \mathbf{n}, \tag{8}$$

where $\mathbf{P} \in \Re^{N \times N}$ is a scalar matrix representing the time-shifted pulses, $\mathbf{h} = [c_1, c_2, \ldots, c_N]^T$ are the channel gain coefficients, and $\mathbf{n}$ are the additive white Gaussian noise (AWGN) terms. Since the UWB channel structure is sparse, $\mathbf{h}$ has only K nonzero coefficients. Similar to (6), the received signal $\mathbf{g}$ can be projected onto a random measurement matrix $\boldsymbol{\Phi} \in \Re^{M \times N}$ so as to obtain $\mathbf{y} \in \Re^M$ as

$$\begin{aligned}\mathbf{y} &= \boldsymbol{\Phi}\mathbf{P}\mathbf{h} + \boldsymbol{\Phi}\mathbf{n} \\ &= \mathbf{A}\mathbf{h} + \mathbf{z}.\end{aligned} \tag{9}$$

Due to the presence of the noise term $\mathbf{z}$, the channel $\mathbf{h}$ can be estimated as

$$\hat{\mathbf{h}} = \text{argmin} \, \|\mathbf{h}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{h}\|_2 \leq \epsilon \,, \tag{10}$$

where ϵ is related to the noise term as $\epsilon \geq \|\mathbf{z}\|_2$. The ℓ_1-*norm* minimization problem in (10) can be recast as a second-order cone program (SOCP) and solved[1] with a generic log-barrier algorithm.

[1] For the implementation of (10), the codes provided by Romberg and Candes publicly availble at http://users.ece.gatech.edu/~justin/l1magic/ are used.

3.2 Bayesian compressive sensing

In this section, the CS problem will be presented from a Bayesian perspective for UWB channel estimation. In the BCS framework proposed in [11,12], the statistical information about the compressible signal and the additive noise is considered, where ℓ_1-*norm* minimization does not consider these factors. Considering sparsity prior of $\mathbf{h}$ and the noise model assumption together with the signal model in (9), BCS can be used[2] for UWB channel estimation. Taking into consideration (9), the full posterior distribution over all unknowns of interest for the problem at hand becomes

$$p(\mathbf{h}, \boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{h}, \boldsymbol{\beta}, \sigma^2)\ p(\mathbf{h}, \boldsymbol{\beta}, \sigma^2)}{p(\mathbf{y})}, \tag{11}$$

where $\boldsymbol{\beta}$ represents hyperparameters that control the inverse variance of each channel coefficient, and σ^2 is the variance of each noise term in $\mathbf{z}$. Unfortunately, this full posterior term is not tractable since the integral

$$p(\mathbf{y}) = \int\int\int p(\mathbf{y} \mid \mathbf{h}, \boldsymbol{\beta}, \sigma^2)\ p(\mathbf{h}, \boldsymbol{\beta}, \sigma^2)\ d\mathbf{h}\, d\boldsymbol{\beta}\, d\sigma^2 \tag{12}$$

cannot be computed analytically. Hence, we decompose the full posterior distribution as

$$p(\mathbf{h}, \boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) \equiv p(\mathbf{h} \mid \mathbf{y}, \boldsymbol{\beta}, \sigma^2)\ p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}). \tag{13}$$

In (9), the noise term $\mathbf{z}$ can be modeled probabilistically as independent zero-mean Gaussian random variables:

$$p(\mathbf{z}) = \prod_{m=1}^{M} \mathcal{N}(z_m \mid 0, \sigma^2). \tag{14}$$

This noise model infers Gaussian likelihood for observation $\mathbf{y}$:

$$p(\mathbf{y} \mid \mathbf{h}, \sigma^2) = (2\pi\sigma^2)^{-M/2} \exp\left(\frac{-\|\mathbf{y} - \boldsymbol{\Phi}\mathbf{h}\|^2}{2\sigma^2}\right). \tag{15}$$

Since this Gaussian likelihood is inferred by AWGN term $\mathbf{z}$, a conjugate[3] prior distribution has to be defined for computational convenience so that the associated Bayesian inference may be performed in closed form [24]. Therefore, suppose that a zero-mean Gaussian prior distribution is defined on channel coefficients with $\{\beta_n\}$:

$$p(\mathbf{h} \mid \boldsymbol{\beta}) = \prod_{n=1}^{N} \mathcal{N}(h_n \mid 0, \beta_n^{-1}) = (2\pi)^{-N/2} \prod_{n=1}^{N} \beta_n^{1/2} \exp\left(-\frac{\beta_n h_n^2}{2}\right). \tag{16}$$

$\{\beta_n\}$'s are independent hyperparameters that form the $\boldsymbol{\beta} = [\beta_1, ..., \beta_N]^T$ vector and control the strength of the prior over associated channel coefficients individually.

