CrossMark

Channel estimation and channel tracking for correlated block-fading channels in massive MIMO systems

Suresh Dahiya [*], Arun Kumar Singh [**]

Department of Electrical Engineering, Indian Institute of Technology, Jodhpur, Rajasthan, 342011, India

ABSTRACT

This paper presents a channel estimation and tracking method for correlated block-fading channels in massive MIMO wireless cellular systems. In order to conserve resources, the proposed algorithm requires the uplink pilot signal only once, at the start of communication. By utilizing the temporal correlation between consecutive Resource Blocks (RBs) and the error correction capability of turbo codes, the channel matrix in subsequent RBs is estimated at the Base Station (BS) itself using the uplink data of current the RB and the estimated channel matrix of previous the RB. Compared to existing blind estimation methods, the proposed method places fewer limitations on the system settings such as the number of BS antennas, the number of users, and the number of coherent channel usage compared to existing blind estimation methods. Simulation results show that the proposed algorithm provides better performance for a moderate RB size, a high-order of QAM scheme, and a smaller ratio of the number of BS antennas and mobile terminals (N/K). For a reasonably small N/K (order of 10), the proposed scheme achieves a lower symbol error probability than the conventional pilot-based estimation approach.

1. Introduction

Massive Multiple Input Multiple Output (massive MIMO) is a promising candidate for 5th-generation wireless mobile communication systems [1–3]. Challenges such as antenna spacing constraints, channel estimation, and limited coherence block-size still require better solutions before this technology comes into deployable form. Massive MIMO systems described in the literature are considered in conjunction with linear processors, such as Match Filtering (MF) and Zero Forcing (ZF) for decoding and precoding the multiple data streams for different users. Instantaneous Channel State Information (ICSI) is the prime requirement for the aforementioned linear processors. Moreover, the accuracy of ICSI is a key performance factor in high SNR and high spectral efficiency regimes [4]. However, the ICSI changes with time and frequency; thus, the communication takes place in blocks that have finite durations and frequency bands in which the ICSI is almost constant. These blocks are termed as RBs. Along with transmission of data, the ICSI must be estimated in each RB. To bring the ICSI overheads within feasible limits, the massive MIMO cellular system proposed in the literature works in Time Division Duplex (TDD) mode. The conventionally proposed system uses Pilot Based Estimation (PBE) to estimate ICSI; this estimation is performed by a BS using the pilot signal transmitted from Mobile Terminals

(MTs). The estimated ICSI is then used for decoding the uplink data and precoding the downlink data [3,5].

Although feasible, PBE has a severe impact on system performance owing to the following factors: [6] First, a fraction of the RB is consumed by pilot training which significantly lowers the spectral efficiency in scenarios with smaller coherence time. Second, the pilots must be reused across the cells during the training phase, which creates interference called pilot contamination. Research on pilot-free massive MIMO systems is already in progress. A subspace projection-based blind pilot decontamination approach reported by Muller et al. [7] utilizes the bulk decomposition of the singular values of channel matrix under certain system parameter conditions. However, their performance analysis was presented for a large ratio (60) of the number of BS antennas to mobile terminals (N/K). Another method of blind channel estimation reported by Ngo et al. [8] utilizes eigenvalue decomposition (EVD) in conjunction with Iterative Least Square Projection (ILSP) [9]. The EVD-based method of [8] relies on a large number of receiving antennas and a large coherence block size. The ILSP gives insignificant improvement after second iteration, owing to the high correlation between the estimates of subsequent consecutive iterations. Thus, the scheme primarily relies on the initial estimate obtained from the EVD-based method. The joint estimation of channel and data decoding presented by Alshamar et al. [10] uses

an algorithm of non-exponential complexity in coherence time. The reduction in complexity as compared to exhaustive search is significant. However, owing to high complexity, the scheme presented in Ref. [10] is useful in scenarios with a small RB size (T) and a large N/K ratio.

There are a few issues with the aforementioned blind channel estimation methods. They rely on asymptotic settings either in the number of BS antennas (N) or in the number of symbols in RB (T), or both. The method described in [8] has a critical requirement, i.e., distinct mapping of MTs with the eigenvalues of the estimated channel-covariance-matrix used in their algorithm. The suggested solution is to use path loss values of MTs for mapping. However, this solution is valid for infinitely large N and T values. For finite N and T values, instead of distinct mapping, only a range based mapping from users to eigenvalues/singular values is possible. For practical settings, this mapping is feasible for a very small number of MTs (K) per cell, with the condition that path loss values of different MTs have enough separation. As K increases, there is a high probability that more than two MTs and/or multiple pairs of MTs will have path loss values that are closer than the minimum range allowed for making the mapping feasible. Owing to shadow fading and several concentric rings of constant path loss around BSs, it is unrealistic to assume that the average channel gains will lie in the distinct intervals of the minimum allowed range that makes the aforementioned mapping feasible.

The method described in [7] is blind only in the sense of removing pilot contamination. Pilot training is considered that is similar to a conventional $K \times K$ system. An array gain without CSI is projected as a major advantage of the scheme. However, for a realistic N/K ratio (for example $N = 200, K = 10$ to 20), the array gain for desired signals compared to interfering signal streams is insignificant. Moreover, the concept of massive MIMO is already lost by converting the $N \times K$ channel matrix into a $K \times K$ matrix. Another issue is that the existing methods do not take advantage of the channel coherence from one Resource Block (RB) to next RB. Moreover, the combined range of different system parameters such as N, K, T, SNR (ρ_u), and temporal correlation in consecutive RBs (α_b) must be investigated while using a blind channel estimation method.

The performance of blind channel estimation techniques depends on system parameters, such as N, K, T, and QAM order. In recent years, with the appearance of the "turbo principle" [11], iterative receivers are becoming more popular. Iterative channel estimation is advocated for existing MIMO-OFDM technology in several research papers [12–14]. Slow variation (or temporal correlation) of the channel remains the basis for iterative estimation in these studies. Still, there are few such works for massive MIMO systems in the literature. The performance of iterative algorithms mainly depends on two components: the initial estimation of the channel and the convergence process. In existing methods, the initial estimate is obtained by processing the received signal using mathematical tools such as EigenValue Decomposition (EVD) and subspace projection. The convergence process employs error correction techniques such as Least Square Projection (LSP) and Least Reliable Layer (LRL) to establish the convergence.

1.1. Techniques and contribution

In the present work, a method is designed to track and blindly correct the channel matrix in massive MIMO systems by taking advantage of the inherent temporal correlation between two consecutive RBs and employing the error correction code (turbo code). The proposed algorithm connects existing techniques, such as ZF estimator, turbo coding, LLR detection, iterative estimation of temporally correlated channels, and the theory of large random matrices, to solve the prominent problem of massive MIMO systems, i.e., channel estimation. The computationally less complex channel estimation method uses the best linear unbiased estimator [15] (ZF estimator) and has better applicability (fewer limitations on system parameters) compared with existing state-of-art blind channel estimation methods; thus it is more useful in practical system

design. Analyzing the impact of a large number of system parameters (such as N, K, T, ρ_u, α_b, and C (explained later)) on the performance of the proposed scheme supports the practical usability of the work. The analysis of the different error components in iterative channel estimation and tracking—supported via simulation—leads to a better understanding of iterative channel estimation in the context of massive MIMO systems. The proposed blind channel tracking is self-sustainable, which improves the spectral efficiency of massive MIMO systems by approximately a factor of $T/(T - K)$. The performance of the proposed Iteration-Based Estimation (IBE) algorithm is compared with that of PBE.

