EFFICIENT COLLABORATIVE SPARSE CHANNEL ESTIMATION IN MASSIVE MIMO

Mudassir Masood[⋆], *Laila H. Afify*[⋆], *and Tareq Y. Al-Naffouri*[⋆∗]

[⋆]King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[∗]King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia

ABSTRACT

We propose a method for estimation of sparse frequency selective channels within MIMO-OFDM systems. These channels are independently sparse and share a common support. The method estimates the impulse response for each channel observed by the antennas at the receiver. Estimation is performed in a coordinated manner by sharing minimal information among neighboring antennas to achieve results better than many contemporary methods. Simulations demonstrate the superior performance of the proposed method.

Index Terms— massive MIMO, OFDM, sparse channel

1. INTRODUCTION

Most wireless channels can be modeled as discrete multipath channels with large delay spread and few significant paths. This implies sparsity of channel impulse response (CIR) [1–3]. This leads from the fact that scatterers are sparsely distributed in space. Thus, it is essentially beneficial to account for such a sparse channel model when performing channel estimation. We aim to use this property in the context of MIMO-OFDM systems. The deployment of multiple antennas, offers key advantages to wireless systems performance in terms of power gains, channel robustness, diversity etc. [4]. Specifically, the use of very large antenna arrays has very recently emerged. Such systems, known as *massive* MIMO [5],[6], have also the potential to scale down the transmission power because of the use of small low-power active antennas. Thus, large antenna arrays can play a key role in exploiting the true potential of traditional MIMO systems.

In large-scale MIMO the major performance bottleneck is the availability of CIR. Several algorithms exist that take advantage of the sparsity and the assumption that channel support does not vary as we move across the antenna grid, however with some drawbacks. For example, the algorithms assume common support throughout antenna array which is not true for large arrays. The readers are directed to [7–14] for some work on MIMO and massive MIMO channel estimation. In this work, we utilize the property of loosely space-invariant channel support along with the sparsity property to propose an efficient pilot-aided Bayesian approach to estimate sparse CIR in the massive-MIMO setup. In this approach each receiving antenna collaborates with its direct neighbors to estimate its unknown sparse channel. The neighboring antennas share their knowledge of most significant taps (MST) to reach a consensus about the CIR support.

This paper is organized as follows. In Section II, we present the system model and formulate the problem. In Section III we introduce a simple Bayesian approach for channel estimation which leads us to present the proposed coordinated channel recovery algorithm in Section IV. Simulation results are discussed in Section V and Section VI concludes the paper. A detailed version of this paper is also available [15].

This work was supported jointly by King Abdullah University of Science and Technology and the Deanship of Scientific Research (DSR) at King Fahd University of Petroleum & Minerals through project No. RG-1314.

2. SYSTEM MODEL AND PROBLEM FORMULATION

2.1. Preliminaries

We consider a MIMO-OFDM system, in which the base station (BS) is equipped with a large two-dimensional antenna array consisting of $R = M \times G$ antennas distributed across M rows and G columns.[1] OFDM is adopted as the signaling mechanism. In an OFDM system, serially incoming bits are divided into N parallel streams and mapped to a Q-ary QAM alphabet $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_Q\}$. This results in an N-dimensional data vector denoted by $\boldsymbol{\mathcal{X}} = [\mathcal{X}(1), \mathcal{X}(2), \ldots, \mathcal{X}(N)]^{\mathsf{T}}$. The equivalent time-domain signal $\mathbf{x} = \mathbf{F}^{\mathsf{H}}\boldsymbol{\mathcal{X}}$ is transmitted. Here $\mathbf{F}$ is an $N \times N$ unitary DFT matrix whose (c, d)th entry is $f_{c,d} = \frac{1}{\sqrt{N}} \exp\left(-j\frac{2\pi}{N}cd\right)$, and N is the number of subcarriers.