The first term of (13), $p(\mathbf{h} \mid \mathbf{y}, \boldsymbol{\beta}, \sigma^2)$, the posterior distribution over the channel coefficients, can be expressed via Bayes' rule as

$$p(\mathbf{h} \mid \mathbf{y}, \boldsymbol{\beta}, \sigma^2) = \frac{p(\mathbf{y} \mid \mathbf{h}, \sigma^2)\ p(\mathbf{h} \mid \boldsymbol{\beta})}{p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2)}. \tag{17}$$

Considering Gaussian likelihood together with Gaussian prior, this posterior distribution is also $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where

$$\boldsymbol{\Sigma} = (\boldsymbol{\Lambda} + \sigma^{-2}\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}, \quad \boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T\mathbf{y}, \tag{18}$$

with $\boldsymbol{\Lambda} = diag(\beta_1, \beta_2, \ldots, \beta_N)$ and is analytically tractable. To compute the full posterior distribution approximately, hyperparameter posterior $p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y})$, the second term in (13), needs to be approximated. This approximation is provided by type-II ML procedure. This procedure, also known as the evidence approximation or the emprical Bayes, is used to estimate hyperparameters by maximizing the marginal likelihood function (LF) [25]. According to the Bayes' theorem, hyperparameter posterior $p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y})$ can be expressed as:

$$p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2)\ p(\boldsymbol{\beta}, \sigma^2). \tag{19}$$

Using appropriately selected uniform[4] hyperpriors for $\boldsymbol{\beta}$ and σ^2 (*i.e.*, $p(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2)$), the estimates of $\boldsymbol{\beta}$ and σ^2 can be found by maximizing marginal likelihood function (LF) $p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2)$ as a consequence of type-II ML procedure. The marginal LF can be obtained by integrating over the channel coefficients $\mathbf{h}$ as:

$$p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2) = \int_{-\infty}^{\infty} p(\mathbf{y} \mid \mathbf{h}, \sigma^2)\ p(\mathbf{h} \mid \boldsymbol{\beta}) d\mathbf{h}. \tag{20}$$

Maximization of the marginal LF with respect to $\boldsymbol{\beta}$ or equivalently, its logarithm can be expressed as:

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2) = \log p(\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2) = \log \int_{-\infty}^{\infty} p(\mathbf{y} \mid \mathbf{h}, \sigma^2)\ p(\mathbf{h} \mid \boldsymbol{\beta}) d\mathbf{h} = -\frac{1}{2}\left[M\log(2\pi) + \log|\boldsymbol{C}| + \mathbf{y}^T\boldsymbol{C}^{-1}\mathbf{y}\right] \tag{21}$$

[2] For the implementation of BCS, the codes provided by Shihao Ji publicly available at http://people.ee.duke.edu/~lcarin/BCS.html are used.

[3] In Bayesian probability theory, if the resulting posterior distributions $p(\mathbf{h}|\,\mathbf{y})$ are in the same class as prior probability distributions $p(\mathbf{h})$, then that class of $p(\mathbf{h})$ is said to be conjugate to the class of likelihood functions $p(\mathbf{y}|\,\mathbf{h})$ [12].

[4] Uniform or flat hyperpriors are known as noninformative hyperpriors [24] which have a minimum effect on the hyperparameter posterior and they can be ignored.

where $\boldsymbol{C} = \sigma^2\mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Phi}^T$ and $\mathbf{I} \in \Re^{M\times M}$ is an identity matrix. Differentiating $\mathcal{L}(\boldsymbol{\beta}, \sigma^2)$ with respect to $\boldsymbol{\beta}$ and σ^2, and equating it to zero yields the following expressions which can be solved iteratively:

$$\boldsymbol{\beta}_n^{new} = \frac{\gamma_n}{\mu_n^2}, \qquad \sigma^{2^{new}} = \frac{\|\mathbf{y} - \boldsymbol{\Phi}\boldsymbol{\mu}\|_2^2}{M - \sum_{n=1}^{N}\gamma_n}, \tag{22}$$

where $\gamma_n \in [0, 1]$ is defined as $\gamma_n = 1 - \beta_n \sum_{nn}$ with $\sum_{nn}$ being the nth diagonal element of the posterior coefficient covariance from (18) and μ_n is the nth posterior coefficient mean from (18).

By employing re-estimates of hyperparameters, an iterative systematic approach is used to determine which basis vectors should be included in the model and which should be removed to promote sparsity [12]. Further details and steps of the BCS algorithm can be found in [11].

4 Performance bound

As in any estimation application, it is useful to quantify the best performance that may be achieved from channel estimator approach proposed. Performance bounds can serve as a benchmark with the goal of facilitating performance comparisons of the various estimation techniques under consideration. Such bounds may also indicate characteristics of the problem that require extra attention for optimal performance.

The CRLB is a widely used performance bound in order to indicate the minimum achievable total variance of any *unbiased* estimator of deterministic parameter vector [19]. Since MSE becomes equal to variance for *unbiased* (zero bias) estimators, CRLB also provides a benchmark on the estimation error for this type of estimators. However, the BCS estimator proposed for UWB channel estimation in our study is a *biased* estimator as well as being Bayesian. Accordingly, restricting ourselves to *unbiased* approach of the lower bound for the problem at hand leads to unreasonable performance results. It is necessary to determine a lower bound on the estimation error which characterizes both the total variance and the bias of the *biased* estimator. Hence, we will provide an MSE lower bound for *biased* Bayesian estimators ($MSE_{l,b}$, where subscript l stands for lower bound and subscript b stands for *biased* Bayesian) by making use of bound in [20], which is based on *biased* CRLB in [26]. In literature, PCRLB or Bayesian CRLB [18] was defined for *unbiased* Bayesian estimators considering prior information about the parameter vector that we want to estimate. In addition to CRLB, PCRLB also takes into account prior probability distribution of the parameter vector. Nevertheless, PCRLB is a lower bound on the variance of the *unbiased* Bayesian estimator not on the estimation error. Accordingly, the $MSE_{l,b}$ that we will provide considering bias with the prior information of channel vector will become a lower bound on the estimation error of *biased* Bayesian estimators. Note that the bias and the prior distribution of the parameter vector are included in the derivation of the performance bound presented below, however, the sparsity conditions are not incorporated into the model and are subject for future research. Next, the MSE of *general* and *Bayesian* biased estimators are presented, respectively.

