This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

A Comprehensive Survey of Pilot Contamination in Massive MIMO - 5G System

Olakunle Elijah, *Student Member, IEEE,* Chee Yen Leow, *Member, IEEE,* Tharek Abdul Rahman, *Member, IEEE*, Solomon Nunoo, *Member, IEEE,* Solomon Zakwoi Iliya *Member, IEEE,*

***Abstract*—Massive MIMO has been recognized as a promising technology to meet the demand for higher data capacity for mobile networks in 2020 and beyond. Although promising, each base station needs accurate estimation of the channel state information (CSI), either through feedback or channel reciprocity schemes in order to achieve the benefits of massive MIMO in practice. Time division duplex (TDD) has been suggested as a better mode to acquire timely CSI in massive MIMO systems. The use of non-orthogonal pilot schemes, proposed for channel estimation in multi-cell TDD networks, is considered as a major source of pilot contamination in the literature due to the limitations of coherence time. Given the importance of pilot contamination in massive MIMO systems, we provide an extensive survey on pilot contamination; and identify other possible sources of pilot contamination, which include hardware impairment and non-reciprocal transceivers. We review established theories that have analyzed the effect of pilot contamination on the performance of massive MIMO systems, particularly on achievable rates. Next, we categorize the different proposed mitigation techniques for pilot contamination using the following taxonomy: pilot-based approach and subspace-based approach. Finally, we highlight the open issues, such as training overhead, deployment scenario, computational complexity, use of channel reciprocity and conclude with broader perspective and a look at future trends in pilot contamination in massive MIMO systems.**

***Index Terms*—5G network, channel state information (CSI), massive multiple-input multiple-output (MIMO), pilot contamination, time-division duplex system (TDD), frequency-division duplex system (FDD).**

I. Introduction

THE increasing demand for higher data rates in wireless mobile communication systems, and the emergence of services like internet of things (IoT), machine-to-machine communication (M2M), e-health, e-learning and e-banking have prompted the need for new technologies that are capable of providing higher capacity compared to the existing cellular network technologies. It is projected that mobile traffic will increase in the next decade in the order of thousands compared to current demand; hence, the need for next generation networks that can deliver the expected capacity compared to existing network deployment [2]. According to Cisco networking index, global mobile data traffic grew 69 percent in 2014, making it 30 times the size of the entire global internet in 2000. The index also shows that wireless data explosion is real and increasing at an exponential rate, which is driven largely by the increased use of smart phones and tablets, and video streaming [3].

The current 4G cellular networks identified by the international telecommunication union (ITU), i.e. 3rd Generation Partnership Project (3GPP) LTE-Advance and IEEE 802.16m cellular networks [4] were designed with the key attributes to support peak spectral efficiency of 15 bps/Hz, cell average spectral efficiency of 2 bps/Hz, bandwidth of 100 MHz and ultra-low latency. However, the projected traffic growth far exceeds the capabilities of current 4G, hence the need for 5G cellular network. 5G is expected to introduce new technology components, such as massive MIMO, ultra dense and reliable networks, device-to-device communication and massive machine communication [5]. Collaborative projects such as Mobile and Wireless Communications Enablers for the Twenty-twenty Information Society (METIS) and 5GNOW, in addition to other research groups mentioned in [6] mainly drawn from the academia, industry and public-private partnership, have taken up the challenge to meet the intense demand [7]. Key technology components that have been identified under the METIS that require significant advancement are radio links, multi-node/multi-antenna technologies, multi-layer and multi-RAT networks, and spectrum usage [8]. In multi-node/ multi-antenna technologies, massive MIMO is currently being studied in order to deliver very high data rates and spectral efficiency, as well as enhanced link reliability, coverage and energy efficiency [9].

Massive MIMO, as illustrated in Fig. 1, is a communication system where a base station (BS) with a few hundred antennas array simultaneously serve many tens of user terminals (UTs), each having a single antenna, in the same time-frequency resource [10]. The BS with multiple antennas sends independent data streams to multiple terminals in the same time-frequency resource. Use of massive MIMO is a promising technology that is expected to deliver high data rates as well as enhanced link reliability, coverage, and/or energy efficiency; and has therefore attracted lots of research interests. Several studies have been carried out to explore its benefits as well as limitations and challenges through theoretical studies,

Part of the materials in this paper was presented at the IEEE 10th Asian Control Conference (ASCC) 2015 [1]. This research is supported by the Ministry of Science, Technology and Innovation Malaysia (MOSTI), the Ministry of Education Malaysia (MOE) and Universiti Teknologi Malaysia under Project Vote No. 4S079, 4F261, 05H39 and 07085.

