



King's Research Portal

DOI:

[10.1186/s13638-016-0796-9](https://doi.org/10.1186/s13638-016-0796-9)

Document Version

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Al-Salihi, H. Q. K., & Nakhai, M. R. (2017). Efficient Bayesian Compressed Sensing-based Channel Estimation Techniques for Massive MIMO-OFDM Systems. *EURASIP Journal on Wireless Communications and Networking (EURASIP JWCN)*. <https://doi.org/10.1186/s13638-016-0796-9>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

RESEARCH

Open Access



Efficient Bayesian compressed sensing-based channel estimation techniques for massive MIMO-OFDM systems

Hayder AL-Salihi* and Mohammad Reza Nakhai

Abstract

Efficient and highly accurate channel state information (CSI) at the base station (BS) is essential to achieve the potential benefits of massive multiple input multiple output (MIMO) systems. However, the achievable accuracy that is attainable is limited in practice due to the problem of pilot contamination. It has recently been shown that compressed sensing (CS) techniques can address the pilot contamination problem. However, CS-based channel estimation requires prior knowledge of channel sparsity to achieve optimum performance, also the conventional CS techniques show poor recovery performance for low signal to noise ratio (SNR). To overcome these shortages, in this paper, an efficient channel estimation approach is proposed for massive MIMO systems using Bayesian compressed sensing (BCS) based on prior knowledge of statistical information regarding channel sparsity. Furthermore, by utilizing the common sparsity feature inherent in the massive MIMO system channel, we extend the proposed Bayesian algorithm to a multi-task (MT) version, so the developed MT-BCS can obtain better performance results than the single task version. Several computer simulation based experiments are performed to confirm that the proposed methods can reconstruct the original channel coefficient more effectively when compared to the conventional channel estimator in terms of estimation accuracy.

Keywords: Massive multiple input multiple output (MIMO), Channel estimation, Bayesian compressed sensing (BCS), Pilot contamination, Channel state information (CSI), Multi-task Bayesian compressed sensing (MTBCS)

1 Introduction

The main activity of recent research has identified that the major targets for the next generation of mobile communications, the so-called fifth generation of mobile communications, are to achieve 1000 times the system capacity and 10 times the spectral efficiency, energy efficiency and data rate, and 25 times the average cell throughput [1]. From a high-level perspective, there is a promising technology that enables reaching higher fifth generation targets, called a massive multiple input multiple output (MIMO). A massive MIMO can be defined as a system using a large number of antennas at the base station; accordingly, a significant beamforming can be achieved and the system capacity can serve a large number of users [2].

When comparing massive MIMO to the conventional MIMO systems, massive MIMO shows several advantageous aspects. Firstly, as the number of the antennas at the base station goes to high values, the simplest coherent combiner and linear precoder turn out to be optimal. Secondly, by exploiting the features of the channel reciprocity, additional antennas increase the network capacity significantly without the need for additional feedback overhead. Thirdly, enabling the power reduction in the uplink and in the downlink can provide the potential for small-cell size shrinking [3].

The major limiting factor in massive MIMO is the availability of accurate, instantaneous channel state information (CSI) at the base station. The CSI is typically acquired by transmitting predefined pilot signals and estimating the channel coefficients from the received signals by applying an appropriate estimation algorithm [1–3].

Channel estimation accuracy depends on having perfect orthogonal pilots allocated to the users; however,

*Correspondence: ana.al-salihi@kcl.ac.uk
King's College London, Department of Informatics, London, Strand, WC2R 2LS, London, UK

to achieve high spectral efficiency, the same carrier frequency should be used in the neighbouring cells by following a specific reuse pattern. This leads to the creation of a spatially correlated inter-cell interference, known as pilot contamination, which reduces the estimation performance and spectral efficiency [1–3].

The pilot contamination problem was analyzed in [4] and it has shown that the precoding downlink signal of the base station in the serving cell contaminated the received signal of the users roaming in other cells. The authors of [5] analyzed the pilot contamination problems in multi-cell massive MIMO systems relying on a large antennas at the base station, and demonstrated that the pilot contamination problem persisted in large-scale MIMO [6].

However, pilot contamination could be reduced by reducing the number of pilots. A multi-user scenario therefore needs to reduce the number of pilots without affecting the channel impulse response (CIR) quality. Hence, the development of efficient channel estimation techniques for massive MIMO that are computationally less complex and require a fewer number of pilots is a challenge that should be thoroughly addressed [7].

Recently, compressed sensing (CS) techniques have received attention since they can recover the unknown signals from only a small number of measurements, thus using significantly far fewer samples than is possible via the conventional Nyquist rate, which is the signal recovery scheme developed for CS to exploit the sparse nature of signals (that is, only a small number of components in a signal vector are non-zero). CS allows for accurate system parameter estimation with fewer pilots; thereby, addressing the pilot contamination problem and improving the bandwidth efficiency [8, 9]. However, classical CS algorithms require prior knowledge of channel sparsity, which is usually unknown in practical scenarios. In addition, to apply CS algorithms, the sampling matrix must satisfy the restricted isometry property (RIP) for guaranteeing reliable estimators. Such a condition cannot be easily verified because it results computational demanding [10, 11].

To overcome the scarcity of CS-based channel estimation in massive MIMO systems, in this paper, we propose an improved channel estimation scheme based on the theory of Bayesian CS (BCS) that introduces relevance vector machines (RVM) and statistical learning information (SLI) into standard CS; whereby, probabilistic a priori information regarding the channel sparsity can be exploited for more reliable channel recovery to mitigate the pilot contamination problem. Also, the sampling matrix condition is efficiently overcome based on probabilistic formulation [12–14].

Compared with the classical based scheme, our simulation results indicate that the proposed channel estimation methods provide improved estimation accuracy and can address the pilot contamination problem.

Furthermore, by exploiting the common statistical sparsity inherent in different multipath signals, we extend the BCS algorithms to a multi-task version for simultaneously reconstructing multiple signals, thus leading to MT-BCS [15, 16].

The main contributions of this paper are summarised as follows:

- The BCS-based channel estimation algorithm has been proposed for massive MIMO to address the pilot contamination problem.
- We have also proposed to enhance the performance of the BCS-based estimator through the principle of thresholding to select the most significant taps to improve the channel estimation accuracy.
- In addition, we have exploited the common statistical sparsity distribution to enhance the estimation accuracy performance through the proposed MT-BCS-based estimator.
- To provide the benchmark for the minimum performance error of the BSC and MT-BCS, the Cramer Rao bound (CRB) has been drawn for BCS and it has been derived and drawn for MT-BCS.

The remainder of this paper is organized as follows. The multi-cell massive MIMO system model is presented in Section 2. The BSC-based and the MT-BSC based channel estimation details are reviewed in Sections 3 and 4, respectively. In section 5, we provide the Cramer-Rao bound analysis. Section 6 presents the simulation results. Finally, the final conclusions are drawn in Section 7.