2.2. Channel Model

The channel through which the transmitted signal $\mathbf{x}$ is received at the receive antenna $r = (m, g)$ (where $m \in \{1, 2, \ldots, M\}$ and $g \in \{1, 2, \ldots, G\}$) as shown in Fig. 1 is denoted by $\mathbf{h}_r \in \mathbb{C}^L$. We shall assume that $\mathbf{h}_r$ has a sparse structure and is modeled as $\mathbf{h}_r = \mathbf{h}_A \circ \mathbf{h}_B$ where $\circ$ indicates element-by-element multiplication. The vector $\mathbf{h}_A$ consists of elements that are drawn from some unknown distribution and $\mathbf{h}_B$ is a Bernoulli random vector where its ith element has an active probability of $p(\mathbf{h}_B(i) = 1) = \lambda_i$.

Therefore, the entries of $\mathbf{h}_B$ form a collection of iid Bernoulli random variables. Thus, $\mathbf{h}_r$ is an L-tap discrete-time sparse channel where no assumption whatsoever is made about the distribution of its non-zero complex-valued coefficients.[2] Moreover, depending upon factors such as antenna separation and transmission bandwidth, the MST locations of $\mathbf{h}_r$'s in the array may or may not vary. The array for which the $\mathbf{h}_r$'s have common support are termed space-invariant arrays (SIA) while the arrays for which this is not true are called space-varying arrays (SVA).

The received signal at the rth antenna is best described in the frequency domain and is given by

$$\boldsymbol{\mathcal{Y}}_r = \mathrm{diag}(\boldsymbol{\mathcal{X}})\boldsymbol{\mathcal{H}}_r + \boldsymbol{\mathcal{W}}_r, \tag{1}$$

where $\boldsymbol{\mathcal{Y}}_r$ is the Fourier transform of the received vector, $\boldsymbol{\mathcal{W}}_r \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2\mathbf{I})$ is the frequency-domain noise vector and diag is an

[1]Depending on the value of M and G, the antennas could have a linear or a planar configuration. Further, we would like to stress that we confine our attention to regular configurations out of convenience as our approach applies to any two-dimensional or even three-dimensional configuration of antennas.

[2]The coefficients could be derived iid or non-iid from a Gaussian or non-Gaussian distribution. The implementation in this paper is agnostic to the distribution of channel coefficients.

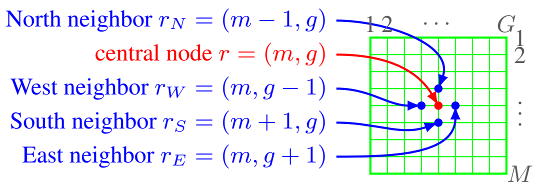


Fig. 1: Center antenna r and its 4-neighbors in 2-D antenna grid.

operator that produces a diagonal matrix by spreading the elements of $\boldsymbol{\mathcal{X}}$ along the diagonal. Moreover, $\boldsymbol{\mathcal{H}}^r = \mathbf{F}[\mathbf{h}_r^\top \; \mathbf{0}_{1\times N-L}]^\top = \underline{\mathbf{F}}\mathbf{h}_r$ is the $N \times 1$ channel frequency response vector where $\underline{\mathbf{F}}$ is the truncated Fourier matrix of size $N \times L$ formed by selecting the first L columns of $\mathbf{F}$. Finally, we can rewrite (1) as $\boldsymbol{\mathcal{Y}}_r = \mathbf{A}\mathbf{h}_r + \boldsymbol{\mathcal{W}}_r$, where $\mathbf{A} \triangleq \text{diag}(\boldsymbol{\mathcal{X}})\underline{\mathbf{F}}$ is an $N \times L$ matrix.