4.1 MSE of a biased estimator

In what follows, MSE of the *biased* estimator is expressed as a sum of the squared norm of bias and trace of covariance matrix for the channel vector $\mathbf{h}$ with given linear signal model in (9),

$$MSE(\hat{\mathbf{h}}_b) = E\left\{\left\|(\hat{\mathbf{h}}_b) - \mathbf{h}\right\|^2\right\} = \|\mathbf{b}(\mathbf{h})\|^2 + Tr(\mathbf{C}_{\hat{\mathbf{h}}_b}), \tag{23}$$

where bias vector, $\mathbf{b}(\mathbf{h}) \in \Re^N$, and covariance matrix of the *biased* estimator, $\mathbf{C}_{\hat{\mathbf{h}}_b} \in \Re^{N\times N}$, can be denoted respectively as

$$\mathbf{b}(\mathbf{h}) = E\left\{\hat{\mathbf{h}}_b\right\} - \mathbf{h}, \tag{24}$$

$$\mathbf{C}_{\hat{\mathbf{h}}_b} = \mathrm{Cov}(\hat{\mathbf{h}}_b) = E\left\{\left[\hat{\mathbf{h}}_b - E(\hat{\mathbf{h}}_b)\right]\left[\hat{\mathbf{h}}_b - E(\hat{\mathbf{h}}_b)\right]^T\right\}, \tag{25}$$

and $\hat{\mathbf{h}}_b \in \Re^N$ corresponds to estimated channel vector. Regarding suitability of regularity condition on $p(\mathbf{y}|\mathbf{h})$ [19]

$$E_{\mathbf{y}}\left[\frac{\partial \ln p(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}}\right] = 0 \qquad \forall\, \mathbf{h}, \tag{26}$$

biased CRLB in [26] for any *biased* estimator with a given bias can be obtained for the vector case as follows:

$$\mathbf{C}_{\hat{\mathbf{h}}_b} \geq \left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}\right)\left(E_{\mathbf{y}|\mathbf{h}}\left\{\left[\frac{\partial \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}}\right]^T \times \left[\frac{\partial \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}}\right]\right\}\right)^{-1}\left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}\right)^T \tag{27}$$

where $\frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}$ represents the bias gradient matrix as will be explained while presenting the assumption on the bias vector. The second term in (27) can also be denoted as

$$E_{\mathbf{y}|\mathbf{h}}\left\{\left[\frac{\partial \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}}\right]^T\left[\frac{\partial \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}}\right]\right\} = -E_{\mathbf{y}|\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}^T\partial \mathbf{h}}\right]\right\}. \tag{28}$$

4.2 MSE of Bayesian biased estimator

Counterpart of the *biased* CRLB in Bayesian framework can be expressed for Bayesian estimators as

$$\mathbf{C}_{\hat{\mathbf{h}}_b} \geq \left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}\right)\left(-E_{\mathbf{y},\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{y},\mathbf{h}}(\mathbf{y},\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right]\right\}\right)^{-1} \times \left(\mathbf{I} + \frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}\right)^T. \quad (29)$$

Moreover, we can decompose the second term in (29) into two parts using the Bayes' rule:

$$-E_{\mathbf{y},\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{y},\mathbf{h}}(\mathbf{y},\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right]\right\} = -E_{\mathbf{y}|\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right]\right\} - E_{\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{h}}(\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right]\right\} \quad (30)$$

which can be expressed in matrix form as

$$\mathbf{J}_H = \mathbf{J}_D + \mathbf{J}_P, \quad (31)$$

where $\mathbf{J}_H \in \Re^{N\times N}$, $\mathbf{J}_D \in \Re^{N\times N}$ and $\mathbf{J}_P \in \Re^{N\times N}$ correspond to Bayesian Fisher information matrix (FIM), observation data ($\mathbf{y}$) information matrix and prior information matrix, respectively. Considering our linear signal model in (9) with (14) and (15), observation data information matrix $\mathbf{J}_D$ can be expressed as

$$\mathbf{J}_D = -E_{\mathbf{y}|\mathbf{h}}\left\{\left[\frac{\partial^2 \ln p_{\mathbf{y}|\mathbf{h}}(\mathbf{y}\,|\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right]\right\} = \mathbf{A}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{A}, \quad (32)$$

where $\mathbf{C}_{\mathbf{z}} = \sigma^2 \mathbf{I} \in \Re^{M\times M}$ is the covariance matrix of the noise term $\mathbf{z}$ and $\mathbf{A} \in \Re^{M\times N}$ is the measurement matrix which is also a full rank matrix. Exploiting assumption ($\mathbf{h} \sim \mathcal{N}(0, \mathbf{C}_{\mathbf{h}})$) in (16), prior information matrix $\mathbf{J}_P$ is equal to inverse of covariance matrix of the channel vector $\mathbf{C}_{\mathbf{h}} \in \Re^{N\times N}$:

$$\mathbf{J}_P = -E_{\mathbf{h}}\left\{\frac{\partial^2 \ln p_{\mathbf{h}}(\mathbf{h})}{\partial \mathbf{h}^T \partial \mathbf{h}}\right\} = \mathbf{C}_{\mathbf{h}}^{-1}. \quad (33)$$

$\mathbf{C}_{\mathbf{h}}$ is a diagonal matrix and each diagonal element is formed by inverse of the hyperparameters

$$\mathbf{C}_{\mathbf{h}} = diag\left\{\beta_n^{-1}\right\}, \qquad n \in \{1, 2, ..., N\}. \quad (34)$$