The following notation is adopted throughout the paper: \mathbb{C} denotes the complex number field. For $A \in \mathbb{C}$, we have $A = A_R + jA_I$, where $j = \sqrt{-1}$, while A_R and A_I are the real and imaginary parts of A , respectively. For any matrix \mathbf{A} , A_{ij} denotes the (i,j) th element. The transpose, inverse and Hermitian transpose operators are denoted by $(\cdot)^T$, $(\cdot)^{-1}$, and $(\cdot)^H$, respectively. Upper bold font are used to denote matrices while lower light font are used to denote vectors, lower and upper case represents the time domain and frequency domain, respectively. The \mathbf{I} denotes an identity matrix, $\text{diag}\{\mathbf{X}\}$ denotes the diagonal matrix with the diagonal entries equal to the elements of \mathbf{X} and \hat{X} represents the estimate of X . The Frobenius and spectral norms of a matrix \mathbf{x} are denoted by $\|\mathbf{x}\|_F$ and $\|\mathbf{x}\|_2$ respectively. $E\{\cdot\}$ has been employed to denote expectation with regard to all random variables within the brackets. A Gaussian stochastic variable \mathbf{o} is denoted by $\mathbf{o} \sim N(r, q)$, where r is the mean and q is the variance. Also, a random vector \mathbf{x} having the proper complex Gaussian distribution of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ is indicated by $\mathbf{x} \sim CN(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $N(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\det(\pi \boldsymbol{\Sigma})} e^{-(\mathbf{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$, for simplicity we refer to $CN(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ as $\mathbf{x} \sim CN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

2 Massive MIMO system model

We consider a time division duplexing (TDD) multi-cell massive MIMO system with C cells as shown in Fig. 1. Each cell comprises of M antennas at the BS and N single antenna users. To improve the spectral efficiency, orthogonal frequency division multiplexing (OFDM) is adopted [17, 18].

At the beginning of the transmission, all mobile stations in all cells synchronously transmit OFDM pilot symbols to their serving base stations. Let the OFDM pilot symbol of user n in the c -th cell be denoted by $\mathbf{x}_c^n = [X_c^n[1] X_c^n[2] \cdots X_c^n[K]]^T$, where K is the number of subcarriers. The OFDM transmission partition the multipath channel between the user and each antenna of the BS into K parallel independent additive white Gaussian noise (AWGN) sub-channels in the frequency domain. Each sub-channel is associated with a subcarrier. Let $H_{c^*,c,i}^n[k]$ denote the k -th sub-channel coefficient between the n -th user in the c -th cell and the i -th antenna of the BS of cell c^* in the uplink.

The received signal $Y_{c^*,i}$ by the i -th antenna element of the cell c^* at the k -th subcarrier can be expressed as

$$Y_{c^*,i}[k] = \sum_{n=1}^N H_{c^*,c^*,i}^n[k] X_{c^*}^n[k] + \sum_{c=1, c \neq c^*}^C \sum_{n=1}^N H_{c^*,c,i}^n[k] X_c^n[k] + V_{c^*,i}[k], \quad (1)$$

for all $1 \leq i \leq M$ and $1 \leq c \leq C$, where $V_{c^*,i}[k]$ is the AWGN at the i -th antenna of the BS in cell c^* at the k -th

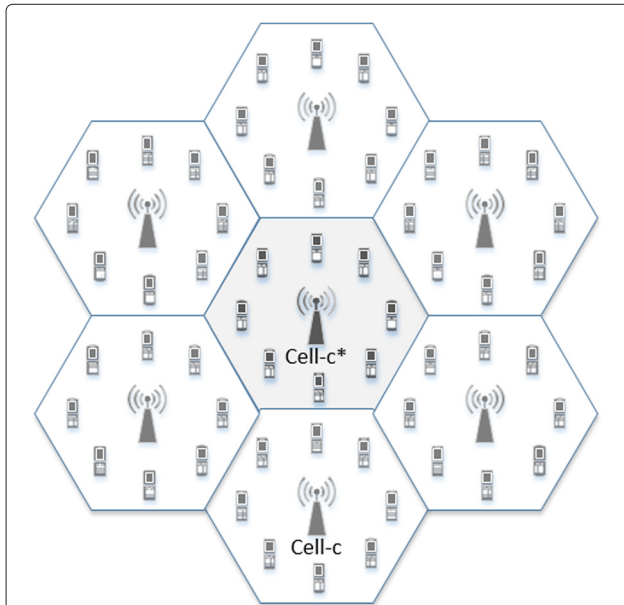


Fig. 1 Illustration of the system model of a multi-cell multi-user massive MIMO

subcarrier. Letting $\mathbf{y}_{c^*,i} = [Y_{c^*,i}[1] \cdots Y_{c^*,i}[K]]^T$, we can write (1) for all subcarriers at the i -th antenna of the BS in cell c^* in the compact form as

$$\mathbf{y}_{c^*,i} = \sum_{n=1}^N \mathbf{X}_{c^*}^n \mathbf{h}_{c^*,c^*,i}^n + \sum_{c=1, c \neq c^*}^C \sum_{n=1}^N \mathbf{X}_c^n \mathbf{h}_{c^*,c,i}^n + \mathbf{v}_{c^*,i}, \quad (2)$$

where $\mathbf{X}_{c^*}^n = \text{diag}\{\mathbf{x}_{c^*}^n\}$, $\mathbf{h}_{c^*,c,i}^n = [H_{c^*,c,i}^n[1] \cdots H_{c^*,c,i}^n[K]]^T$ and $\mathbf{v}_{c^*,i} = [V_{c^*,i}[1] \cdots V_{c^*,i}[K]]^T \sim CN(0, \sigma_v^2)$. Let $\mathbf{g}_{c^*,c,i}^n = [g_{c^*,c,i}^n[1] \cdots g_{c^*,c,i}^n[\ell] \cdots g_{c^*,c,i}^n[L]]^T$ collect the samples of the sampled multipath CIR between the n -th user of the c -th cell and the i -th antenna of the BS in cell c^* , where L is the number of the channel taps and $g_{c^*,c,i}^n[\ell]$ corresponds to the ℓ -th channel tap. The K frequency domain channel coefficients, i.e., $\mathbf{h}_{c^*,c,i}^n$, can be calculated as the K -point DFT of the CIR samples, i.e., $\mathbf{g}_{c^*,c,i}^n \in \mathbb{C}^{L \times 1}$, e.g., [18].

Hence,

$$\mathbf{h}_{c^*,c,i}^n = \mathbf{F} \mathbf{g}_{c^*,c,i}^n, \quad (3)$$

where $\mathbf{F} \in \mathbb{C}^{K \times K}$ represents the discrete Fourier transform (DFT) matrix, whose element in row s and column r is given by $[\frac{1}{\sqrt{K}} e^{-j2\pi s(K-r)(K-s)/K}]$, $1 \leq r \leq K$ and $1 \leq s \leq K$ and $\mathbf{g}_{c^*,c,i}^n \in \mathbb{C}^{K \times 1}$ is $\mathbf{g}_{c^*,c,i}^n \in \mathbb{C}^{L \times 1}$ augmented with $K - L$ zeros. Using (3) in (2), we get

$$\mathbf{y}_{c^*,i} = \sum_{n=1}^N \mathbf{X}_{c^*}^n \mathbf{F} \mathbf{g}_{c^*,c^*,i}^n + \sum_{c=1, c \neq c^*}^C \sum_{n=1}^N \mathbf{X}_c^n \mathbf{F} \mathbf{g}_{c^*,c,i}^n + \mathbf{v}_{c^*,i}. \quad (4)$$

The channel coefficient is modelled as $g_{c^*,c,i}^n[\ell] = \sqrt{\phi_{c^*,c,i}[\ell]} \psi_{c^*,c,i}[\ell]$ for $1 \leq \ell \leq L$, where $\phi_{c^*,c,i}$ model the path-loss and shadowing (large-scale fading), while the term $\psi_{c^*,c,i}$ is assumed to be independent identical distribution (i.i.d) of unknown random variables with $CN(0, 1)$ (small-scale fading) [3].