1.2. Notation and organization of paper

Boldface (lower case) is used for column vectors (e.g., $\mathbf{x}$), and upper case is used for matrices (e.g., $\mathbf{X}$). Italic (lower case) with subscript ij is used for elements of matrices (e.g., x_{ij}); italic (lower case) with subscript i is used for elements of column vectors (e.g., x_i). Italic (upper case) is used for numerical constants. Superscript H is used for denoting a Hermitian transpose (e.g., $\mathbf{X}^H$). The mathematical symbols for expectation, absolute value, Frobenius norm, "tends to," "implies," and "pseudo inverse" are denoted by $\mathbb{E}$, $|.|$, $\|.\|$, $\rightarrow$, $\Rightarrow$, and superscript $\dagger$, respectively. The important symbols used in the paper are listed in Table 6 (Appendix C) for additional convenience.

The remainder of the paper is organized as follows: Section 2 describes the system model and correlated block-fading channel matrix used in the paper. Section 3 presents the communication strategy and the proposed algorithm. Subsections 3.3 and 3.4 present the theoretical analysis of the proposed algorithm by describing the details of the convergence process and estimation error analysis. Section 4 presents a complexity comparison between the proposed algorithm and existing algorithms. Section 5 presents a simulation, numerical results, and a discussion in which the performance of the proposed algorithm is analyzed in conjunction with the system parameters. Section 6 summarizes the important findings.

2. System model

This work considers a correlated block-fading channel model in which the channel is almost flat within an RB on a time-frequency grid, and is correlated from one RB to the next RB along the time-index. RBs across the frequency-index create parallel channels with the underlying assumption that Orthogonal Frequency Division Multiplexing (OFDM) can be used to convert a frequency-selective wideband channel into multiple frequency-flat narrow-band parallel channels with a certain amount of cyclic prefix overhead. The correlation may exist among RBs along the frequency index, but this correlation is not exploited in the proposed algorithm—it will be addressed in future work. For a wireless cellular massive MIMO system having K active MTs and one BS with N antennas, if T symbols (after removing cyclic prefixes) are received at the BS in each RB, the received matrix at the BS in a RB with index (t^b, f^b) can then be written as:

$$\mathbf{Y}(t^b, f^b) = \mathbf{H}(t^b, f^b)\mathbf{X}(t^b, f^b) + \frac{1}{\sqrt{\rho_u}}\mathbf{W}(t^b, f^b) \qquad (1)$$

where $\mathbf{Y}(t^b, f^b)$ is an $N \times T$ received complex matrix in the $(t^b, f^b)^{th}$ RB. $\mathbf{X}(t^b, f^b)$ is a $K \times T$ transmitted complex user data matrix. $\mathbf{W}(t^b, f^b)$ is an $N \times T$ complex AWGN matrix having i.i.d. zero mean and unit variance entries, and ρ_u is the uplink Signal to-Noise-Ratio (SNR) considering a symmetric channel. The system model of (1) is also valid for multi-cell settings with the assumption that inter-cell interference coming from neighboring cells is white and Gaussian distributed. In that case, the ρ_u can be considered as the uplink signal-to-noise plus inter-cell interference ratio. $\mathbf{H}(t^b, f^b)$ is an $N \times K$ complex channel matrix in the $(t^b, f^b)^{th}$ RB with zero mean unit variance entries. Elements of $\mathbf{H}(t^b, f^b)$ are i.i.d.

across antennas and users but have a certain correlation from RB index t^b to $t^b + 1$ defined by α_b such that

$$\mathrm{H}\big(t^b + 1, f^b\big) = \alpha_b \mathrm{H}\big(t^b, f^b\big) + \sqrt{\big\{1 - \alpha_b^2\big\}} \tilde{\mathrm{W}}\big(t^b, f^b\big) \qquad (2)$$

where entries of $\tilde{\mathrm{W}}(t^b, f^b)$ are zero mean and unit variance i.i.d. complex Gaussian random variables. $\tilde{\mathrm{W}}(t^b, f^b)$ is independent of $\tilde{\mathrm{H}}(t^b, f^b)$; thus:

$$\alpha_b = \mathbb{E}\Big[h_{ij}^*\big(t^b + 1, f^b\big) h_{ij}\big(t^b, f^b\big) \Big] \qquad (3)$$

$\forall 1 \le i \le N$ and $1 \le j \le K$.

Each MT transmits its data over F^b OFDM carriers (i.e., over F^b RBs for a given time index t^b). The rate -1/3 turbo code with parameters as per LTE standard is used to encode the user data across all carriers and all T symbols of each RB. Thus, the transmitted user data matrix can be represented as follows:

$$\mathrm{X}_{TX}\big(t^b\big) = F_{apn}\big(\mathrm{X}(t^b, 1), \mathrm{X}(t^b, 2), \ldots, \mathrm{X}(t^b, F^b)\big) \qquad (4)$$

where F_{apn} is a horizontal matrix concatenation function. The ZF decoded user data matrix at BS can be written as follows:

$$\mathrm{Z}\big(t^b\big) = F_{apn}\Big(\tilde{\mathrm{X}}(t^b, 1), \tilde{\mathrm{X}}(t^b, 2), \ldots, \tilde{\mathrm{X}}(t^b, F^b) \Big). \qquad (5)$$

3. Channel estimation method

3.1. Communication strategy and settings

The proposed algorithm is designed to work for a massive MIMO cellular system in TDD mode. Moreover, the size of the RBs is smaller than the coherence time of the channel, such that there exists significant temporal correlation from one block to the next block. The pilot is sent over all the OFDM carriers (i.e., RBs with $t^b = 1$ and $f^b = 1$ to F^b) once at the start of communication for the PBE of the channel. In RBs with $t^b > 1$, the MTs send uplink data encoded over all OFDM carriers with T symbols per carrier; the BS jointly estimates the channel and uplink data, precodes the downlink data using estimated channel matrices, and sends the precoded downlink data in the same RBs in which uplink data are received. MMSE estimation is used to simulate the pilot-based estimate of the channel, as follows [16]:

$$\tilde{\mathrm{H}}_{PBE}\big(t^b, f^b\big) = \frac{\rho_u^\tau}{1 + \rho_u^\tau} \left\{ \mathrm{H}\big(t^b, f^b\big) + \frac{1}{\sqrt{\rho_u^\tau}} \mathrm{W}\big(t^b, f^b\big) \right\} \qquad (6)$$

where τ is the length of the uplink pilot train. PBE is also used for all RBs separately, in order to allow a performance comparison.