Once $\mathbf{J}_D$ and $\mathbf{J}_P$ are obtained, the Bayesian FIM $\mathbf{J}_H$ can be rewritten in compact form as

$$\mathbf{J}_H = \mathbf{A}^T \mathbf{C}_{\mathbf{z}}^{-1} \mathbf{A} + \mathbf{C}_{\mathbf{h}}^{-1}. \quad (35)$$

Since the denominator of (29) is obtained, to form a final expression for the *biased* Bayesian CRLB, an a-priori choice of the bias gradient is required. In [20], estimators with only linear bias vectors are considered instead of taking into account all possible estimators. For its simplicity and tractability, we also consider only linear bias vectors in this study. Advantages of restricting attention to linear bias vectors can be found in [20]. Linear bias vector can be denoted as

$$\mathbf{b}(\mathbf{h}) = \mathbf{S}\mathbf{h}, \quad (36)$$

where $\mathbf{S} \in \Re^{N\times N}$ is the bias gradient matrix defined by

$$\mathbf{S} = \frac{\partial \mathbf{b}(\mathbf{h})}{\partial \mathbf{h}}. \quad (37)$$

Thus, (29) can be rearranged as

$$\mathbf{C}_{\hat{\mathbf{h}}_b} \geq (\mathbf{I} + \mathbf{S})\, \mathbf{J}_H^{-1}\, (\mathbf{I} + \mathbf{S})^T, \quad (38)$$

where $\mathbf{I} \in \Re^{N\times N}$ is an identity matrix. Inserting (36) and (38) into (23), the $MSE_{l,b}$ for *biased* Bayesian estimators can be obtained as

$$MSE_{l,b} = E\left\{\left\|(\hat{\mathbf{h}}_b) - \mathbf{h}\right\|^2\right\} = \mathbf{h}^T \mathbf{S}^T \mathbf{S}\mathbf{h} + Tr\left\{(\mathbf{I} + \mathbf{S})\, \mathbf{J}_H^{-1}\, (\mathbf{I} + \mathbf{S})^T\right\}. \quad (39)$$

Now, the optimal $\mathbf{S}$ matrix needs to be determined to find the achievable smallest MSE over all estimators with linear bias. Since (39) is convex in $\mathbf{S}$, the smallest value of $MSE_{l,b}$ can be found by equating its derivative to zero

$$\frac{\partial\left[\mathbf{h}^T \mathbf{S}^T \mathbf{S}\mathbf{h} + Tr\left\{(\mathbf{I} + \mathbf{S})\mathbf{J}_H^{-1}(\mathbf{I} + \mathbf{S})^T\right\}\right]}{\partial \mathbf{S}} = 0,$$
$$2\mathbf{h}^T \mathbf{h}\mathbf{S} + 2\mathbf{J}_H^{-1} + 2\mathbf{J}_H^{-1}\mathbf{S} = 0, \quad (40)$$

which yields

$$\mathbf{S}(\mathbf{J}_H^{-1} + \mathbf{h}^T \mathbf{h}) = -\mathbf{J}_H^{-1}. \quad (41)$$

Multiplying both sides of (41) with $(\mathbf{J}_H^{-1} + \mathbf{h}^T \mathbf{h})^{-1}$ leaves the matrix $\mathbf{S}$ alone at the left side in (41). Using the matrix inversion lemma, $(\mathbf{J}_H^{-1} + \mathbf{h}^T \mathbf{h})^{-1}$ can be expressed as

$$(\mathbf{J}_H^{-1} + \mathbf{h}^T \mathbf{h})^{-1} = \mathbf{J}_H - \frac{\mathbf{J}_H \mathbf{h}\mathbf{h}^T \mathbf{J}_H}{1 + \mathbf{h}^T \mathbf{J}_H \mathbf{h}}. \quad (42)$$

After multiplying the right side of (41) with (42), the optimal $\mathbf{S}$ matrix can be obtained as follows:

$$\mathbf{S} = -\mathbf{I} + \frac{1}{1 + \mathbf{h}^T \mathbf{J}_H \mathbf{h}} \mathbf{h}\mathbf{h}^T \mathbf{J}_H. \quad (43)$$