The received signal of (4) can be re-written as

$$\mathbf{y}_{c^*,i} = \sum_{n=1}^N \mathbf{X}_{c^*}^n \mathbf{F} \mathbf{g}_{c^*,c^*,i}^n + \mathbf{z}_{c^*,i}, \quad (5)$$

where the term $\mathbf{z}_{c^*,i} = \sum_{c=1, c \neq c^*}^C \sum_{n=1}^N \mathbf{X}_c^n \mathbf{F} \mathbf{g}_{c^*,c,i}^n + \mathbf{v}_{c^*,i}$ in (5) represents the net sum of inter-cell interference

plus the receiver noise, the variance interference σ_I^2 of the inter-cell interference term caused during pilot transmission can be expressed as

$$\sigma_I^2 = E \left\{ \left(\sum_{c=1, c \neq c^*}^C \sum_{n=1}^N \mathbf{X}_c^n \mathbf{F} \mathbf{g}_{c^*, c, i}^n \right) \times \left(\sum_{c=1, c \neq c^*}^C \sum_{n=1}^N \mathbf{X}_c^n \mathbf{F} \mathbf{g}_{c^*, c, i}^n \right)^H \right\}. \quad (6)$$

We define the measurement matrix $\mathbf{A}_{c^*}^n = \mathbf{X}_{c^*}^n \mathbf{F}$, then (5) can be rewritten as

$$\mathbf{y}_{c^*, i} = \sum_{n=1}^N \mathbf{A}_{c^*}^n \mathbf{g}_{c^*, c, i}^n + \mathbf{z}_{c^*, i}. \quad (7)$$

Based on the physical properties of outdoor electromagnetic propagation, the CIR in wireless communications usually contain a few significant channel taps as can be shown in Fig. 2, i.e., the CIR are sparse; hence, the number of non-zero taps of the channel is much smaller than the channel length, then the CS techniques can be applied for sparse channel estimation. This sparse property can be exploited to reduce the necessary channel parameters to be estimated. In this case, we can address the pilot contamination problem by using fewer pilots than the unknown channel coefficients [7, 19, 20].

3 BCS-based channel estimation

In common literature, channel estimation methods are classified into parametric and Bayesian approaches. A standard parametric approach is the best linear unbiased estimator, which is often referred to as least squares channel estimation. In contrast to parametric methods, the Bayesian approach treats the desired parameters as random variable with a-priori known statistics. Clearly, the a priori probability density function (PDF) of the channel is assumed to be perfectly known at the receiver

[21, 22]. Based on the Bayesian channel estimation philosophy, the estimation of unknown parameters is the expectation of the posterior probabilistic distribution that is proportional to the prior probability and the likelihood of the unknown parameters.

In this section, BCS-based channel estimation is presented in the context of massive MIMO channel estimation. Following the general procedure of BCS in [23] and [24], the full posterior distribution over unknown parameters of interest for the problem at hand can be given as

$$P(\mathbf{g}_{c^*, c^*, i}^n, \boldsymbol{\beta}, \sigma^2 | \mathbf{y}_{c^*, i}) = \frac{P(\mathbf{y}_{c^*, i} | \mathbf{g}_{c^*, c^*, i}^n, \boldsymbol{\beta}, \sigma^2) P(\mathbf{g}_{c^*, c^*, i}^n, \boldsymbol{\beta}, \sigma^2)}{P(\mathbf{y}_{c^*, i})}, \quad (8)$$

where $\boldsymbol{\beta}$ represents the hyperparameters that control the sparsity of the channel while σ^2 is the net sum of the noise variance and interference variance.

However, the probability of the observation vector, $P(\mathbf{y}_{c^*, i})$, is defined by the following equation

$$P(\mathbf{y}_{c^*, i}) = \int \int \int P(\mathbf{y}_{c^*, i} | \mathbf{g}_{c^*, c^*, i}^n, \sigma^2, \boldsymbol{\beta}) P(\mathbf{g}_{c^*, c^*, i}^n, \boldsymbol{\beta}, \sigma^2) d\mathbf{g}' d\boldsymbol{\beta} d\sigma^2, \quad (9)$$

cannot be computed analytically. So, the posterior distribution can be decomposed as

$$P(\mathbf{g}_{c^*, c^*, i}^n, \boldsymbol{\beta}, \sigma^2 | \mathbf{y}_{c^*, i}) \equiv P(\mathbf{g}_{c^*, c^*, i}^n | \mathbf{y}_{c^*, i}, \boldsymbol{\beta}, \sigma^2) P(\boldsymbol{\beta}, \sigma^2 | \mathbf{y}_{c^*, i}). \quad (10)$$

The first term of (10), $P(\mathbf{g}_{c^*, c^*, i}^n | \mathbf{y}_{c^*, i}, \boldsymbol{\beta}, \sigma^2)$, the posterior distribution over the channel coefficient can be expressed based on Bayes' rule as

$$P(\mathbf{g}_{c^*, c^*, i}^n | \mathbf{y}_{c^*, i}, \boldsymbol{\beta}, \sigma^2) = \frac{P(\mathbf{y}_{c^*, i} | \mathbf{g}_{c^*, c^*, i}^n, \sigma^2) P(\mathbf{g}_{c^*, c^*, i}^n | \boldsymbol{\beta})}{P(\mathbf{y}_{c^*, i} | \boldsymbol{\beta}, \sigma^2)}. \quad (11)$$

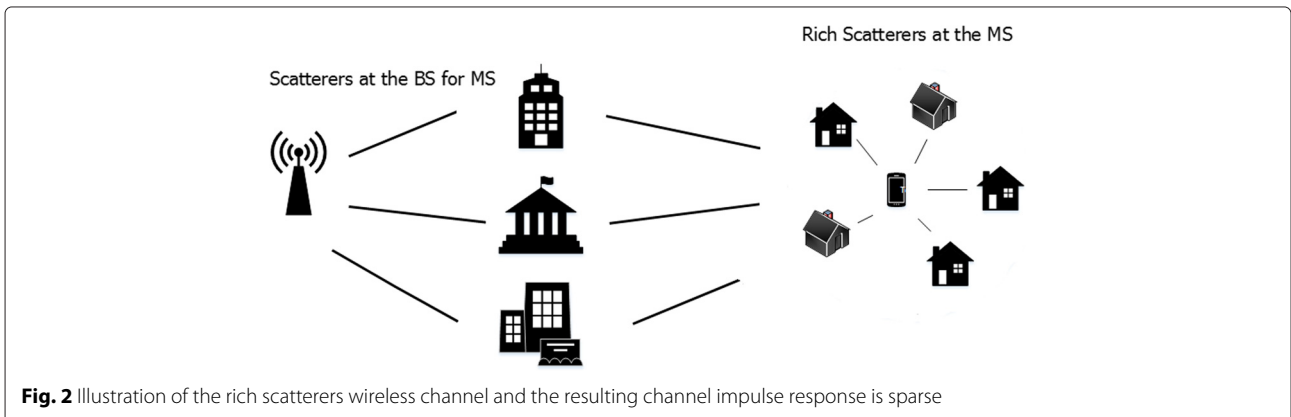


Fig. 2 Illustration of the rich scatterers wireless channel and the resulting channel impulse response is sparse