2.3. Problem Formulation

Let the transmit antenna sends pilots in K subcarriers and the remaining $N - K$ subcarriers are used for data transmission. Let $\mathcal{P}$ represents the set of indices of the K subcarriers over which pilots are transmitted. Thus,

$$\boldsymbol{\mathcal{Y}}_r(\mathcal{P}) = \mathbf{A}(\mathcal{P})\mathbf{h}_r + \boldsymbol{\mathcal{W}}_r(\mathcal{P}) \tag{2}$$

where $\boldsymbol{\mathcal{Y}}_r(\mathcal{P})$ and $\boldsymbol{\mathcal{W}}_r(\mathcal{P})$ are formed, respectively, by selecting entries of $\boldsymbol{\mathcal{Y}}_r$ and $\boldsymbol{\mathcal{W}}_r$ indexed by $\mathcal{P}$. Similarly, $\mathbf{A}(\mathcal{P})$ is a $K \times L$ matrix formed by selecting the rows of $\mathbf{A}$ indexed by $\mathcal{P}$. We aim to solve for $\mathbf{h}_r$ in equation (2). This obviously requires that $K \geq L$. Since the channel delay spread (equivalently L) is usually large, this requires a large number of subcarriers to be reserved for pilots, severely affecting the spectral efficiency of the system. However, by virtue of channels being sparse with large delay spread, we could actually solve for $\mathbf{h}_r$ if $K < L$ as suggested by the compressed sensing theory [16, 17]. We consider a random placement of pilot tones $\mathcal{P}$ over the OFDM subcarriers as it has been found to be optimal for sparse channel estimation [18, 19]. The aforementioned system model will be used in subsequent sections to develop our coordinated approach for estimation of all R channels $\mathbf{h}_r$.

3. SPARSITY-AWARE DISTRIBUTION AGNOSTIC BAYESIAN CHANNEL ESTIMATION

Consider the model presented in (2). For notational convenience, we will drop the symbols r and $\mathcal{P}$ unless required for clarity. Hence,

$$\boldsymbol{\mathcal{Y}} = \mathbf{A}\mathbf{h} + \boldsymbol{\mathcal{W}}, \tag{3}$$

where we are interested in performing Bayesian estimation of the wireless CIR $\mathbf{h}$. Bayesian approaches assume a prior distribution, however given the dynamic nature of wireless channels it is usually impossible to characterize the distribution. Even if the distribution is known it is very difficult or even impossible to estimate the distribution parameters (e.g., mean and variance for Gaussian) especially when the channel statistics are not i.i.d. In that respect, the use of distribution agnostic Bayesian sparse signal recovery method (SABMP) [20] is quite attractive which provides Bayesian estimates even when the prior is non-Gaussian or unknown.

A naive way would be to use SABMP to perform sparse channel recovery at each antenna element in the array. The channels would be estimated independently and the receivers will not take advantage of the additional information of common support. Unlike that we propose a coordinated channel estimation method in Sec. 4 which utilizes the common support information. However, before doing so, we introduce in the following some necessary modifications to the SABMP algorithm presented in [20].

3.1. SABMP for non-iid Bernoulli random vector

The development of the SABMP algorithm assumes that elements of $\mathbf{h}$ are activated with equal probability λ (iid Bernoulli). However, if some elements are more probable than others, it is desirable to assign those elements a higher probability. This requires us to assume a non-iid Bernoulli behavior for $\mathbf{h}$. Thus if $\mathcal{S}$ contains the indices of the active elements of $\mathbf{h}$ (i.e., the support of $\mathbf{h}$), the probability of that support is given by, $p(\mathcal{S}) = \prod_{i\in\mathcal{S}} \lambda_i \prod_{j\in\{1,\ldots,L\}\setminus\mathcal{S}}(1-\lambda_j)$ where λ_i is the active probability of index i. Using this $p(\mathcal{S})$ results in a modified version of the dominant support selection metric of [20] (see eq. (13) therein). The new metric is,

$$\nu(\mathcal{S}) \triangleq -\frac{1}{2\sigma_{\mathbf{n}}^2}\left\|\mathbf{P}_{\mathcal{S}}^{\perp}\boldsymbol{\mathcal{Y}}\right\|_2^2 + \sum_{i\in\mathcal{S}} \ln\lambda_i + \sum_{j\notin\mathcal{S}} \ln(1-\lambda_j) \tag{4}$$

For future reference, let us call the algorithm taking advantage of this new dominant support selection metric `RS1`.