Olakunle Elijah, Chee Yen Leow, Tharek Abdul Rahman are with Wireless Communication Center, Universiti Teknologi Malaysia, Malaysia, (UTM). e-mail: elij_olak@yahoo.com, bruceleow,tharek@fke.utm.my.

Solomon Nunoo is with the Department of Electrical and Electronic Engineering, University of Mines and Technology, Ghana. e-mail: snunoo@ieee.org.

Solomon Zakwoi Iliya, is with the Center for Satellite Technology Development (CSTD), an arm of the National Space Research and Development Agency (NASRDA). email: solozakwoi@gmail.com.

simulation and prototyping [10]–[12].

The basic advantages offered by the features of massive MIMO can be summarized as follows:

- **Multiplexing gain:** Aggressive spatial multiplexing used in massive MIMO makes it theoretically possible to increase the capacity 10 times more [10].
- **Energy efficiency:** The large antenna arrays can potentially reduce uplink (UL) and downlink (DL) transmit powers through coherent combining and an increased antenna aperture [13]. It offers increased energy efficiency in which UL transmit power of each UT can be reduced inversely proportional to the number of antennas at the BS with no reduction in performance [13], [14].
- **Spectral efficiency:** The large number of service antennas in massive MIMO systems and multiplexing to many users rather than beamforming to a single user provides the benefit of spectral-efficiency [13], [15].
- **Increased robustness and reliability:** Basically, the large number of antennas allows for more diversity gains that the propagation channel can provide. This in turn leads to better performance in terms of data rate or link reliability. Also, when the number of antennas increases without bound, uncorrelated noise, fast fading, and intra-cell interference vanish [11], [12].
- **Simple linear processing:** Because BS station antenna is much larger than the UT antenna $(M \gg K)$, simplest linear pre-coders and detectors are optimal [11], [16].
- **Cost reduction in radio frequency (RF) power components:** Due to the reduction in energy consumption, the large array of antennas allows for use of low cost RF amplifiers in the milli-Watt range [12].

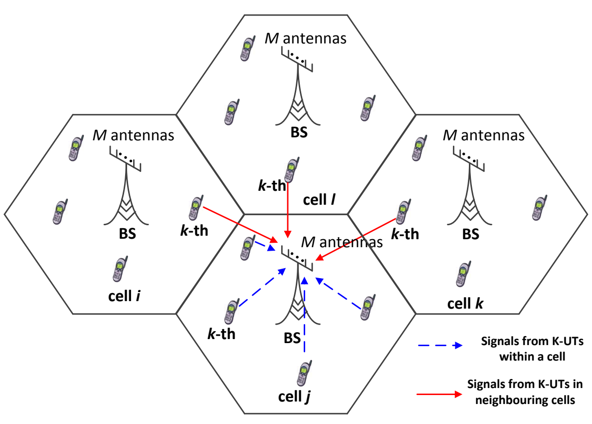


Fig. 1. Illustration of massive MIMO system with uniform distribution of UTs (K) and large number of BS antennas (M) in (L) number of cells

Although promising, practical implementation of massive MIMO poses a research challenge. In this regard, real-world challenges such as channel estimation and pilot design, antenna calibration, link adaptation and propagation effects, theoretical limits in massive MIMO, and use of massive MIMO in millimeter-wave bands have been considered in METIS [9].

To achieve the benefits of massive MIMO in practice, each BS needs accurate estimation of the channel state information (CSI), either through feedback or channel reciprocity schemes. In view of this, frequency-division duplex (FDD) or time-division duplex (TDD) has been considered in the literature. TDD is considered a better mode to acquire timely CSI in advance wireless systems over FDD because TDD requires estimation, which can be done in one direction and used in both directions; while FDD requires estimation and feedback for both forward and reverse directions, respectively [10], [12], [17]. Under METIS, TDD has been suggested for massive MIMO [8] because of efficient usage of radio resources. However, [18]–[21] proposed the use of FDD. The possibility of using FDD scheme will continue to attract more interest because FDD is common among existing cellular networks. Our focus however, is on the use of TDD scheme in a massive MIMO system. This paper intends to examine the sources and major effects of pilot contamination in massive MIMO systems in the TDD scheme with the aim of reviewing proposed methods in eliminating or reducing pilot contamination in multi-cell systems within the last decade.