The posterior distribution given above is Gaussian distribution with mean $\boldsymbol{\mu}_{c^*,c^*,i}^n$ and the variance $\boldsymbol{\Sigma}_{c^*,c^*,i}^n$ are given by

$$\boldsymbol{\mu}_{c^*,c^*,i}^n = \sigma^{-2} \boldsymbol{\Sigma} \mathbf{A}_{c^*}^n \mathbf{y}_{c^*,i}, \quad (12)$$

$$\boldsymbol{\Sigma}_{c^*,c^*,i}^n = \left(\boldsymbol{\zeta} + \sigma^{-2} (\mathbf{A}_{c^*}^n)^H \mathbf{A}_{c^*}^n \right)^{-1}, \quad (13)$$

where $\boldsymbol{\zeta} = \text{diag}\{\beta_1, \beta_2, \dots, \beta_K\}$.

The estimated channel based on Bayesian estimation approaches to minimize the mean square error (MSE) is the expectation of $P(\mathbf{g}_{c^*,c^*,i}^n | \mathbf{y}_{c^*,i}, \boldsymbol{\beta}, \sigma^2)$, so the estimated channel can be expressed as

$$\hat{\mathbf{g}}_{c^*,c^*,i}^n = E(P(\mathbf{g}_{c^*,c^*,i}^n | \mathbf{y}_{c^*,i}, \boldsymbol{\beta}, \sigma^2)) = \boldsymbol{\mu}_{c^*,c^*,i}^n. \quad (14)$$

Now, to obtain the estimated channel $\hat{\mathbf{g}}_{c^*,c^*,i}^n$, we need to find the hyperparameter σ^2 and $\boldsymbol{\beta}$ that can be obtained from the second term on the right-hand side of (10) by applying a type-II maximum likelihood procedure by operating a RVM.

Based on Bayes' theorem, the posterior distribution $P(\boldsymbol{\beta}, \sigma^2 | \mathbf{y}_{c^*,i})$ is proportional $P(\mathbf{y}_{c^*,i} | \boldsymbol{\beta}, \sigma^2)$ [23]. Then, the type-II maximum likelihood is applied to the log marginal likelihood as follows

$$P(\mathbf{y}_{c^*,i} | \boldsymbol{\beta}, \sigma^2) = \int_{-\infty}^{\infty} P(\mathbf{y}_{c^*,i} | \mathbf{g}_{c^*,c^*,i}^n, \sigma^2) P(\mathbf{g}_{c^*,c^*,i}^n | \boldsymbol{\beta}) d\mathbf{g}'. \quad (15)$$

Based on the assumption of the RVM approach in [23], the term $P(\mathbf{g}_{c^*,c^*,i}^n | \boldsymbol{\beta})$ follows zero-mean Gaussian distribution and can be expressed as

$$P(\mathbf{g}_{c^*,c^*,i}^n | \boldsymbol{\beta}) = (2\pi)^{-\frac{K}{2}} \prod_{i=1}^K \beta_k^{\frac{1}{2}} \exp \left[-\frac{1}{2} \mathbf{g}_{c^*,c^*,i}^n \boldsymbol{\beta}_k (\mathbf{g}_{c^*,c^*,i}^n)^H \right], \quad (16)$$

while the Gaussian likelihood function of $\mathbf{y}_{c^*,i}$ according to the probability theory, can be written as

$$P(\mathbf{y}_{c^*,i} | \mathbf{g}_{c^*,c^*,i}^n, \sigma^2) = \left(\frac{2\pi}{\sigma^2} \right)^{-\frac{K}{2}} \exp \left(-\frac{\sigma^2}{2} \|\mathbf{y}_{c^*,i} - \mathbf{A}_{c^*}^n \mathbf{g}_{c^*,c^*,i}^n\|_2^2 \right). \quad (17)$$

By substituting (16) and (17) into (15), marginal likelihood $P(\mathbf{y}_{c^*,i} | \boldsymbol{\beta}, \sigma^2)$ can be expressed as

$$P(\mathbf{y}_{c^*,i} | \boldsymbol{\beta}, \sigma^2) = \log \left\{ \left(\frac{\beta_k}{2\pi} \right)^{\frac{K}{2}} \left(\frac{1}{2\pi} \right)^{\frac{K}{2}} \prod_{k=1}^K \beta_k^{\frac{1}{2}} \int_{-\infty}^{\infty} \exp \left(-\frac{\beta_k}{2} \|\mathbf{y}_{c^*,i} - \mathbf{A}_{c^*}^n \mathbf{g}_{c^*,c^*,i}^n\|_2^2 \right) + \frac{1}{2} (\mathbf{g}_{c^*,c^*,i}^n)^H \boldsymbol{\beta}_k \mathbf{g}_{c^*,c^*,i}^n \right\}, \quad (18)$$

$\boldsymbol{\beta}$ can be obtained by differentiating the log marginal likelihood with regard to σ^2 , and equating it to zero and it can be given as

$$(\beta_k)^{ii} = \frac{I - \beta_k (\boldsymbol{\Sigma}_{c^*,c^*,i}^n)_k}{(\boldsymbol{\mu}_{c^*,c^*,i}^n)_k^2}. \quad (19)$$

While σ^2 is obtained by differentiating (19) with regard to $\boldsymbol{\beta}$ and set these derivations to zero and can be expressed as

$$(\sigma^2)^{ii} = \frac{\|\mathbf{y}_{c^*,i} - \mathbf{A}_{c^*}^n \mathbf{g}_{c^*,c^*,i}^n\|_2^2}{(M - I + \sum_{k=1}^K \beta_k)}. \quad (20)$$

The β_k and σ_k^2 which maximize the log marginal likelihood are then found iteratively by setting $\boldsymbol{\beta}$ and σ^2 to initial values and then finding values for $\boldsymbol{\mu}_{c^*,c^*,i}^n$ and $\boldsymbol{\Sigma}_{c^*,c^*,i}^n$ from (12) and (13). These values are then repeatedly used to calculate a new estimate for β_k and σ^2 and until a convergence criteria is met.

Further details of the BCS algorithm can be found in [23, 24]. The procedure for implementation of the proposed technique is summarized in Algorithm 1.

Algorithm 1 BCS-Based Channel Estimation Algorithm

INPUTS:

- 1) Pilot signal \mathbf{x}^n .
- 2) observation matrix $\mathbf{A}_{c^*}^n = \mathbf{x}_{c^*}^n \mathbf{F}$.

Initial Configuration:

- 1: Select a suitable value for convergence δ .
- 2: $ii \leftarrow 1$
- 3: Select a start value for σ^2 and $\boldsymbol{\beta}$.

- 4: Compute $\boldsymbol{\Sigma}_{c^*,c^*,i}^n = \left(\boldsymbol{\zeta} + \boldsymbol{\beta} (\mathbf{A}_{c^*}^n)^H \mathbf{A}_{c^*}^n \right)^{-1}$.