3.2. All possible combinations of support

Let the set of dominant taps of $\mathbf{h}$ as detected by SABMP be $T = \{\alpha_1, \alpha_2, \cdots, \alpha_{T_{max}}\}$. The SABMP algorithm computes the approximate MMSE estimate of $\mathbf{h}$ for all support sets $\mathcal{S}$ in the set of dominant supports sets $\mathcal{S}^d$. According to the definition provided in SABMP [20], $\mathcal{S}^d = \{\{\alpha_1\}, \{\alpha_1, \alpha_2\}, \cdots, \{\alpha_1, \alpha_2, \cdots, \alpha_{T_{max}}\}\}$ which is a set of cardinality T_{max}. However, due to the reasons that will be mentioned in Section 3.3, we modified the SABMP algorithm to generate all possible $2^{T_{max}} - 1$ combinations of dominant locations in the set T to form as many support sets $\mathcal{S}_v$. Therefore, the new set $\mathcal{S}^d$ to compute $p(\mathcal{S}_v|\boldsymbol{\mathcal{Y}})$ and $\mathbb{E}[\mathbf{h}|\boldsymbol{\mathcal{Y}}, \mathcal{S}_v]$ is,

$$\mathcal{S}^d = \{\{\mathcal{S}_v\}, v = 1, 2, \cdots, V\}, \qquad V = 2^{T_{max}} - 1. \tag{5}$$

3.3. Marginalization

The above-mentioned modified algorithm provides all different combinations of the elements obtained in the set of dominant taps together with their marginalized posterior probabilities. In this section, we aim to use this information to marginalize out each dominant tap and compute $p(\alpha_t), \forall t \in \{1, 2, \cdots, T_{\max}\}$. This could be done by considering the posterior probabilities of only those combinations of support which involve purely the detected dominant taps. This is because only these supports will have significant posteriors and the remaining cases could be safely ignored (see Fig. 7 in [20]). Set $\mathcal{S}^d$ described in the last section corresponds to this set of supports. Thus the marginal probability of tap α_t is computed as follows,

$$p(\alpha_t) = \sum_{\mathcal{S}\in\mathcal{S}^d,\, \alpha_t\cap\mathcal{S}\neq\varnothing} p(\mathcal{S}|\boldsymbol{\mathcal{Y}}). \tag{6}$$

Note that the marginal probabilities of only the MSTs are computed in this way. The other elements that are found to be inactive could be forced to have either very low or even zero probability. The resulting new SABMP algorithm which incorporates the modifications presented in Sections 3.1 - 3.3 is capable of handling non-iid Bernoulli

Algorithm 1 Channel Estimation using Pilots - Posterior-based

1. Run RS2 at each antenna
2. Each antenna, receives marginals from its neighbors
3. Each antenna computes average marginal probabilities
4. Repeat steps 2-3 above, D times
5. Use RS1 to re-estimate all channels using new probabilities

distributed random vectors and also returns the marginal probabilities of detected MSTs. Let us call this algorithm RS2 for reference.

4. ITERATIVE COORDINATED CHANNEL RECOVERY

We will now describe the proposed channel estimation method. In this method the receiver antennas coordinate with each other to determine the MSTs and consequently estimate the channels. A naive approach would be to formulate the problem either in the form of an MMV or a block-sparse recovery problem and utilize available algorithms such as [21,22]. However, in the proposed approach, in order to reduce the communication overhead, the antennas collaborate in a stagewise manner. Basically, each receiver element r and only its immediate 4-neighbors $\mathcal{N} = \{r_N, r_S, r_E, r_W\}$ as shown in Fig. 1 communicate with each other.[3] This process is repeated which effectively diffuses the information present at each antenna to distant antennas. In this manner the collaboration is performed to estimate channels accurately. In the discussion that follows, we present three algorithms for CIR estimation that take advantage of collaboration.