3.2. Channel estimation algorithm

Fig. 1 shows the flowchart of the proposed algorithm (IBE). The algorithm starts in step (1) with an initial estimate of the channel obtained from the pilot-based estimate of the previous RB. Step (2) corresponds to the initialization of the channel estimation process for each RB, using the channel matrix of the previous RB as an initial estimate of the channel matrix of the current RB. Step (3) applies the normalized pseudo-inverse of the estimated channel matrix on the received data matrix to obtain the coded user data matrix corresponding to each OFDM carrier (each f^b index). A complete encoded user data matrix is obtained by horizontally appending the encoded user data matrices corresponding to all OFDM carriers in step (4). Step (5) generates soft bit output using an LLR detector for each of K users. Step (6) uses a turbo decoder to decode bits corresponding to each user with error correction.
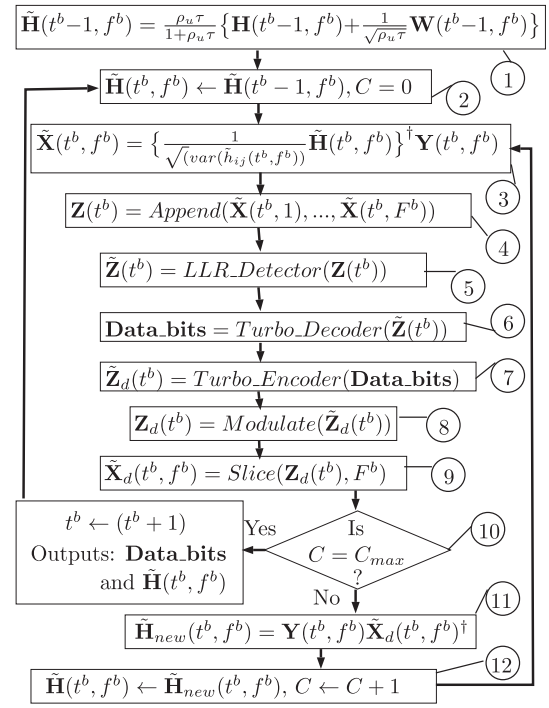

Fig. 1. Joint algorithm for channel estimation, data decoding and channel tracking.

The pseudo-inverse operation in step (3) is the Best Linear Unbiased Estimator (BLUE), which not only reduces the complexity but also performs well under certain limits on N/K, T, α_b, and ρ_u (a measure of SNR). These limits are discussed in the simulation results. Steps (5) and (6) perform error correction on user data (an essential part of an iterative algorithm). The error correction process of the computationally complex least square projection of the existing ILSP algorithm [8], which is exponentially complex in constellation size, is replaced by the computationally less complex LLR soft detection and max-log-MAP based turbo decoder [11,17]. Because the turbo code is one of the best available error correction codes, step (6) produces significant performance improvements. Moreover, because the turbo-coding error-correction is at bit-level, the modulation order does not affect performance, unlike in the ILSP-based method.

Step (7) encodes the error corrected bits using turbo coding for each user, and step (8) generates a modulated symbol matrix by modulating error-corrected encoded bits on the M-QAM constellation. Step (9) is a slicing operation (inverse of the append operation in step (4)) to partition the complete user data matrix into F^b sub-matrices corresponding to each OFDM carrier. Step (10) compares the iteration count with the maximum iteration count to complete the iteration for the current RB. If the iterations for the current RB are completed, the algorithm then switches to the next RB; otherwise, a new channel matrix estimate is calculated in step (11). This new channel matrix estimate is assigned to the initial estimate of the channel in step (12) to run the next iteration for the current RB.

The initial channel estimate for each new RB is obtained using the temporal correlation of RBs. The simplest method is to use the channel matrix of the previous RB as an initial estimate of the channel matrix for the current RB, as done in step (2). However, each channel coefficient ($h_{ij}(t^b, f^b)$) is a band-limited discrete signal with index t^b owing to the finite Doppler spectrum of wireless channel [Ch 2] [18], thus, the extrapolation for band-limited signals [19,20] can be used to obtain an even better initial estimate of the channel matrix for new RB.

3.3. Theory behind the algorithm

Theorem 3.1. *In the current system model (leaving indexes t^b and f^b):*

$$Y = HX + \frac{1}{\sqrt{\rho_u}}W,$$

for a given $\tilde{W}$ with complex zero mean unit variance i.i.d. entries and an $\alpha \geq 0$,

$$\lim_{(N \to \infty, constant\ K)} \left(\sqrt{1+\alpha^2}\right) \left\{ \frac{(H + \alpha\tilde{W})}{\sqrt{1+\alpha^2}} \right\}^{\dagger} Y = X \quad (7)$$

provided that H and $\tilde{W}$ as well as $(H + \alpha\tilde{W})$ and W are independent.

Superscript $\dagger$ represents the pseudo-inverse operation. (Proof of theorem 3.1: See Appendix A). This theorem illustrates that the estimation error can be controlled by the matrix dimensions of a massive MIMO system. When the above theorem is applied with large but finite N, T values, and a small K, there will be a finite residual error. The residual error in step (3) of the algorithm can be approximated as follows:

Let (with dropped t^b and f^b indexes),

$$\tilde{X} = \left(\sqrt{1+\alpha^2}\right) \left\{ \frac{(H + \alpha\tilde{W})}{\sqrt{1+\alpha^2}} \right\}^{\dagger} Y$$

where α is a measure of the deviation between the available channel matrix estimate and the actual channel matrix. In the algorithm, α is a function of iteration count C. However, for notational simplicity, the α is kept free of this index, where it is not required. The initial deviation, i.e., $\alpha(1)$, is a direct function of the correlation between the channel elements of two consecutive RBs (α_b) and the correlation between the estimated and actual channel elements in the previous RB (α_e). For the first block-time index after PBE (at the start of communication), these parameters can be described by the following equations.

$$\alpha_e \alpha_b = \frac{1}{\sqrt{1+\alpha^2(1)}} \quad (8)$$

$$\alpha_e = \mathbb{C}orr\left(h_{ij}(t^b - 1, f^b), \widetilde{h_{ij}}(t^b - 1, f^b)_{PBE} \right) \quad (9)$$

$$\alpha_b = \mathbb{C}orr\left(h_{ij}(t^b, f^b), h_{ij}(t^b - 1, f^b) \right) \quad (10)$$

The pilot based estimate of the channel is considered at the start of communication. By using (6) and (9), α_e can be expressed as follows:

$$\alpha_e^2 = \frac{\rho_u \tau}{1 + \rho_u \tau}. \quad (11)$$

By using the random matrix theory [21] for $N >> K >> 1$ and empirical study, the $\tilde{X}$ can be approximated as follows:

$$\tilde{X} = X + \left(\sqrt{\frac{\alpha^2(C)K}{N-K}} \right)W_1 + \left(\sqrt{\frac{1+\alpha^2(C)}{\rho_u(N-K)}} \right)W_2 \quad (12)$$

$$\tilde{X} = X + \Delta X_A + \Delta X_B \quad (13)$$

where

$$\Delta X_A \triangleq \left(\sqrt{\frac{\alpha^2(C)K}{(N-K)}}\right)W_1 \text{ and } \Delta X_B \triangleq \left(\sqrt{\frac{1+\alpha^2(C)}{\rho_u(N-K)}}\right)W_2.$$