- 5: Compute $\boldsymbol{\mu}_{c^*,c^*,i}^n = \boldsymbol{\beta} \boldsymbol{\Sigma}_{c^*,c^*,i}^n \mathbf{A}_{c^*}^n \mathbf{y}_{c^*,i}$.

repeat

- 6: Update $(\beta_k)^{ii} = \frac{I - \beta_k (\boldsymbol{\Sigma}_{c^*,c^*,i}^n)_k}{(\boldsymbol{\mu}_{c^*,c^*,i}^n)_k^2}$ and $(\sigma^2)^{ii} = \frac{\|\mathbf{y}_{c^*,i} - \mathbf{A}_{c^*}^n \mathbf{g}_{c^*,c^*,i}^n\|_2^2}{(M - I + \sum_{k=1}^K \beta_k)}$.

- 7: **Until** $\delta > \sum_{ii=1} (\sigma^2)^{ii+1} - (\sigma^2)^{ii}$.

- 8: Compute $\hat{\mathbf{g}}_{c^*,c^*,i}^n = \boldsymbol{\mu}_{c^*,c^*,i}^n = E(P(\mathbf{g}_{c^*,c^*,i}^n | \mathbf{y}_{c^*,i}, \boldsymbol{\beta}, \sigma^2))$.

OUTPUTS: Return the Estimated Channel $\hat{\mathbf{g}}_{c^*,c^*,i}^n$.

In contrast to the conventional BCS-based estimator, it can also improve the performance of the BCS estimator based on the principle of thresholding, which can be applied to keep the most significant taps. The proposed algorithm applies a threshold approach by retaining the channel taps that have energy above a threshold value of ϱ and set the other taps to zero. The value of ϱ is the energy of the channel impulse response.

4 Multi-task BCS based channel estimation

With a high probability of user movements, the massive MIMO system channel may vary. Consequently, the channels at different time instants/locations are different but share the same common statistical property. As a result, to estimate the current channel, we can exploit the previous compressive vectors in addition to the current compressive vector [15].

Given the system model in II, the received signals of (7) can have the following formulation

$$\mathbf{y}_{c^*,i,j} = \sum_{n=1}^N \mathbf{A}_{c^*,j}^n \mathbf{g}_{c^*,c^*,i,j}^n + \mathbf{z}_{c^*,i,j}, \quad (21)$$

for $j = 1, 2, \dots, J$ where J is the number of the task, $\mathbf{A}_{c^*,j}^n$, $\mathbf{g}_{c^*,c^*,i,j}^n$ and $\mathbf{z}_{c^*,i,j}$ represents the j th measurement matrices, channel vector and the noise vector, respectively [15].

The main target is to estimate the channel $\mathbf{g}_{c^*,c^*,i,j}^n$ which can be computed based on Bayesian channel estimation philosophy as the mean of the channel posterior distribution that can be represented as

$$\hat{\mathbf{g}}_{c^*,c^*,i,j}^n = E(P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{y}_{c^*,i,j}, \mathbf{\Xi}_j, \xi_0)), \quad (22)$$

where ξ_0 represents the inverse of the net sum of the noise variance and interference variance, while $\mathbf{\Xi}_j$ represent the hyperparameters that control the sparsity of the channel. Based on Bayes' rule the posterior distribution can be given as

$$\begin{aligned} & P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{y}_{c^*,i,j}, \mathbf{\Xi}_j, \xi_0) \\ &= \frac{P(\mathbf{y}_{c^*,i,j} | \mathbf{g}_{c^*,c^*,i,j}^n, \xi_0) P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{\Xi}_j)}{\int P(\mathbf{y}_{c^*,i,j} | \mathbf{g}_{c^*,c^*,i,j}^n, \xi_0) P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{\Xi}_j) d\mathbf{g}^n} \\ & \sim N(\boldsymbol{\mu}_{c^*,i,j}^n, \boldsymbol{\Sigma}_{c^*,i,j}^n), \end{aligned} \quad (23)$$

the mean and covariance can be given by

$$\boldsymbol{\mu}_{c^*,i,j}^n = \xi_0 \boldsymbol{\Sigma}_{c^*,i,j}^n \mathbf{A}_{c^*,j}^n \mathbf{y}_{c^*,i,j}, \quad (24)$$

$$\boldsymbol{\Sigma}_{c^*,i,j}^n = (\boldsymbol{\Psi} + \mathbf{\Xi}_j (\mathbf{A}_{c^*,j}^n)^H \mathbf{A}_{c^*,j}^n)^{-1}, \quad (25)$$

where $\boldsymbol{\Psi} = \text{diag}(\psi_0, \psi_1, \psi_2, \dots, \psi_K)$.

The likelihood function for the parameter $\mathbf{g}_{c^*,c^*,i,j}^n$ and ξ_0 based on the received signal $\mathbf{y}_{c^*,i,j}$ and can be expressed as

$$\begin{aligned} P(\mathbf{y}_{c^*,i,j} | \mathbf{g}_{c^*,c^*,i,j}^n, \xi_0) &= \left(\frac{2\pi}{\xi_0} \right)^{\frac{-N}{2}} \\ & \exp\left(\frac{-\xi_0}{2} \|\mathbf{y}_{c^*,i,j} - \mathbf{A}_{c^*,j}^n \mathbf{g}_{c^*,c^*,i,j}^n\|_2^2 \right). \end{aligned} \quad (26)$$

The channel coefficients $\mathbf{g}_{c^*,c^*,i,j}^n$ are assumed to be drawn from a product of zero-mean Gaussian distributions that are shared by all tasks as follow

$$\begin{aligned} P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{\Xi}_j) &= \prod_{i=1}^N (P(\mathbf{g}_{c^*,c^*,i,j}^n | 0, \mathbf{\Xi}_j^{-1})) \\ &= (2\pi)^{\frac{-N}{2}} \prod_{i=1}^N \mathbf{\Xi}_j^{\frac{1}{2}} \\ & \times \exp\left[\frac{-1}{2} (\mathbf{g}_{c^*,c^*,i,j}^n)^H \mathbf{\Xi}_j \mathbf{g}_{c^*,c^*,i,j}^n \right]. \end{aligned} \quad (27)$$

To obtain the estimated channel, we need to estimate $\mathbf{\Xi}_j$ and ξ_0 by applying the same procedure in Section 3 to the posterior distribution $P(\mathbf{y}_{c^*,i,j} | \mathbf{\Xi}_j, \xi_0)$ that can be inference as [16]

$$\begin{aligned} P(\mathbf{y}_{c^*,i,j} | \mathbf{\Xi}_j, \xi_0) &\equiv P(\mathbf{y}_{c^*,i,j} | \mathbf{g}_{c^*,c^*,i,j}^n, \xi_0) \\ & P(\mathbf{g}_{c^*,c^*,i,j}^n | \mathbf{\Xi}_j). \end{aligned} \quad (28)$$

Now, by maximizing the log marginal likelihood and then differentiating with respect to $\mathbf{\Xi}_j$ and ξ_0 and setting to zero yields

$$(\mathbf{\Xi}_j)^{new} = \frac{J - \mathbf{\Xi}_j \sum_{j=1}^J \boldsymbol{\Sigma}_{c^*,c^*,i,j}^n}{\sum_{j=1}^J (\boldsymbol{\mu}_{c^*,c^*,i,j}^n)^2}, \quad (29)$$

$$(\xi_0)^{new} = \frac{\sum_{j=1}^J (K - J + \sum_{i=1}^J \boldsymbol{\Sigma}_{c^*,c^*,i,j}^n \mathbf{\Xi}_j)}{\sum_{j=1}^J \|\mathbf{y}_{c^*,i,j} - \mathbf{A}_{c^*,j}^n \mathbf{g}_{c^*,c^*,i,j}^n\|_2^2}. \quad (30)$$

Further information on MT-BCS can be found in [16].