4.1. Algorithm 1: Channel Estimation using Pilots

We seek to solve the problem mentioned in (2) by using the pilot observations. This algorithm starts by estimating the sparse channels $\mathbf{h}_r$ at each antenna element r using the RS2 algorithm. We initialize the algorithm by assuming that all taps of $\mathbf{h}_r$ have equal active probability λ_{init} throughout the array. Therefore, $p\left(\mathbf{h}_{\text{B}}(l) = 1\right) = \lambda_l = \lambda_{\text{init}}, \forall l \in \{1, 2, \cdots, L\}$.

Let $T^r = \{\alpha_1^r, \alpha_2^r, \cdots, \alpha_{T_{\max}}^r\}$ be the set of active taps of channel $\mathbf{h}_r$ as detected by RS2. Note that since λ_{init} is same throughout the array, the number of detected active taps $T_{\max}$ will also be equal for all the receivers i.e., the cardinality $|T^r| = T_{\max}, \forall r$. The RS2 algorithm also returns the marginal probabilities $p(\alpha_t^r), t \in \{1, 2, \cdots, T_{\max}\}$. Each antenna r, acting as central antenna, collects these probabilities from its 4-neighbors and computes the average for each tap $\alpha_i, i \in \{1, 2, \cdots, L\}$ as follows

$$p(\alpha_i) = \begin{cases} \sum\limits_{j \in \mathcal{N}^+} p(\alpha_i^j) \Big/ |\mathcal{N}^+|, & \text{if } \alpha_i \in \bigcup\limits_{j \in \mathcal{N}^+} T^j \\ p_{\text{small}}, & \text{otherwise} \end{cases}, \qquad (7)$$

where $\mathcal{N}^+ = \mathcal{N} \cup r$ and p_{small} is an arbitrarily small value assigned to the taps which have not been detected by any of the neighbors and the central antenna. This averaging step is repeated D times by each antenna where the value of D depends on whether the array under consideration is SIA or SVA. In the SIA case, since the MST locations do not vary across the array, contribution from as many antennas as possible will strengthen our belief in these locations. Therefore, we may select $D = \max(M, G)$ which equals to the largest dimension of the antenna array which ensures that each antenna receives information from every other antenna in the array. On the other hand, for SVA, the array configuration and other parameters should be taken into consideration to decide a proper value of D. Specifically, according to lemma 1 in [23] if observations from q antennas are used to recover n-sparse channel vectors using K pilots then for a unique solution $n \leq \lceil (K+q)/2 \rceil - 1$ holds which simplifies to the condition on D as $D > \sqrt{n - \frac{K}{2} - \frac{1}{4}} - \frac{1}{2}$. Here $\lceil \cdot \rceil$ denotes the ceiling operation. Finally, each antenna uses the newly computed probabilities as new initial probabilities with the RS1 algorithm to get refined sparse channel estimates. The algorithm is summarized in Algorithm 1.

We would like to point out that sharing posteriors puts a high communication load on the massive-MIMO system because it needs floating point numbers to be communicated. The communication cost could be reduced significantly by sharing just the integers. We therefore, propose a variant which uses integers for communication among receiver antennas. As explained next, this algorithm has an additional advantage of low computational complexity.

4.2. Algorithm 2: Low Communication/Computational Cost

In this algorithm we do not calculate marginal probabilities. The algorithm starts by estimating channel at each receiver using the original SABMP algorithm. At each receiver, the algorithm assigns scores to the channel tap locations based on the detected amplitudes. Since there are $T_{\max}$ possible channel taps detected by SABMP, the algorithm starts by assigning a maximum score of $T_{\max}$ to the tap location with highest absolute amplitude, moves downward until a score of 1 is assigned to the tap location with the least amplitude among the top $T_{\max}$ taps. All other tap locations are assigned a score of zero. Each antenna, acting like a central antenna, collects the $T_{\max}$ scores from each neighbor and finds an average score $\psi(\alpha_i)$ for each tap α_i in a fashion similar to that in (7). Finally, after repeating the process D times, a *belief* measure $b(\alpha_i) = \psi(\alpha_i)/T_{\max}$ is computed to be used by the RS1 algorithm. $b(\alpha_i)$ is the estimated belief that the ith tap is active. The beliefs $b(\alpha)$ are used in place of the marginal probabilities to re-estimate the channels following a strategy similar to that explained in Algorithm 1. This belief-based algorithm reduces the communication cost significantly because we totally avoid communicating floating point numbers. Moreover, since the algorithm does not compute marginal probabilities it has lower computational complexity. We now move on to suggest another level of refinement for the posteriors/scores by selecting reliable data carriers to perform channel estimation.