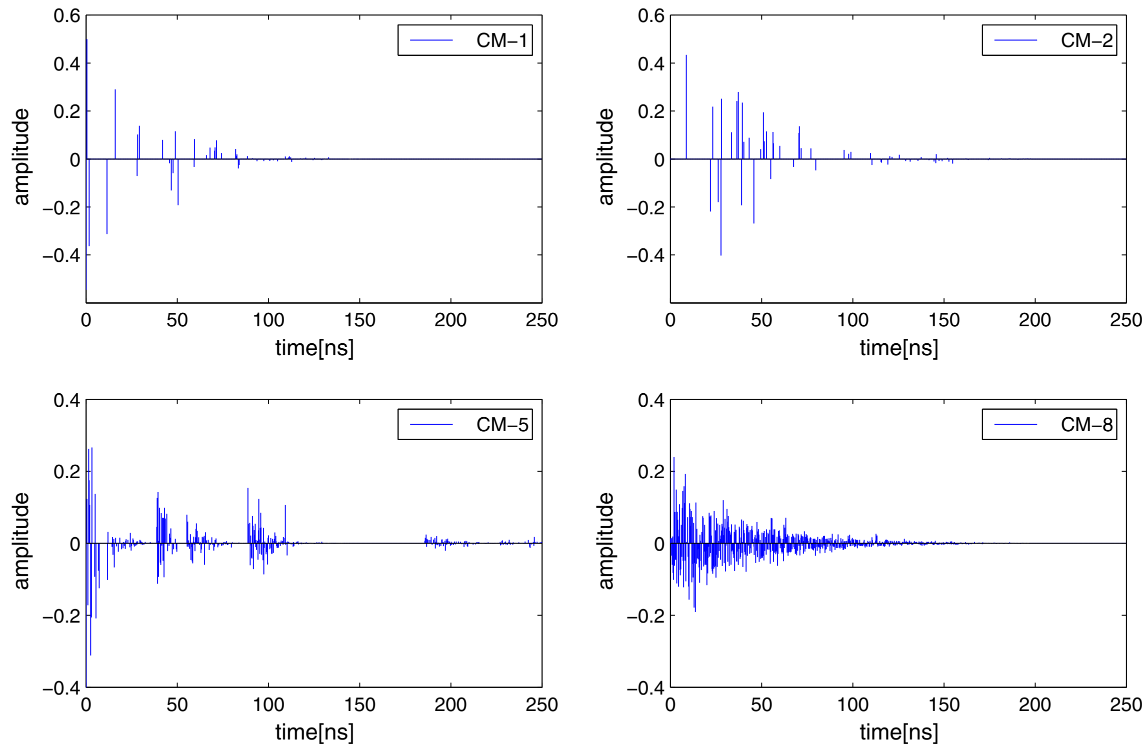

Fig. 1 Realizations of channel models for T_c = 250 ns and T_s = 0.25 ns

Note that when $\mathbf{S} = 0$, which is the zero bias case, CRLB for *unbiased* Bayesian estimators (i.e., PCRLB) is obtained: $MSE_{l,b=0} = Tr(\mathbf{J}_H^{-1})$. Therefore $MSE_{l,b}$ also includes *unbiased* Bayesian estimation as a special case.

5 Performance results

In this section, we investigate the effects of number of measurements, SNR regions, and the IEEE 802.15.4a channel models on the BCS channel estimation performance, and compare the results to the performance of the ℓ_1-*norm* minimization results. As the performance measure, we evaluate the MSE of the estimated channel vector. To remove the path loss effect and to treat each channel model fairly, we normalize the channel coefficients as $\sum_{n=1}^{N} c_n^2 = 1$. For the simulations, the channel length and resolution are fixed to T_c = 250 ns and T_s = 0.25 ns, respectively, resulting in the discrete-time channel length $N = T_c/T_s$ = 1,000. According to these T_c and T_s values, single channel realizations of CM-1, CM-2, CM-5 and CM-8 with the parameters given in [22] are plotted in Fig. 1 for illustrative purposes. The performances are evaluated for $M = \{250, 500, 750\}$ measurements in the [0, 30]dB SNR region. Here, M/N can be regarded as the compression ratio (i.e., the ratio of number of measurements to the length of the equivalent discrete-time channel) and K/N can be regarded as the sparsity ratio (i.e., the ratio of number of nonzero coefficients to the length of the equivalent discrete-time channel). The channel models' sparsity ratios, which are acquired by averaging over 200 channel realizations, for fixed T_c and T_s values are given in Table 1. The elements of the measurement matrix $\mathbf{\Phi}$ are obtained from the $\mathcal{N}(0, 1)$ distribution, and the basis where the channel vector is sparse is defined as $\mathbf{\Psi} = \mathbf{I}$ in our simulations.

Table 1 Sparsity ratios of channel models when T_c = 250 ns and T_s = 0.25 ns

Channel model	Sparsity ratio (K/N)
CM-1	0.06
CM-2	0.09
CM-5	0.47
CM-8	0.79

In Figs. 2, 3, 4 and 5, the channel estimation performances of BCS and ℓ_1-*norm* minimization are compared for various number of measurements and SNR values for the channel models CM-1, CM-2, CM-5 and CM-8, respectively. The best channel estimation performance for both methods is obtained for CM-1, as it exhibits the sparsest structure among these channel models (see Fig. 1; Table 1). BCS outperforms ℓ_1-*norm* minimization in the sparser channel models CM-1 and CM-2 for SNR values greater than 12-13dB for all measurements considered. This can be explained as for the