Here, W_1 and W_2 are AWGN matrices with complex zero mean unit variance i.i.d. entries. In practice, multiplier $\sqrt{1+\alpha^2(C)}$ may not be available; however, it can be estimated using SNR estimation techniques. LLR detection and turbo-decoding further improve the accuracy of the

estimate of the data matrix; thus, the quantized output in step (9), i.e. $\tilde{X}_d$, is a better estimate of X. Similarly, the new estimate of the channel matrix in step (11) is:

$$\tilde{H}_{new} = Y\tilde{X}_d^{\dagger} \quad (14)$$

$$\tilde{H}_{new} = H + \Delta H_A + \Delta H_B \quad (15)$$

where $\Delta H_B = W_4/\sqrt{\rho_u(T-K)}$ while ΔH_A does not have an explicit relation with K, T, and ρ_u (although ΔH_A is caused by the error in $\tilde{X}_d$). ΔX_A can be minimized by improving the estimate of the channel matrix, and ΔH_A can be minimized by improving the estimate of the quantized user data matrix $\tilde{X}_d$. However, ΔX_B and ΔH_B cannot be minimized simply by improving the estimates of H and X, although they can be controlled by N, K, T, and ρ_u. Because the primary sources of errors are AWGN matrix W and the initial deviation in the channel matrix (related to $\tilde{W}$), the subsequent estimates of H are correlated to primary random matrices (W and $\tilde{W}$); thus, the algorithm cannot converge on its own without the help of LLR detection and turbo-coding error correction. In fact, the LLR detector and turbo decoders are responsible for maintaining the convergence while the system parameters (N, K, T, ρ_u) are responsible for controlling the rate of convergence. The scheme works under a finite combined boundary on N, K, T, and ρ_u. As the setting moves close to this boundary, the required number of iterations (as well as coding length) increases tremendously. On the other hand, as setting moves slightly away from this boundary, the number of required iterations is reduced significantly.

3.4. Estimation error analysis

The analysis of estimation errors in data and channel matrices versus iteration counts is useful in quantifying the required number of iterations in a given system setting. For a given iteration index and block-time index, the estimation errors in data matrix (E_X) and in channel matrix (E_H) are defined as follows:

$$E_X^2 = \mathbb{E}_{f^b}\left[\frac{1}{KT} \| \tilde{X}(t^b, f^b) - X(t^b, f^b) \|^2 \right] \quad (16)$$

$$E_H^2 = \mathbb{E}_{f^b}\left[\frac{1}{NK} \| \tilde{H}(t^b, f^b) - H(t^b, f^b) \|^2 \right] \quad (17)$$

Fig. 2 shows the simulated estimation errors along with a linearly fitted curve for errors. Because the behavior of E_H^2 is similar to that of E_X^2, we analyze only the error in the data matrix, which has a detailed analytical expression. From (12), the E_X^2 can be written as follows:

$$E_X^2 = \left(\frac{\alpha^2(C)K}{N-K} \right) + \left(\frac{1+\alpha^2(C)}{\rho_u(N-K)} \right) \quad (18)$$

The actual behavior of the estimation error with respect to the iteration count is less significant. Rather, the required number of iterations for convergence is more important. Thus, we use a simplified linear curve fitting to find the empirical relation between the iteration counts and the system parameters. The linear fit for E_X^2 can be written as follows:

$$E \overset{(q_1)}{=} E_1 + \frac{E_2 - E_1}{C_2 - C_1}(C - C_1) \quad (19)$$

At the start, $E_1 \overset{(q_2)}{=} \left(\frac{\alpha^2(1)K}{N-K} \right) + \left(\frac{1+\alpha^2(1)}{\rho_u(N-K)} \right) \quad (20)$

At the end, $E_2 \overset{(q_3)}{=} \left(\frac{\alpha^2(\infty)K}{N-K} \right) + \left(\frac{1+\alpha^2(\infty)}{\rho_u(N-K)} \right) \quad (21)$
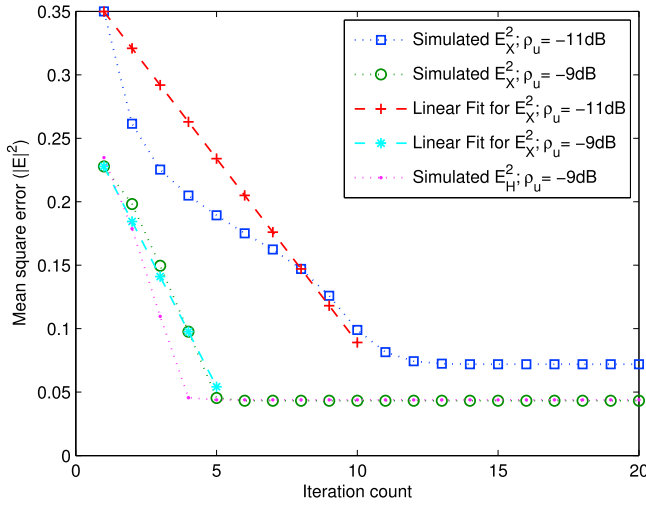
Fig. 2. Data and channel estimation error versus iteration counts for different values of ρ_u (SNR) with $\alpha_b = 0.8$, $K = 20$, $N = 200$ and $T = 200$.

Table 1
Parameters values (rounded) for establishing empirical expression for C_2 ($\alpha_b = 0.8$), $N = 200$, $K = 20$, and ($T = 200$).

$\{\rho_u\}$	$\alpha(1)$	$\alpha(\infty)$	C_2	$\frac{\alpha^2(1)+\alpha^2(\infty)}{\rho_u}$	PF
0.08	1.24	0.26	10.23	20.36	0.50
0.13	1.09	0.21	5.08	9.75	0.52
0.13	1.52	0.21	8.95	18.67	0.48

where $\alpha^2(\infty) \overset{(q_4)}{=} \dfrac{1}{\rho_u(T-K)}$. (22)

Here, q_1 is the linear curve fitting equation for E_X^2. The initial estimation error (E_1) corresponds to the first iteration count ($C_1 = 1$) and E_2 is the estimation error when convergence completes. Parameters q_2 and q_3 follow from (18) and q_4 follows from (15) with $\Delta H_A = 0$.

Now we fit (19) on the simulated mean square error plots in the data matrix, using the Least Absolute Residual (LAR) method to obtain the required iteration count C_2. Using the curve fitting data, an empirical relation is established as follows:

$$C_2 = PF\frac{\alpha^2(1)+\alpha^2(\infty)}{\rho_u}$$ (23)

where PF is a proportionality factor corresponding to the capability of the error correction code (in this case, turbo code). For the current setup, the empirically obtained PF is approximately 0.5, as shown in Table 1.