4.3. Algorithm 3: Using Reliable Carriers

The estimated channels from previous sections could be used to perform equalization and recover the transmitted data. A significant improvement in channel estimation could be achieved by incorporating this available information of user data. However, it is a challenging task to determine which data carriers are reliable enough to be considered. In this respect, we seek to assign a reliability measure $\mathfrak{R}(i),\ i \in \{1, \cdots, N\} \backslash \mathcal{P}$ to each of the $N - |\mathcal{P}|$ data carriers. For this purpose, we use the reliability measure suggested in [24] to compute carrier reliabilities. The reliability values are then sorted and the carriers corresponding to the top U values of $\mathfrak{R}$ are considered in calculations. Let $\mathcal{R}^r$ contains the indices of the top U reliable carriers for receiver r. Collaboration among receiver antennas could be performed to further strengthen the belief in the reliable carriers. In order to do so, each antenna r_C, acting as central antenna, collects the indices of the reliable carriers from its 4-neighbors $\mathcal{N}$ and

[3] For the elements lying at the edges of the array the number of neighbors are different. We use $\mathcal{N}$ to denote the set of neighbors irrespective of the position of r and therefore $2 \leq |\mathcal{N}| \leq 4$.

Algorithm 2 Channel Estimation using $\mathcal{R}^{\star} \cup \mathcal{P}$

1. Run `Algorithm 1` or `Algorithm 2`
2. Each antenna r uses its estimated channel to find top U reliable carriers $\mathcal{R}^r$ and sends $\mathcal{R}^r$ to its central antenna
3. Each antenna finds the intersection of received reliable carriers $\mathcal{R} = \bigcap_{r\in\{r_C, r_N, r_S, r_E, r_W\}} \mathcal{R}^r$ and sends it back to its neighbors
4. Each antenna sends data on $\mathcal{R}$ back to its central antenna
5. Each antenna further refines the reliable carriers by selecting only those with same data. Call this list $\mathcal{R}^{\star}$.
6. Each antenna uses $\mathcal{R}^{\star} \cup \mathcal{P}$ to perform SABMP recovery

selects only those which are common to all antennas under consideration $\mathcal{R} = \bigcap_{r\in\{r_C \cup \mathcal{N}\}} \mathcal{R}^r$, where $\mathcal{R}^r$ are the indices of reliable carriers of antenna r. $\mathcal{R}$ is then transmitted to the neighbors which then send back the corresponding data. Further refinement is done by retaining only those carriers which carry same data. Let us represent these carriers by $\mathcal{R}^{\star}$. The central antenna uses this final list of reliable carriers plus the pilots i.e., $\mathcal{R}^{\star} \cup \mathcal{P}$ to solve,

$$\boldsymbol{\mathcal{Y}}_r(\mathcal{R}^{\star} \cup \mathcal{P}) = \mathbf{A}(\mathcal{R}^{\star} \cup \mathcal{P})\mathbf{h}_r + \boldsymbol{\mathcal{W}}_r(\mathcal{R}^{\star} \cup \mathcal{P}) \qquad (8)$$

and estimate channel $\mathbf{h}_r$. Thus the pilots and reliable carriers are used together to reach at better estimates of channels which is evident from the simulation results presented in Section 5. The resulting algorithm is presented in Algorithm 2. Note that the proposed algorithms are independent of the antenna grid topology as the only information required by an antenna is that of its neighbors.