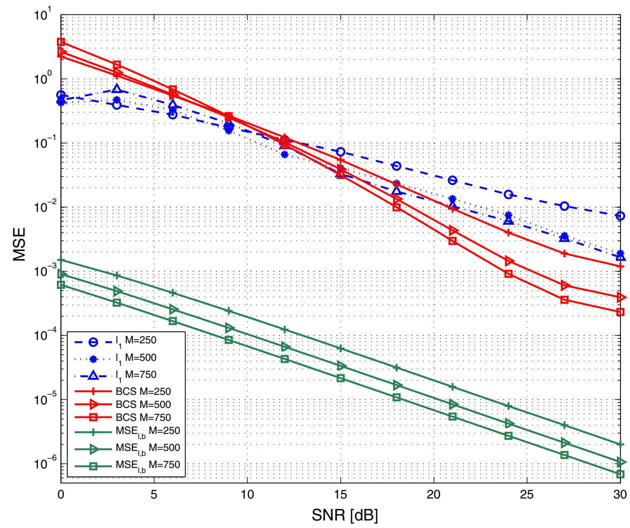

Fig. 2 MSE performance comparison of BCS and ℓ_1-*norm* minimization for CM-1

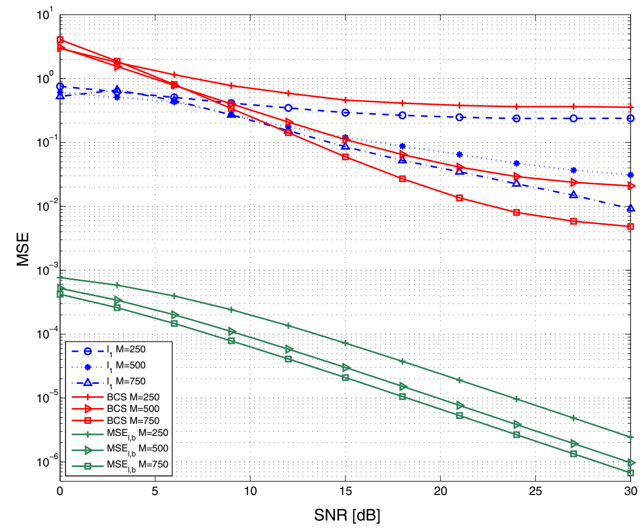

Fig. 4 MSE performance comparison of BCS and ℓ_1-*norm* minimization for CM-5

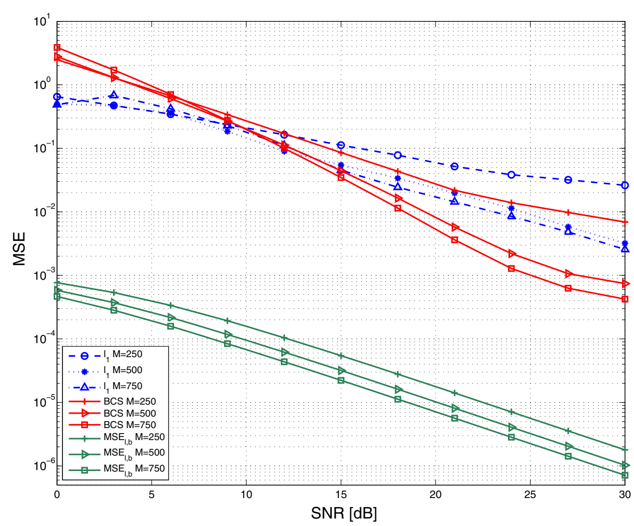

Fig. 3 MSE performance comparison of BCS and ℓ_1-*norm* minimization for CM-2

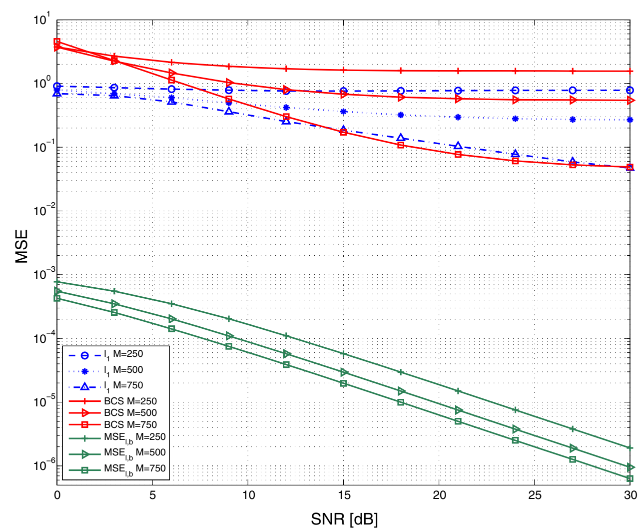

Fig. 5 MSE performance comparison of BCS and ℓ_1-*norm* minimization for CM-8

higher SNR regions posterior density function over the channel coefficients and noise is beneficial to the channel coefficient estimation, whereas for lower SNR regions the uncertainty in the estimation is higher. As for CM-5, which is a less sparse channel, the number of measurements should be greater than $M = 500$ in order for BCS to have a superior performance at higher SNR regions. As for CM-8, which is not a sparse channel model, as the multipaths arrive almost in every time bin, the BCS performs inferior compared to the ℓ_1-*norm* minimization for almost all conditions. In summary, BCS can be an effective channel estimation method for sparser channel models at high SNR regions. This is mainly due to BCS considering the channel and noise statistics and providing a posterior density function over noise and the channel coefficients, whereas the ℓ_1-*norm* minimization method not utilizing such statistics.

Next, we compare the MSE performance of BCS with the MSE lower bound, $MSE_{l,b}$, in Figs. 2, 3, 4 and 5. It can be observed that the MSE lower bound performance improves with the number of measurements M as expected. On the other hand, for M fixed the MSE bounds are similar for different channel models. This can be explained as follows. When quantified, the $MSE_{l,b}$ term in (39) is observed to be dominated by the second term, which depends on $\mathbf{J}_H = \mathbf{J}_D + \mathbf{J}_P$. Here, the observation data information matrix $\mathbf{J}_D$ has more significant contribution compared to the prior information matrix $\mathbf{J}_P$ that carries the channel model information. Therefore, $MSE_{l,b}$ values have found to be similar despite channel model differences. Lastly, we observed a performance gap between the MSE performance of BCS and the MSE lower bound as in [18], where they compared their proposed CS based block maximum-a-posteriori least mean squares (CS-BMAP-LMS) method to the Bayesian CRLB. A tighter bound for our implementation may be obtained if the sparsity knowledge of the channel can be incorporated into the lower bound computation and the linearly assumed bias vector can be generalized to cover nonlinear bias vectors. Both considerations are non-trivial to implement, however, are expected to provide tighter bounds and subject to further investigation. In the next section, we will present the second part of simulation results, which is related to computation times of both methods.

6 Computation efficiency

Before presenting numerical values for computation times, we provide a short discussion on the comparison of computation efficiencies of both BCS and ℓ_1-*norm* minimization in this section.