3.5. Improvement in spectral efficiency

The proposed scheme provides an advantage in spectral efficiency because the degrees of freedom used for pilots in PBE are now free for user data transmission. The spectral efficiency for a massive multi-user MIMO system can be given by the following [22], [eq:(14)].

$$SE = \sum_{k=1}^{K}\left(\frac{B}{\beta}\right)\left(\frac{T_{slot}-T_{pilot}}{T_{slot}}\right)\left(\frac{T_u}{T_{slot}}\right)\log_2(1+SIR_k)$$ (24)

where B = total bandwidth, β = frequency reuse factor, $T_{slot} = T$ = number of symbols in RB, $T_{pilot} = K$ = number of pilot symbols in RB, SIR_k is the signal-to-interference ratio for k^{th} MT, and T_u is the useful OFDM slot. Using the above equation, there is a gain in spectral efficiency by a factor of at least $T/(T-K)$ in IBE compared to PBE, for system parameter ranges in which Symbol Error Probability (SEP) in IBE

is less than or equal to the SEP in PBE.

4. Complexity analysis

This section compares the number of multiplications in the proposed algorithm with those in existing EVD-ILSP based channel estimation algorithms. The number of multiplications (and divisions) per bit for turbo decoders with the number of states S and turbo-iterations count C_T is given by $8SC_T$ [23], [Table 2, Fig. 2]. For the algorithm, the number of multiplications per symbol with C number of iterations can be written as follows:

$$MUL_{TDC} = 8SC_T C \log_2(M).$$ (25)

The complexity (per symbol) associated with linear operation $\tilde{H}(t^b, f^b)^{\dagger} Y(t^b, f^b)$ is:

$$MUL_{lop1} = \frac{(2NK+K^2+NT)}{T}C.$$ (26)

Similarly, the complexity (per symbol) associated with linear operation $Y(t^b, f^b)\tilde{X}_d(t^b, f^b)^{\dagger}$ is:

$$MUL_{lop2} = \left(2K+K^2/T+N\right)C.$$ (27)

In the proposed algorithm, operations corresponding to Eqs. (25–27) are key contributors to the complexity. Thus, the overall number of multiplications in the proposed algorithm can be approximated as follows:

$$MUL_{IBE} = \left\{\left(2NK+K^2+NT\right)/T + \left(2K+K^2/T+N\right)\right. $$
$$\left. + 8SC_T\log_2(M)\right\}C.$$ (28)

The least square solution for the ILSP-based iterative algorithm used in Ref. [8] requires 2^M multiplications (M: modulation order). Thus, the complexity associated with ILSP-based channel estimation when using a turbo decoder (after the last iteration) can be written as follows:

$$MUL_{ILSP} = \left\{\left(2NK+K^2+NT\right)/T + \left(2K+K^2/T+N\right)+2^M\right\}C $$
$$+ 8SC_T \log_2(M).$$ (29)

The EVD-based blind (almost) method has been used for initial estimation in existing ILSP-based estimation methods. The complexity associated with EVD-based estimation can be written as follows:

$$MUL_{EVD} = \left\{\frac{N^2}{K}+\frac{N^2}{T}+\frac{NK}{T}+N+K+\frac{8T^2}{K}+\frac{8NT}{K}+\frac{4N}{K}+\frac{4T}{K}\right\}$$ (30)

The calculations for determining the number of multiplications in the mathematical operations in (30) are provided in Appendix B. Fig. 3 shows a comparison of the complexity associated with the proposed algorithm, the EVD-based algorithm and the ILPS-based algorithm. The turbo iteration count (C_T) is 4, the number of states in the turbo-encoder (S) is 16, and the algorithm iteration count (C) is 8. The comparison plot shows that ILPS is not suitable above the 4-QAM scheme, and the EVD-based algorithm is more complex than the proposed algorithm. Moreover, the EVD-based algorithm provides the only initial estimate of the channel, which in the case of the proposed algorithm is derived from the previous estimate without computationally complex operations.

5. Simulation results and discussion

The simulation results of the current scheme (i.e., IBE) are compared with the generic channel estimation technique for massive MIMO systems, i.e., PBE. PBE is a well-explored scheme of the non-blind class of channel estimation; thus, comparison with PBE is assumed to be useful. However, there are few existing results (as described in the introduction) that simulate the performance of their schemes, and they are considered

less suitable for comparison because of their use of specific settings, such as large N/K ratios (for better orthogonality of channel vectors), small RBs (for lower complexity), large RBs (for better estimation of the covariance matrix), and lower modulation orders. Table 2 summarizes the settings used for the presented simulation.
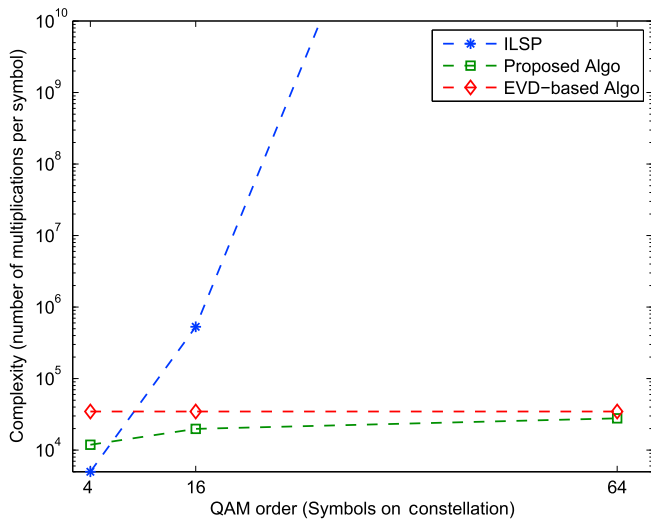
5.1. SEP verses iteration count

Fig. 4 and Fig. 5 present results that support the simulation for the working of the algorithm and the section "Theory behind the algorithm." Convergence in SEP with iteration reflects the strength of the algorithm. As shown in Fig. 4, SEP plots fall very quickly with iteration count for better SNRs ($-9.5dB$ in the figure). Further, the system can support as
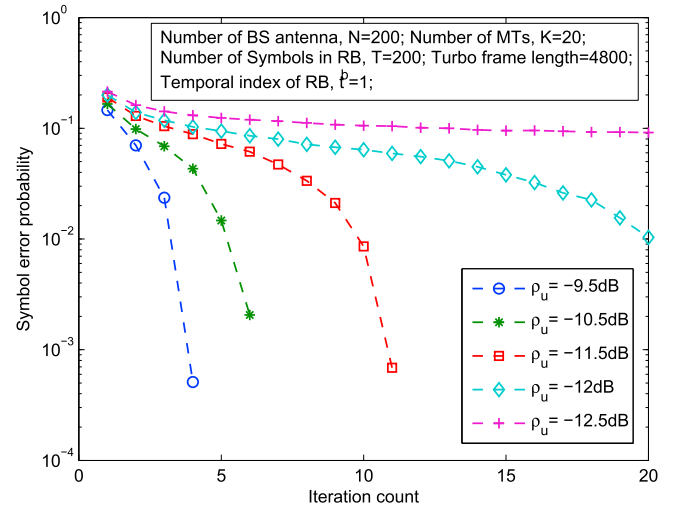
Table 2
Summary of simulation parameters.