5 CRB for BCS-based estimator

In this section, we analyse the CRB for the proposed BCS and MT-BCS based channel estimation techniques to provide a benchmark for the minimum estimation error that can be achieved by the proposed algorithm. The CRB on the covariance of any estimator $\hat{\boldsymbol{\theta}}$ can be given as

$$E\left\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^H\right\} \geq \mathbf{J}^{-1}(\boldsymbol{\theta}), \quad (31)$$

where $J(\theta)$ is the Fisher information matrix (FIM) corresponding to the observation f , and can be given as

$$J(\theta) = E \left(\frac{\partial}{\partial \theta} \log l(\theta, f) \right) \left(\frac{\partial}{\partial \theta} \log l(\theta, f) \right)^T, \quad (32)$$

where $l(\theta, f)$ is the likelihood function corresponding to the observation f , parameterized by θ [25].

Therefore, given the system model in 2, the closed form expression of the Bayesian CRB (BCRB) for the proposed BCS can be given as

$$J(\mathbf{g}_{c^*,c^*,i}^n) \geq \left(\frac{1}{\beta} + \frac{\mathbf{A}_{c^*,c^*}^n (\mathbf{A}_{c^*,c^*}^n)^H}{\sigma^2} \right)^{-1}. \quad (33)$$

Theorem 1 Given (28), the closed form expression of the BCRB for the proposed MT-BCS can be given as

$$J(\mathbf{g}_{c^*,c^*,ij}^n) \geq \left(\frac{1}{\xi_j} + \frac{\mathbf{A}_{c^*,j}^n (\mathbf{A}_{c^*,j}^n)^H}{\xi_0} \right)^{-1}. \quad (34)$$

Proof See Appendix 1. □

6 Simulation results

To verify the accuracy of our analytical results, the simulation parameters can be summarized as follows: the number of antennas is 100, the number of users is 100, the number of the channel taps is 500, the number of subcarrier K is 4096 and the convergence δ is 10^{-6} . The simulation results are obtained by averaging over 1000 realizations.

To compare the accuracy of the channel estimation techniques, the normalized (MSE) is used for performance evaluation and is computed as

$$MSE = \frac{\|\hat{\mathbf{g}}_{c^*,c^*,ij}^n - \mathbf{g}_{c^*,c^*,ij}^n\|_2^2}{\|\mathbf{g}_{c^*,c^*,ij}^n\|_2^2}. \quad (35)$$

Figure 3 shows the MSE performance comparison among a BCS-based channel estimation of three scenarios under small pilot contamination ($\phi_{c^*,c^*,i} = 1$ and $\phi_{c^*,c,i} = 0.1$), strong pilot contamination ($\phi_{c^*,c^*,i} = 1$ and $\phi_{c^*,c,i} = 0.5$), very strong pilot contamination ($\phi_{c^*,c^*,i} = 1$ and $\phi_{c^*,c,i} = 0.9$), regularized least square (RLS)-based estimator with no pilot contamination as a benchmark and the BCRB for BCS as a reference line. The results have shown significant improvement in estimation accuracy and addressing the pilot contamination problem for SNR values of -40 to 40 dB for the proposed technique compared with R-LS. This is a result of exploiting the prior statistical of channel sparsity. Furthermore, the results still show enhanced estimation performance for high SNR.

Figure 4 shows the (MSE) performance versus SNR with a different value of setting to the number of subcar-

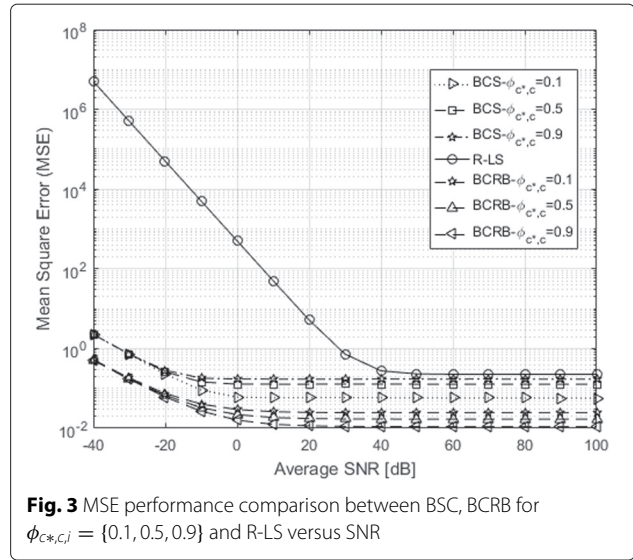


Fig. 3 MSE performance comparison between BCS, BCRB for $\phi_{c^*,c,i} = \{0.1, 0.5, 0.9\}$ and R-LS versus SNR

rier $K = \{100, 200$ and $300\}$, so the compression ratio (CR) (i.e., L/K) is to be $CR = \{0.2, 0.1$ and $0.06\}$, while the experiment is run under small pilot contamination ($\phi_{c^*,c^*,i} = 1$ and $\phi_{c^*,c,i} = 0.1$). The results prove that the estimation accuracy is better performed by decreasing the values of the number of subcarriers, accordingly with increasing CR.

Figure 5 demonstrates the MSE of the BSC-based channel estimation versus SNR for three scenarios of different settings to the number of antennas at the base station $M = \{100, 200$, and $300\}$, the system under strong pilot contamination ($\phi_{c^*,c^*,i} = 1$ and $\phi_{c^*,c,i} = 0.7$). The results show that the estimation accuracy of the proposed algorithm is enhanced by increasing the number of antennas.