In ℓ_1-*norm* minimization, whose computational complexity is proportional to $O(N^3)$ [27], the basis vectors are added to the model and never removed during the channel coefficient estimation. Therefore, not only the K basis vectors that correspond to nonzero coefficients but all basis vectors are considered during the channel estimation process. This situation apparently increases the computational complexity of this method. However, in BCS, whose computational complexity is proportional to $O(NK^2)$, there is an iterative update approach which sequentially adds or removes basis vectors to the model until all K basis vectors have been included [12]. Thus, BCS is computationally more efficient compared to the ℓ_1-*norm* minimization.

To justify this argument, computation times of both methods are provided. The average computation times of the channel estimators for both methods are compared based on the publicly available codes, where their main structures are not modified but adapted to IEEE 802.15.4a channel estimation. In Tables 2, 3, 4 and 5, the computation times of both methods are presented for different number of measurements in CM-1, CM-2, CM-5 and CM-8, respectively. The simulations were run on a computer that has a 3.4 GHz Intel Core i7 CPU and a 3.88 GB RAM. It can be observed that the computation time of BCS is significantly shorter than the ℓ_1-*norm* minimization for every channel model and number of measurements. It can be further observed that the computation time of ℓ_1-*norm* minimization does not change much with sparsity or the number of measurements. This can be explained by the computational complexity of ℓ_1-*norm* minimization not depending on the number of nonzero coefficients (K) but only on the discrete-time channel length (N), which is the same for all channel models considered. Unlike ℓ_1-*norm* minimization, the computational complexity of BCS depends on K, and therefore, the computation time of BCS changes remarkably with sparsity and the number of measurements. The computation times of both methods are summarized in Table 6 for different channel models when the number of measurements is fixed to $M = 250$. Considering CM-1, which has the sparsest structure, and CM-8, which has the least sparse structure among the channel models, the computation time of ℓ_1-*norm* minimization in CM-8 increases 6.53 % compared to CM-1 but for BCS this ratio becomes 104.24 %. Similar observations were made for the cases $M = \{500, 750\}$. This remarkable increase in the computation time of BCS is a result of its computational complexity depending on K. Nev-

Table 2 Computation times of both methods for CM-1

Number of measurements	ℓ_1-*norm* minimization (s)	Bayesian CS (s)
M = 250	3.5911	0.13607
M = 500	3.6684	0.2892
M = 750	3.5778	0.76564

Table 3 Computation times of both methods for CM-2

Number of measurements	ℓ_1-*norm* minimization (s)	Bayesian CS (s)
M = 250	3.6073	0.15767
M = 500	3.627	0.31896
M = 750	3.4591	0.82328

Table 4 Computation times of both methods for CM-5

Number of measurements	ℓ_1-*norm* minimization (s)	Bayesian CS (s)
M = 250	3.748	0.22791
M = 500	3.5745	0.47146
M = 750	3.2783	1.1099

Table 5 Computation times of both methods for CM-8

Number of measurements	ℓ_1-*norm* minimization (s)	Bayesian CS (s)
M = 250	3.8257	0.27791
M = 500	4.0806	0.84026
M = 750	3.6359	1.9952

Table 6 Computation times of both methods for channel models when M=250

Channel model	ℓ_1-*norm* minimization $\sim O(N^3)$ (s)	Bayesian CS $\sim O(NK^2)$ (s)
CM-1	3.5911	0.13607
CM-2	3.6073	0.15767
CM-5	3.748	0.22791
CM-8	3.8257	0.27791

ertheless, BCS is computationally more efficient compared to ℓ_1-*norm* minimization as shown with practical examples.

7 Conclusion

In this paper, we considered the application of Bayesian CS to UWB channel estimation, and studied its channel estimation performance for various UWB channel models and noise conditions. Specifically, we investigated the effects of the sparse structure of standardized IEEE 802.15.4a channel models, SNR regions, and number of measurements on the BCS channel estimation performance, and compared them to the results of the conventional ℓ_1-*norm* minimization based estimation. We also (i) provided an MSE lower bound on the estimation error for *biased* Bayesian estimators with linear bias vectors, and (ii) compared the computational efficiencies of both BCS and ℓ_1-*norm* minimization for channel estimation.

The results of this study show that BCS exhibits superior performance at sparser channel models and higher SNR regions as it utilizes the statistics of channel coefficients and noise. Furthermore, the computational efficiency of BCS has been found to be significantly better than ℓ_1-*norm* minimization for the cases considered. Based on the results of this study, the implementation conditions of BCS can be determined for practical cases.