Parameters	Values	Details
N	100 to 300	Number of BS antennas
K	18 to 24	Number of active MTs
T	150 to 300	Number of symbols in RB
F_b	8 to 15	Number of RBs on frequency axis
T_b	5	Number of RBs considered on time axis
M	64	QAM-order
α_b	0.6 to 0.9	Time correlation in RBs

SNR that is $2dB$ lower($-11.5dB$ in the figure) by increasing iteration counts to around 12. However, in further lowering the SNR, the boundary values are reached; consequently, an exponential increase in iteration counts is required. Fig. 5 shows the strength of the algorithm against different correlation values from one RB to the next RB. The convergence is made possible by increasing the iteration count in a certain range of correlation α_b for workable values of N, K, and SNR, as shown by different plots in Fig. 5. However, the behavior of SEP vs the iteration count plot is slightly different in the case of decreasing α_b compared with the case of decreasing ρ_u, because the α_b is connected to both of the unwanted terms in (12) while ρ_u is connected to only one term.

The plots in Figs. 4 and 5 correspond to the first RB ($t^b = 1 \ \forall \ 1 \le f^b \le F_b$), where the initial estimate of the channel matrix is obtained from the PBE of the previous block. Becasue the channel matrix estimate coming from the previous RB also improves because of this scheme, there are further performance improvements in terms of the required iterations or supported SNR in the steady state (after few block times).
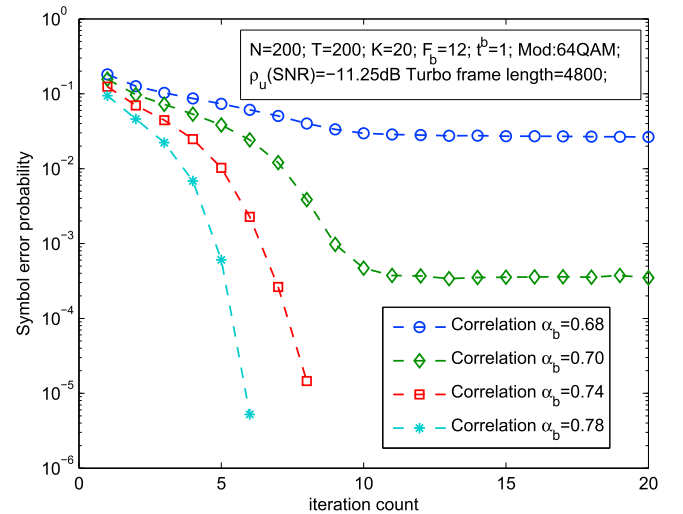

Fig. 4. Symbol error probability versus iteration counts for different values of ρ_u (SNR) with $\alpha_b = 0.8$, $K = 20$, $N = 200$ and $T = 200$.

5.2. Impact of the number of MTs

Fig. 6 shows the SEP versus SNR performance of PBE and IBE for different numbers of active MTs. The number of IBE iterations is set to 10. SEP in IBE falls toward zero significantly earlier than PBE when the number of active MTs decreases. The performance of IBE is better than that of PBE down to the conventionally proposed N/K ratio (i.e. (≈ 10). Because N, K, T, and ρ_u form a common boundary, there can be a trade-off among these parameters to allow more users, lower coherence time (i.e., lower block size), or lower SNR. The impact of the number of BS antennas is similar but opposite to the impact of K, which is also in line with theory because N and K are coupled in estimation error expressions (cf. equation (12)).

5.3. Impact of the mobility

The limited coherence time in massive MIMO systems caused by the mobility of MTs is one of the fundamental bottlenecks in the performance of the system. The lower coherence time results in reduced temporal correlation between consecutive RBs as well as channel variations inside the RB. The number of symbols (T) in an RB must be reduced in order to maintain the time invariant consideration of the channel inside the RB. T

Fig. 3. Complexity versus constellation size with $K = 20$, $N = 200$ and $T = 200$.

Fig. 5. Symbol error probability versus iteration counts for different values of α_b with $\rho_u = -11.25dB$, $K = 20$, $N = 200$ and $T = 200$.
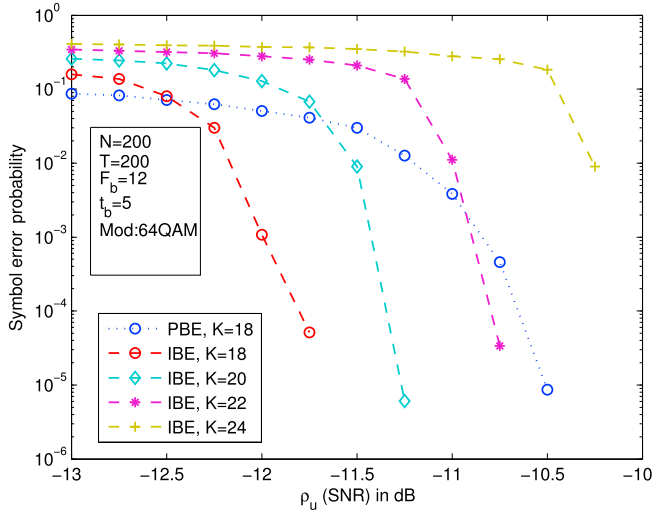
Fig. 6. Symbol error probability versus ρ_u (SNR) for different values of K with $\alpha_b = 0.8$, $N = 200$, $T = 200$ and $t^b = 5$.

Fig. 8. Symbol error probability versus ρ_u (SNR) for different values of T with $K = 20$, $N = 200$, $\alpha_b = 0.8$ and $t^b = 5$.
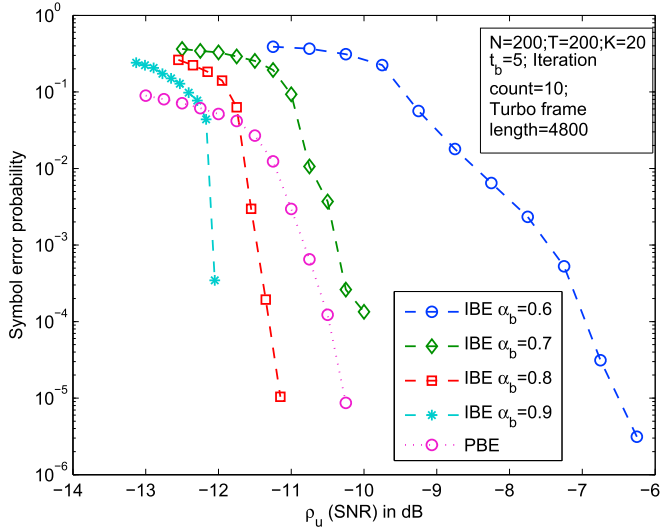

Fig. 7. Symbol error probability versus ρ_u (SNR) for different values of α_b with $K = 20$, $N = 200$, $T = 200$ and $t^b = 5$.

is directly connected to the spectral efficiency in PBE and the error performance in IBE. The proposed algorithm is connected to both the temporal correlation (α_b) and the number of symbols in the RB (T). Thus, we study the impact of both α_b and T on the performance.