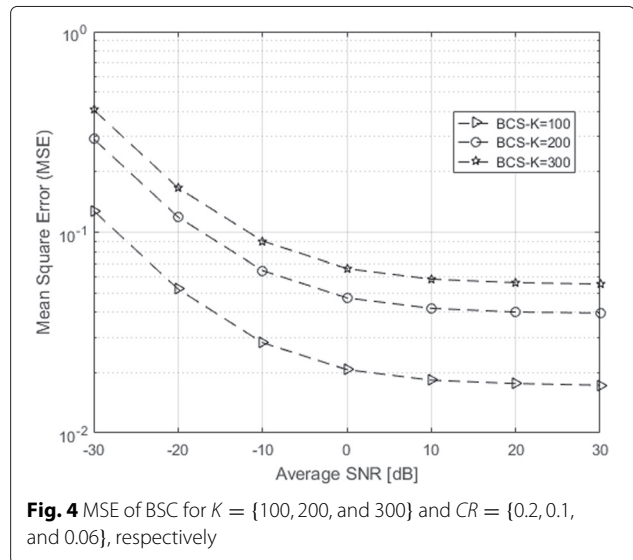
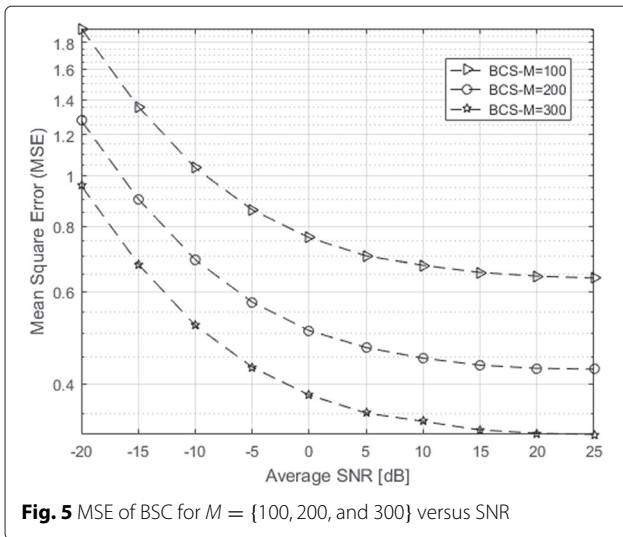


Fig. 4 MSE of BSC for $K = \{100, 200$, and $300\}$ and $CR = \{0.2, 0.1$, and $0.06\}$, respectively



Thus, according to the law of large numbers, more coordinated BS antennas could provide more accurate support estimation.

Figure 6 shows the (MSE) performance versus SNR for BCS with different values for the number of pilots: 1000, 500, 100, 50, and 25, where the number of subcarrier K is 4096. The number of the CIR path is 500 while the experiments run under strong pilot contamination. For cases of the number of the pilots is greater than the number of channel taps (i.e., 1000 and 500), the BCS provides inefficient estimation accuracy, while for the other cases of the number of the pilot of (100, 50, and 25), which is less than 500, the estimation accuracy is enhanced significantly. In addition, there is no significant improvement for the cases of the number of the pilots 100, 50, and 25. In these cases, we can address pilot contamination by employing small values for the number of the pilot, i.e., 25.

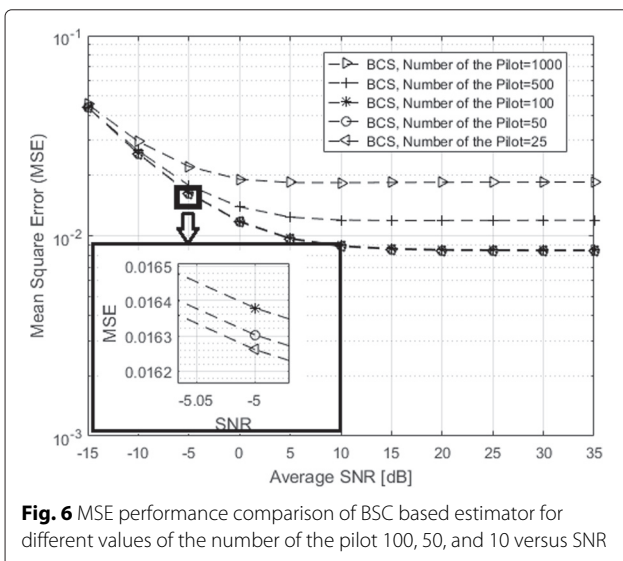


Figure 7 compares the (MSE) performance versus SNR among BCS, threshold-BSC, MT-BCS, LS, OMP and the Bilinear Approximate Message Passing (Bi-AMP) [26]. The number of subcarrier K is 1024 and the number of the CIR path is 100. Results show the proposed MT-BCS enjoys significant performance improvement over all the other estimators as a result of exploiting the statistical prior information on a large scale. However, this advantage is at the expense of a relatively high complexity of BCS and MT-BCS over other estimators as depicted in Table 1, which compares the computational complexity Bi-AMP [26], BCS [23], OMP [27], LS [28], and the MT-BCS [16]. Also, the results showed that the thresholding approach enhances the estimation accuracy of the conventional BCS, as the CIR contains so many taps with no significant energy. By setting the threshold and neglecting these taps, a huge part of the noise and interference from pilot contamination will be eliminated.

7 Conclusions

To address the pilot contamination problem in massive MIMO systems, we proposed a BCS-based channel estimation algorithm for the multi-cell multi-user massive MIMO. The simulation results have revealed that the BCS-based channel estimation algorithm has tremendous improvement over conventional-based channel estimation algorithms and can address the pilot contamination problem. Furthermore, the proposed technique can be enhanced by thresholding the CIR to a certain value and also by exploiting the common sparsity feature inherent in the system channel. In addition, the number of antennas and the compression ratio should be selected wisely to achieve optimum estimation accuracy.

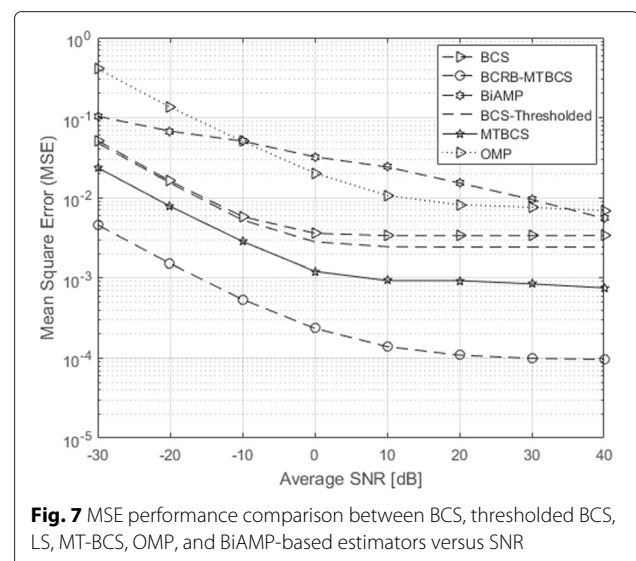


Table 1 Complexity analysis

Estimators	Computation complexity
R-LS	$O(L^2)$
BCS	$O(KL^2)$
MTBCS	$O(KL^3)$
OMP	$O(L \log(K))$
BiAMP	$O(LK + K)$

Appendix 1: Proof of Theorem 1

Following Section 5, we can write the FIM as

$$J(\mathbf{y}_{c^*,i,j}) \geq -E \left(\frac{\partial^2 \log(P_{\mathbf{y}_{c^*,i,j}|\Xi_j, \xi_0}(P(\mathbf{y}_{c^*,i,j}|\Xi_j, \xi_0)))}{\partial^2 \mathbf{g}'} \right)^{-1} \quad (36)$$