5. SIMULATION RESULTS

We simulated a MIMO-OFDM system with a 10×10 receive antenna grid. Moreover, the number of sub-carriers used is $N = 256$ and the pilot carriers $\mathcal{P}$ are chosen randomly. We use 4-QAM modulation and the Gaussian noise statistics are adjusted according to the desired SNR. Channels of sparsity 3 and varying length L are generated using IlmProp channel modeling tool [25, 26]. All results were averaged over 100 trials. We conducted three different experiments.

5.1. Methods for Performance Comparisons

The channel vectors $\mathbf{h}_r$ are estimated using the *a*) LS with known true MST locations (oracle-LS), *b*) block-sparse recovery method (BR), *c*) proposed Posterior-based channel estimation using pilots (PB-P), *d*) proposed Integer-based channel estimation using pilots (IB-P), and *e*) proposed Posterior- or Integer-based channel estimation using pilots and reliable carriers (PB-R / IB-R),

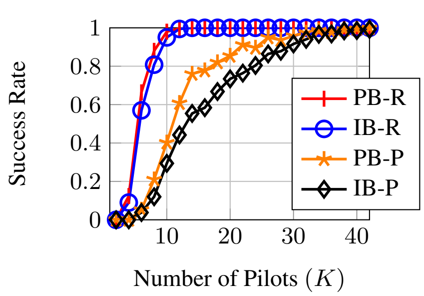
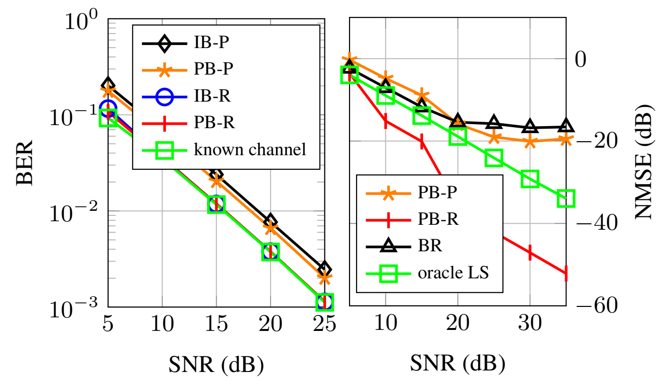


Fig. 2: How many pilots are needed to successfully recover the CIR?

Oracle-LS and BR are used for benchmark purpose. Oracle-LS knows the channel support at each antenna and hence the only burden is tap estimation using the available pilots. The block sparse recovery method (BR) works in the space-invariant case and uses the fact that the channel support is the same across the array. It casts the problem as several block sparse problems and use the block sparse Bayesian learning algorithm (BSBL) [21] to solve them.



(a) PB vs IB (b) PB vs BR and oracle-LS

Fig. 3: Performance comparison of algorithms.

5.2. Experiment 1 - How many pilots?

In this experiment, we are interested in finding the required number of pilots for successful recovery of channels of length $L = 64$. The graph in Fig. 2 shows the channel recovery success rate vs varying number of pilots. Note that both pilot-based and data-aided versions of PB and IB algorithms were simulated. Here, success rate is defined as the ratio of the number of successful trials to the number of total trials, where a trial was considered successful when NMSE < -10 dB. SNR was fixed at 10 dB and the number of pilots varied from 2 to 42. It is evident from the figure that just 6 pilots are needed by both PB-R and IB-R to achieve a success rate $> 50\%$ and only 12 pilots to achieve a 100% success rate. This is a small fraction of the channel length $L = 64$ (i.e., 9.37% and 18.75% respectively).

5.3. Experiment 2 - Comparison between PB and IB

In this experiment, we compare the performance of the proposed PB and IB channel estimation algorithms. Channels of length $L = 64$ were estimated using 16 pilots. Fig. 3a shows the BER of the recovered data using the estimated CIRs. The figure shows that incorporating reliable carriers results in significant performance gains.