References

1. Win, M. Z., & Scholtz, R. A. (2000). Ultra-widebandwidth time-hopping spread-spectrum impulse radio for wireless multiple-access communications. *IEEE Transactions on Communications*, *48*, 679–691.
2. Porcino, D., & Hirt, W. (2003). Ultra-wideband radio technology: Potential and challenges ahead. *IEEE Communications Magazine*, *41*(7), 66–74.
3. Ragoubi, K., Jin, M., Saha, G., & Yang, Y. (2011). Recent advances in UWB systems: Theory and applications. *Journal of Telecommunication Systems*. doi:10.1007/s11235-011-9628-8.
4. IEEE. (2007). Standart 802.15.4a-2007: Part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs).
5. De Nardis, L., Fiorina, J., Panaitopol, D., & Di Benedetto, M. G. (2011). Combining UWB with time reversal for improved communication and positioning. *Journal of Telecommunication Systems*. doi:10.1007/s11235-011-9630-1.
6. Islam, S. M. R., & Kwak, K. S. (2011). Preamble-based improved channel estimation for multiband UWB systems in presence of interferences. *Journal of Telecommunication Systems*. doi:10.1007/s11235-011-9440-5.
7. Candès, E. J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, *52*, 489–509.
8. Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, *52*, 1289–1306.
9. Parades, J., Arce, G. R., & Wang, Z. (2007). Ultra-wideband compressed sensing: Channel estimation. *IEEE Journal on Selected Topics in Signal Processing*, *1*, 383–395.
10. Başaran, M., Erküçük, S., & Çırpan, H.A. (2011). The effect of channel models on compressed sensing based UWB channel estimation. In *IEEE international conference on ultra-wideband (ICUWB)* (pp. 375–379).
11. Tipping, M.E., & Faul, A.C. (2003). Fast marginal likelihood maximization for sparse Bayesian models. In *Proceedings of 9th international workshop on artificial intelligence and statistics* (pp. 1–13).
12. Ji, S., Xue, Y., & Carin, L. (2008). Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, *56*(6), 2346–2355.
13. Babacan, S.D., Molina, R., & Katsaggelos, A.K. (2009). Fast Bayesian compressive sensing using Laplace priors. In *IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 2873–2876).
14. Yang, D., Li, H., & Peterson, G.D. (2011). Decentralized Turbo Bayesian compressed sensing with application to UWB systems. *EURASIP Journal on Advances in Signal Processing*, *2011*, article ID 817947.
15. Tang, L., Zhou, Z., & Shi, L. (2011). Ultra-wideband channel estimation based on distributed Bayesian compressive sensing. *International Journal of Digital Content Technology and its Applications (JDCTA)*, *5*(2), 1–8.
16. Shi, L., Zhou, Z., Tang, L., Yao, H., & Zhang, J. (2010). Ultra-wideband channel estimation based on Bayesian compressive sensing. In *Proceedings of 10th international symposium on communications and information technologies (ISCIT)* (pp. 779–782).
17. Özgör, M., Erküçük, S., & Çırpan, H.A. (2012). Bayesian compressive sensing for ultra-wideband channel models. In *35th international conference on telecommunications and signal processing (TSP)* (pp. 320–324).
18. Zayyani, H., Babaie-Zadeh, M., & Jutten, C. (2009). Compressed sensing block MAP-LMS adaptive filter for sparse channel estimation and a Bayesian Cramér-Rao bound. In *IEEE international workshop on machine learning and signal processing (MLSP)* (pp. 1–6).
19. Kay, S. M. (1993). *Fundamentals of statistical signal processing: Estimation theory*. Upper Saddle River, NJ: Prentice Hall.
20. Eldar, Y. C. (2006). Uniformly improving the Cramér-Rao bound and maximum likelihood estimation. *IEEE Transactions on Signal Processing*, *54*(8), 2943–2956.
21. Erküçük, S., Kim, D.I., & Kwak, K.S. (2007). Effects of channel models and Rake receiving process on UWB-IR system performance. In *IEEE international conference on communications (ICC)* (pp. 4896–4901).
22. Molisch, A. F., et al. (2006). A comprehensive standardized model for ultrawideband propagation channels. *IEEE Transactions on Antennas and Propagation*, *54*, 3151–3166.
23. Candès, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, *25*, 21–30.
24. Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: CRC Press.
25. Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd ed.). New York: Springer.
26. Van Tress, H. L. (1968). *Detection, estimation and modulation theory (Part I)*. New York: Wiley.
27. Baraniuk, R. G. (2007). Compressive sensing. *IEEE Signal Processing Magazine*, *24*(4), 118–121.

Mehmet Özgör received the B.Sc. and M.Sc. degrees with high honors in Electrical and Electronics Engineering from Karadeniz Technical University, Trabzon, Turkey in 2010 and in Telecommunication Engineering from İstanbul Technical University, Turkey in 2013, respectively. He is currently pursuing his Ph.D. degree in Telecommunication Engineering at İstanbul Technical University. From 2010 to 2012 he was a research and teaching assistant at Doğuş University, İstanbul, Turkey before joining İstanbul Kültür University, Turkey as a research and teaching assistant in January 2012. He has been working for P.I.Works Inc., a mobile network performance monitoring and optimization company, as a telecommunications engineer since June 2014. His primary research interests are self-organizing networks, mobile network optimization, compressive sensing and ultra-wideband communications.

Serhat Erküçük received the B.Sc. and M.Sc. degrees in Electrical Engineering from Middle East Technical University, Ankara, Turkey and from Ryerson University, Toronto, ON, Canada, in 2001 and 2003, respectively, and the Ph.D. degree in Engineering Science from Simon Fraser University, Burnaby, BC, Canada in 2007. He was an NSERC postdoctoral fellow at the University of British Columbia, Vancouver, BC, Canada until September 2008. Since then, he has been with Kadir Has University, Istanbul, Turkey, where he is currently an associate professor. His research interests are in physical layer design of emerging communication systems, wireless sensor networks and communication theory. Dr. Erküçük serves as an area editor for AEÜ – International Journal of Electronics and Communications.

Hakan Ali Çırpan received the B.S. degree in 1989 from Uludag University, Bursa, Turkey, the M.S. degree in 1992 from the University of Istanbul, Istanbul, Turkey, and the Ph.D. degree in 1997 from the Stevens Institute of Technology, Hoboken, NJ, USA, all in electrical engineering. From 1995 to 1997, he was a Research Assistant with the Stevens Institute of Technology, working on signal processing algorithms for wireless communication systems. In 1997, he joined the faculty of the Department of Electrical-Electronics Engineering at The University of Istanbul. In 2010 he has joined to the faculty of the department of Electronics & Communication Engineering at Istanbul Technical University. His general research interests cover wireless communications, statistical signal and array processing, system identification and estimation theory. His current research activities are focused on machine learning, signal processing and communication concepts with specific attention to channel estimation and equalization algorithms for future wireless systems.