Based on Bayes' rule in (32), the FIM can be decomposed into two terms

$$\begin{aligned} & -E \left(\frac{\partial^2 \log(P_{\mathbf{y}_{c^*,i,j}|\Xi_j, \xi_0}(P(\mathbf{y}_{c^*,i,j}|\Xi_j, \xi_0)))}{\partial^2 \mathbf{g}'} \right) = - \\ & E \left(\frac{\partial^2 \log(P_{\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0}(P(\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0)))}{\partial^2 \mathbf{g}'} \right) - \\ & E \left(\frac{\partial^2 \log(P_{\mathbf{g}_{c^*,c^*,i,j}^n|\Xi_j}(P(\mathbf{g}_{c^*,c^*,i,j}^n|\Xi_j)))}{\partial^2 \mathbf{g}'} \right), \end{aligned} \quad (37)$$

using (28), the first term can be computed as follow

$$\begin{aligned} & \frac{\partial^2 \log(P_{\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0}(P(\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0)))}{\partial^2 \mathbf{g}'} = \\ & \frac{\partial}{\partial \mathbf{g}'} \left[-\log(2\pi)^{\frac{1}{2}} \xi_0^{-1} - \frac{\xi_0}{2} \|\mathbf{y}_{c^*,i,j} - \mathbf{g}_{c^*,c^*,i,j}^n \mathbf{A}_{c^*,j}^n\|_2^2 \right], \end{aligned} \quad (38)$$

$$\frac{\partial^2 \log(P_{\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0}(P(\mathbf{y}_{c^*,i,j}|\mathbf{g}_{c^*,c^*,i,j}^n, \xi_0)))}{\partial^2 \mathbf{g}'} = \frac{\mathbf{A}_{c^*,j}^n (\mathbf{A}_{c^*,j}^n)^H}{\xi_0}. \quad (39)$$

By applying the same procedure in (38 and 39) to the second term of (37) gives

$$E \left(\frac{\partial^2 \log(P_{\mathbf{g}_{c^*,c^*,i,j}^n|\Xi_j}(P(\mathbf{g}_{c^*,c^*,i,j}^n|\Xi_j)))}{\partial^2 \mathbf{g}'} \right) = (\Xi_j)^{-1}. \quad (40)$$

Acknowledgements

This work is supported by the Iraqi Higher Committee of Educational Development (HCED). The authors would like to acknowledge its financial support.

Competing interests

The authors declare that they have no competing interests.

Received: 3 October 2016 Accepted: 12 December 2016

Published online: 23 February 2017

References

1. C-X Wang, et al, Cellular architecture and key technologies for 5G wireless communication Networks. *IEEE Comm. Mag.* **52**(2), 122–130 (2014)
2. V Jungnickel, K Manolakis, W Zirwas, B Panzner, V Braun, M Lossow, M Sternad, R Apelfrojd, T Svensson, The role of small cells, coordinated multipoint, and massive MIMO in 5G. *IEEE Commun. Mag.* **52**(5), 44–51 (2014)
3. H Zhang, S Gao, D Li, H Chen, L Yang, On superimposed pilot for channel estimation in multi-cell multiuser MIMO uplink: large system analysis. *IEEE Trans. Vehicular Technol.* **65**, 99 (2015)
4. J Jose, A Ashikhmin, TL Marzetta, et al, Pilot contamination problem in multi-cell TDD systems. *Proc. IEEE Int.Symp. Inf. Theory, Seoul.* **28**, 2184–2188 (2009)
5. J Jose, A Ashikhmin, TL Marzetta, et al, Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans. Wireless Commun.* **10**(8), 2640–2651 (2011)
6. J Zhang, B Zhang, S Chen, X Mu, M El-Hajjar, L Hanzo, Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems. *IEEE J. Selected Topics Signal Process.* **8**(5), 759–772 (2014)
7. M Masood, L Afify, TY Al-Naffouri, Efficient coordinated recovery of sparse channels in massive MIMO. *IEEE Trans. Signal Process.* **63**(1), 104–118 (2015)
8. N Sinh. Compressive sensing for multi-channel and large scale MIMO networks. PhD. thesis, Dept. of Elect. And Comp. Eng., Concordia Univ., (Montreal, 2014), p. 13
9. C Qi, G Yue, L Wu, Y Huang, A Nallanathan, Pilot design schemes for sparse channel estimation in OFDM systems. *IEEE Trans. Veh. Technol.* **64**(4), 1493–1505 (2015)
10. M Carlin, P Rocca, G Oliveri, F Viani, A Massa, Directions-of-arrival estimation through Bayesian compressive sensing strategies. *IEEE Trans. Antennas Propagat.* **61**(7), 3828–3838 (2013)
11. RG Baraniuk, Compressive sampling. *IEEE Signal Process. Mag.* **24**(4), 118–124 (2007)
12. Z Fan, Z Lu, Y Han, in *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. Accurate channel estimation based on Bayesian compressive sensing for next-generation wireless broadcasting systems, (Beijing, 2014), pp. 1–5
13. Z Fan, Z Lu, Y Han, *Reliable channel estimation based on Bayesian compressive sensing for TDS-OFDM systems*, (Macau, 2014), pp. 620–624
14. X Cheng, YL Guan, G Yue, S Li, in *Proc. IEEE GLOBECOM*. Enhanced Bayesian compressive sensing for ultra-wideband channel estimation, (California, 2012), pp. 4065–4070
15. X Cheng, M Wang, Y Guan, Ultra wideband channel estimation: a Bayesian compressive sensing strategy based on statistical sparsity. *IEEE Trans. Vehicular Technol.* **64**(5), 1819–1832 (2015)
16. S Ji, D Dunson, L Carin, Multi-task compressive sensing. *IEEE Trans. Signal Process.* **57**(1), 92–106 (2009)
17. X Guo, S Chen, J Zhang, X Mu, L Hanzo, Optimal pilot design for pilot contamination elimination/reduction in large-scale multiple-antenna aided OFDM systems. *IEEE Trans. Wireless Commun.* **15**(11), 7229–7243 (2016)
18. MR Nakhai, Multicarrier transmission. *IET Signal Process.* **2**(1), 114 (2008)
19. W Ding, F Yang, W Dai, J Song, Time–frequency joint sparse channel estimation for MIMO-OFDM systems. *IEEE Commun. Lett.* **19**(1), 58–61 (2015)
20. C Qi, G Yue, L Wu, Y Huang, A Nallanathan, Pilot design schemes for sparse channel estimation in OFDM systems. *IEEE Trans. Veh. Technol.* **64**(4), 1493–1505 (2015)
21. A Scherb, K Kammeyer, in *Proc. IEEE Workshop Smart Antennas*. Bayesian channel estimation for doubly correlated MIMO systems, (Vienna, 2007)
22. SM Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. (PTR Prentice-Hall, Englewood Cliffs, 1993)
23. S Ji, Y Xue, L Carin, Bayesian compressive sensing. *IEEE Trans. Signal Process.* **56**(6), 2346–2356 (2008)
24. ME Tipping, Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **1**, 211–244 (2001)
25. AK Jagannatham, BD Rao, Whitening-rotation-based semi-blind MIMO channel estimation. *IEEE Trans. Signal Process.* **54**(3), 861–869 (2006)

26. JT Parker, P Schniter, V Cevher, Bilinear generalized approximate message passing. *IEEE Trans. Signal Process.* **62**(22), 5839–5853 (2014)
27. JA Tropp, AC Gilbert, Signal recovery from partial information via orthogonal matching pursuit. *IEEE Trans. Inf. Theory.* **53**(12), 4655–4666 (2007)
28. OO Oyerinde, Reweighted regularised variable step size normalised least mean square-based iterative channel estimation for multicarrier-interleave division multiple access systems. *IET Signal Process.* **10**(8), 947–954 (2016)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