5.4. Experiment 3 - Comparison with BR and oracle-LS

In this experiment, we benchmark the performance of the proposed algorithms against BR and oracle-LS. Here the signal is passed through a channel of length $L = 32$. Moreover, we use $K = 8$ pilots. Fig. 3b shows that the proposed PB-R algorithm has the best performance among all algorithms.

6. CONCLUSION

A channel estimation procedure in the massive-MIMO setup which is agnostic to the distribution of channel taps was proposed. It uses a modified version of SABMP to exploit the sparse common support property and share information in a stagewise manner to perform channel recovery. The approach results in lower communication and computational complexity. Simulation results show the superiority over other methods.

7. REFERENCES

[1] W. Bajwa, A. Sayeed, and R. Nowak, "Sparse multipath channels: Modeling and estimation," in *Proc. Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE)*, 2009, pp. 320–325.

[2] H. Minn and V. Bhargava, "An investigation into time-domain approach for OFDM channel estimation," *IEEE Trans. Broadcast.*, pp. 240–248, 2000.

[3] S.-S. Sadough, M. Ichir, P. Duhamel, and E. Jaffrot, "Wavelet-based semiblind channel estimation for ultrawideband OFDM systems," *IEEE Trans. Veh. Technol.*, pp. 1302–1314, 2009.

[4] H. Holma and A. Toskala, *LTE Advanced: 3GPP Solution for IMT-Advanced*. Wiley, 2012.

[5] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, pp. 40–60, Jan. 2013.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 160–171, Feb. 2013.

[7] T. Y. Al-Naffouri and A. A. Quadeer, "A forward-backward Kalman filter-based STBC MIMO OFDM receiver," *EURASIP J. Adv. Signal Process*, vol. 2008, pp. 203:1–203:14, Jan. 2008. [Online]. Available: http://dx.doi.org/10.1155/2008/158037

[8] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 4, Nov. 2004, pp. 2487–2492.

[9] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 815–830, Oct 2014.

[10] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[11] H. Q. Ngo, T. Marzetta, and E. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, May 2011, pp. 3464–3467.

[12] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[13] S. L. H. Nguyen, A. Ghrayeb, and M. Hasna, "Iterative compressive estimation and decoding for network-channel-coded two-way relay sparse ISI channels," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 1992–1995, Dec. 2012.

[14] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, "Semi-blind channel identification and equalization in OFDM: an expectation-maximization approach," in *2002 IEEE 56th Vehicular Technology Conference. Proceedings. VTC 2002-Fall.*, vol. 1, pp. 13–17.

[15] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 104–118, Jan 2015.

[16] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, Apr. 2006.

[17] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[18] C. Qi and L. Wu, "Optimized pilot placement for sparse channel estimation in OFDM systems," *IEEE Signal Process. Lett.*, vol. 18, no. 12, pp. 749–752, Dec. 2011.

[19] G. Taubock and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, Apr. 2008, pp. 2885–2888.

[20] M. Masood and T. Y. Al-Naffouri, "Sparse reconstruction using distribution agnostic Bayesian matching pursuit," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5298–5309, Nov. 2013.

[21] Z. Zhang and B. Rao, "Recovery of block sparse signals using the framework of BSBL," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Process. (ICASSP)*, 2012, pp. 3345–3348.

[22] M. Masood and T. Y. Al-Naffouri, "Support agnostic Bayesian matching pursuit for block sparse signals," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 4643–4647.

[23] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, July 2005.

[24] E. Al-Safadi and T. Y. Al-Naffouri, "Pilotless recovery of clipped OFDM signals by compressive sensing over reliable data carriers," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2012, pp. 580–584.

[25] "IlmProp," http://www2.tu-ilmenau.de/nt/en/ilmprop//, [Online; accessed January 10, 2014].

[26] G. Del Galdo, M. Haardt, and C. Schneider, "Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements," *Advances in Radio Science - Kleinheubacher Berichte*, vol. 50, no. 5, pp. 117–126, October 2003.