Fig. 7 shows the SEP versus SNR performance of PBE and IBE for different time correlation values between consecutive RBs. The SEP performance improves with correlation between two consecutive RBs. A smaller time correlation between consecutive RBs introduces more deviation into the initial estimate of the channel matrix. Consequently, the error components in IBE grow without bounds. On the other hand, provided that the channel estimate in the previous RB is good (i.e., a small α_e), the increasing correlation α_b minimizes the error components in IBE that are dependent on temporal correlation.

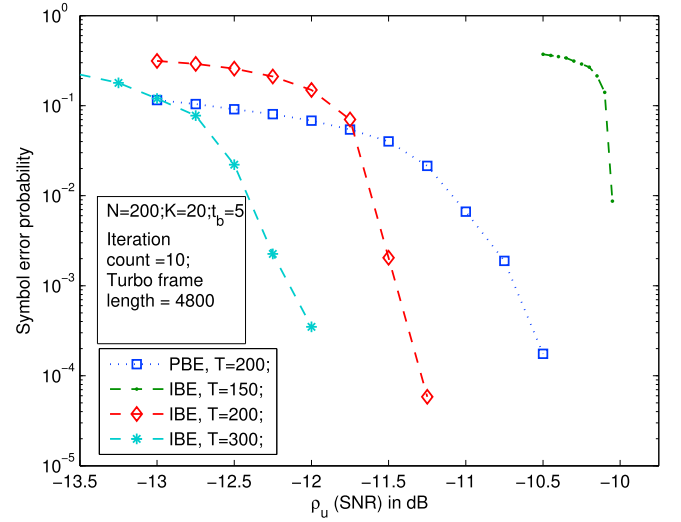Fig. 8 shows the SEP versus SNR performance of PBE and IBE for

different numbers of symbols in an RB. For IBE, the SEP versus SNR plots shift to a lower SNR when T increases, because the error components in IBE depend on T. The SEP in PBE does not depend on T, but there is a penalty on spectral efficiency as given in (24) for a smaller T. Another observation is that the SEP falls more quickly with ρ_u for a smaller T in IBE. The reason behind this sharp fall is the dominance of error component ΔH_A. This error component does not depend directly on the SNR and it is reduced effectively by the algorithm.

However, the proposed scheme can handle high mobility cases up to an extent by adjusting the iteration count (C_{max}), N, K, and ρ_u, as shown in different performance plots. The present work only considers the processing at the BS. When the processing at MTs is considered in conjunction with that of the BS, the mobility of MTs could be exploited in favor of massive MIMO systems, because opportunistic communication is advocated as a beneficial strategy in literature for mobile networks [24–26].

6. Conclusion

This work provides a simplified solution to a very fundamental problem in cellular massive MIMO systems i.e., estimation of the channel matrix. An iterative blind channel estimation and tracking algorithm is presented, which exploits the temporal correlation of the channel and the capability of turbo code. The theoretical and simulation based analysis of the algorithm provides insights on the convergence process and different error components in the estimation. The algorithm has lower computational complexity compared to previous similar algorithms [8,10]. The simulation study of the proposed algorithm's performance, along with a comparison against pilot-based estimation, provides a detailed evaluation of the combined effect of system parameters (N, K, T, ρ_u, and α_b). SEP plots of uplink data for a practical value of N/K ($\approx$10) from the proposed scheme indicate better performance than the SEP plots for PBE.

Acknowledgments

Appendix A

Proof of Theorem 3.1: Let

$$\widetilde{Y} = \lim_{\substack{N\to\infty,\\ \text{constant } K}} \sqrt{1+\alpha^2}\left\{\left(\frac{H+\Delta H}{\sqrt{1+\alpha^2}}\right)^H \left(\frac{H+\Delta H}{\sqrt{1+\alpha^2}}\right)\right\}^{-1} \times \left(\frac{H+\Delta H}{\sqrt{1+\alpha^2}}\right)^H Y$$

Columns of $\frac{H+\Delta H}{\sqrt{1+\alpha^2}}$ are asymptotically orthogonal, consequently $\lim_{\substack{N\to\infty,\\ \text{constant } K}} \left\{\left(\frac{H+\Delta H}{\sqrt{1+\alpha^2}}\right)^H \left(\frac{H+\Delta H}{\sqrt{1+\alpha^2}}\right)\right\}\Big/ N\to I$ so

$$\widetilde{Y} = \lim_{\substack{N\to\infty,\\ \text{constant } K}} \frac{(H+\Delta H)^H Y}{N}$$

$$\widetilde{Y} = \lim_{\substack{N\to\infty,\\ \text{constant } K}} \frac{(H+\Delta H)^H (HX)}{N} + \frac{(H+\Delta H)^H W}{N\sqrt{\rho_u}}$$

$$\widetilde{Y} = \lim_{\substack{N\to\infty,\\ \text{constant } K}} \frac{H^H HX}{N} + \frac{\Delta H^H HX}{N} + \frac{(H+\Delta H)^H W}{N\sqrt{\rho_u}}$$

By representing matrix H as $[h_{ij}]$ with $1 \le i \le N$ & $1 \le j \le K$, ΔH as $[\delta h_{ij}]$ with $1 \le i \le N$ & $1 \le j \le K$ and W as $[w_{ij}]$ with $1 \le i \le N$ & $1 \le j \le T$,

$$\widetilde{Y} = X + \lim_{\substack{N\to\infty,\\ \text{constant } K}}\left[\frac{1}{N}\sum_{n=1}^N \delta h_{ni}^* h_{nj}\right]X + \lim_{\substack{N\to\infty,\\ \text{constant } K}}\left[\frac{1}{N\sqrt{\rho_u}}\sum_{n=1}^N \left(h_{ni}^* + \delta h_{ni}^*\right)w_{nj}\right]$$

Because δh_{ni}^* and h_{nj} as well as $(h_{ni}^* + \delta h_{ni}^*)$ and w_{nj} are independent,

$$\widetilde{Y} = X + [0]X + [0] \Rightarrow \widetilde{Y} = X.$$

Appendix B

Block indexes (t^b) and (f^b) are omitted to keep the notation simple.

B.1 Complexity calculation for equation (26)

Table 3 shows the complexity calculation for equation (26).

Table 3
Number of multiplications for pseudo inverse operation using channel matrix.

Operation	No. of multiplications
$\bar{H}^H \bar{H}$	$K^2 N$
$(\bar{H}^H \bar{H})^{-1}$	$K^2 N + K^3$
$(\bar{H}^H \bar{H})^{-1}\bar{H}^H$	$K^2 N + K^3 + K^2 N$
$(\bar{H}^H \bar{H})^{-1}\bar{H}^H Y$	$K^2 N + K^3 + K^2 N + NKT$
Complexity per symbol	$(2KN + K^2 + NT)/T$

B.2 Complexity calculation for equation (27)

Table 4 shows the complexity calculation for equation (27).

Table 4
Number of multiplications for pseudo inverse operation using data matrix.

Operation	No. of multiplications
$\bar{X}_d \bar{X}_d^H$	$K^2 T$
$(\bar{X}_d \bar{X}_d^H)^{-1}$	$K^2 T + K^3$
$\bar{X}_d^H (\bar{X}_d \bar{X}_d^H)^{-1}$	$K^2 T + K^3 + K^2 T$
$Y\bar{X}_d^H (\bar{X}_d \bar{X}_d^H)^{-1}$	$K^2 T + K^3 + K^2 T + NKT$
Complexity per symbol	$2K + K^2/T + N$

B.3 *Complexity calculation for EVD-based Channel estimation*

Definitions of variables in Table 5 follow from Ref. [8]. Using Table 5, the total number of multiplications associated with EVD-based channel estimation are:

$$N^2T + N^2K + NK^2 + NKT + K^2T + 8T^3 + 8NT^2\nu + 4NTL + 4T^2.$$

With $\nu = 1$ and $L = 1$, complexity per symbol is:

$$\frac{N^2}{K} + \frac{N^2}{T} + \frac{NK}{T} + N + K + \frac{8T^2}{K} + \frac{8NT}{K}\nu + \frac{4N}{K} + \frac{4T}{K}.$$

Table 5
Number of multiplications for EVD-based channel estimation algorithm.

Operation	No. of multiplications
$R_y = \frac{1}{T}\sum_{t=1}^{T}yy^H$	N^2T
R_yH_{ll}	N^2K
$U_{ll}\Xi$	NK^2
$A_n = \sqrt{P_t}U_{ll}D^{-1/2}X_n$	$NKT + K^2T$
$\widehat{\xi} = (\sum_{n=1}^{\nu}\overline{A}_n^T\overline{A}_n)^{-1}\sum_{n=1}^{L}\overline{A}_n^Ty_n$	$8T^3 + 8T^2N\nu + 4NTL + 4T^2$

Appendix C

Table 6 lists the different symbols used in this work. The symbols H, X, and W have different variants ($\tilde{H}$, $\tilde{X}$, W_1, W_2 etc.) for representing the similar matrices corresponding to the details given in Table 6.

Table 6
List of symbols.

Symbol	Details
N	Number of BS antennas
K	Number of active MTs
T	Number of symbols in a RB
T_b	Number of RBs on time axis
F_b	Number of RBs on frequency axis
t^b	Time index of RB
f^b	Frequency index of RB
H	Channel matrix
X	User data matrix
W	AWGN matrix
ρ_u	Measure of uplink SNR
τ	Length of uplink pilot
α_e	Correlation between estimated and true channel elements in previous RB
α_b	Correlation between current and previous RBs' channel elements
α	Deviation in initial estimate of channel matrix from true value
M	QAM-order

References

[1] L. Lu, et al., An overview of massive mimo: benefits and challenges, IEEE J. Sel. Top. Signal Process 8 (5) (2014) 742–758.
[2] E.G. Larsson, O. Edfors, F. Tufvesson, T.L. Marzetta, Massive mimo for next generation wireless systems, IEEE Commun. Magz (2014) 186–195.
[3] F. Rusek, D. Persson, B.K. Lau, E.G. Larsson, T.L. Marzetta, O. Edfors, F. Tufvesson, Scaling up mimo opportunities and challenges with very large arrays, IEE Sig. Proces. (2013) 40–60.
[4] X. Jiang, F. Kaltenberger, L. Deneirey, How accurately should we calibrate a massive mimo tdd system?, in: IEEE ICC2016-workshops: W09-workshop on 5G RAN Design, IEEE, 2016, pp. 706–711.
[5] R.C. de Lamare, Massive Mimo Systems: Signal Processing Challenges and Research Trends, arXiv:1310.7282, Oct. 2013.
[6] E. Bjornson, E.G. Larsson, T.L. Marzetta, Massive Mimo 10 Myths and One Grand Question, arXiv:1503.06854v1 [cs.IT], March 2015.
[7] R.R. Muller, L. Cottatellucci, M. Vehkaper, Blind pilot decontamination, IEEE J. Sel. Top. Sig. Proces. 8 (5) (2014) 773–886.
[8] H.Q. Ngo, E.G. Larsson, Evd-based channel estimation in multicell multiuser mimo systems with very large antenna arrays, in: International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 3249–3252.

[9] S. Talwar, M. Viberg, A. Paulraj, Blind separation of synchronous co-channel digital signals using an antenna array-part i, IEEE Trans. Signal Process 44 (5) (1996) 1184–1197.
[10] H.A.J. Alshamary, W. Xu, Efficient optimal joint channel estimation and data detection for massive mimo systems, in: Information Theory (ISIT), IEEE, 2016.
[11] C. Berrou, A. Glavieux, P. Thitimajshima, Near shannon limit error-correcting coding and decoding: turbo-codes, in: ICC, IEEE, 1993, pp. 1064–1070.
[12] F. Kadrija, M. Simko, M. Rupp, Iterative channel estimation in lte systems, in: Smart Antennas (WSA), IEEE, 2013.
[13] Y. Liu, S. Sezginer, Iterative compensated mmse channel estimation in lte systems, in: Communications (ICC), IEEE, 2012.
[14] Z. Du, X. Song, J. Cheng, N.C. Beaulieu, A convergence study of iterative channel estimation algorithms for ofdm systems in dispersive time-varying channels, in: Wireless Communications and Networking Conference (WCNC), IEEE, 2009.
[15] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall PTR, 1993.
[16] H.Q. Ngo, E.G. Larsson, T.L. Marzetta, Energy and spectral efficiency of very large multiuser mimo systems, IEEE Trans. Commun. 61 (4) (2013) 1436–1449.
[17] Y. Wu, On the complexity of turbo decoding algorithms, in: Vehicular Technology Conference, 2001, pp. 1439–1443.
[18] D. Tse, P. Viswanath, Fundamentals of Wireless Communication, Cambridge university press, 2005.

[19] S. Dharanipragada, K.S. Arun, Bandlimited extrapolation using time-bandwidth dimension, IEEE Trans. Signal Proces. 45 (12) (1997) 2951–2966.

[20] J. Shi, X. Sha, Q. Zhang, N. Zhang, Extrapolation of bandlimited signals in linear canonical transform domain, IEEE Trans. Signal Proces. 60 (3) (2012) 1502–1508.

[21] A. Tulino, S. Verdu, Random Matrix Theory and Wireless Communications, Foundation and Trends InCommunications and Information Theory, Now Publishers, Inc., Delft, The Netherlands, 2004.

[22] T.L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, IEEE Trans. Wirel. Commun. 9 (11) (2010) 3590–3600.

[23] I.A. Chatzigeorgiou, M.R.D. Rodrigues, I.J. Wassell, R.A. Carrasco, Comparison of convolutional and turbo coding for broadband fwa systems, IEEE Trans.. Broadcast. 53 (2) (2007) 494–503.

[24] Y. Liu, A.M.A.E. Bashar, F. Li, Y. Wang, K. Liu, Multi-copy data dissemination with probabilistic delay constraint in mobile opportunistic device-to-device networks, in: World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2016.

[25] W. Zirwas, Opportunistic comp for 5g massive mimo multilayer networks, in: Smart Antennas (WSA 2015), VDE, 2015.

[26] Y. Liu, Z. Yang, T. Ning, H. Wu, Efficient quality-of-service (qos) support in mobile opportunistic networks, IEEE Trans. Veh. Technol. 63 (9) (2014) 4574–4584.