In TDD architecture, the use of channel reciprocity and training signals (pilot) in the UL are key features for its application. Using the notion of reciprocity, it is assumed that the forward channel is equal to the transpose of the reverse channel for the purpose of theoretical studies. Hence, the required channel information is obtained from transmitted pilots on the reverse link from UTs [22]. However, in practice, an antenna calibration scheme must be implemented at the transmitter side and/or the receiver side owing to the different characteristics of transmit or receive RF chains [23].

The number of pilot symbols for CSI estimation are considered in [22], [24]–[26]. In [22], the minimum number of UL pilot symbols equal to the number of UTs was suggested, while in [24], it was shown that optimal number of training symbols can be larger than the number of antennas if training and data power are required to be equal. In the majority of studies carried out on pilot contamination, it is assumed that the same size of pilot signals are used in all cells. Contrary to this assumption, the studies in [25], [26] have shown that arbitrary pilot allocation is possible in multi-cell system.

Spectral efficiency in wireless communication has been an important research topic which requires appropriate frequency or time or pilot reuse factors in order to maximize system throughput [27]. The reuse of frequency has been shown to provide more efficient use of the limited available spectrum, but it also introduces unavoidable co-channel interference [28]. In a massive MIMO TDD system, the pilot signals which are used to estimate the channels can be contaminated as a result of reuse of non-orthogonal pilot signals in a multi-cell system [29], [30]. This phenomenon causes the inter-cell interference that is proportional to the number of BS antennas [31], which in turn reduces the achievable rates in the network [32] and affect the spectrum efficiency.

There are several studies on eliminating inter-cell interference in multi-cell systems in which it is assumed that the BSs are aware of CSI. For instance, coordinated beamforming have been proposed in multi-cell multi-antenna wireless systems in eliminating inter-cell interference with the assumption that the CSI of each UT is available at the BS [33]–[36]. However, in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

practical implementation, estimation of channel state information is required. In the asymptotic regime, where the BS is equipped with unlimited number of antennas and there is no cooperation in the cellular network, it was shown that not all interference vanishes because of reuse of orthogonal training sequences across adjacent cells leading to inter-cell interference [11], [37], [38]. Several methods have been proposed on reduction of inter-cell interference with emphasis on mitigation of pilot contamination in channel estimation. Although most authors have focused on the reuse of non-orthogonal training sequence as the only source of pilot contamination, others sources of pilot contamination have been identified in recent times. In a practical operation regime, other sources of pilot contamination could be hardware impairments due to in-band and out-of-bound distortions that interfere with training signals [39]–[43] and non-reciprocal transceivers due to internal clock structures of the radio frequency chain. Given the importance of pilot contamination in massive MIMO in next generation communication systems, we provide in this paper an extensive survey of state-of-the-art proposals and approaches to mitigating pilot contamination from the literature.

Table I shows the summary of the notations used throughout this paper. The rest of this paper is organized as follows: in Section II, we give a short introduction to CSI, taking into consideration training methods, TDD and FDD schemes. Section III describes the multi-cell system model based on independent channel and spatial correlated channel, and the common assumptions made therein. In Section IV, the communication scheme which includes the UL and DL phases is explained. Following this, we present the sources of pilot contamination, analyze the problem of pilot contamination and explain the capacity and achievable rate in multi-cell system in Section V. In Section VI, we review the different methods for mitigation of pilot contamination. Open issues on pilot contamination in massive MIMO are listed and some broader perspectives are also discussed in Section VII, and Section VIII concludes the paper.

TABLE I
NOTATIONS

Notation	Description
$var(.)$	variance operation
I_N	size-N identity matrix
$\mathbb{E}\left\|\cdot\right\|$	expectation operator
$(\cdot)^T$	transpose operator
$tr\{\cdot\}$	trace of a matrix
$(\cdot)^H$	Hermitian transpose
$\mathbb{C}$	set of complex numbers
$\|\cdot\|$	two-norm
$*$	convolution
bold font upper case	matrices
bold font lower case	vectors

II. CHANNEL STATE INFORMATION

Acquisition of timely and accurate CSI at the BS is very important in a wireless communication system and it is perhaps the central activity of massive MIMO [44]. Good CSI helps to maximize network throughput by focusing of transmit power on the DL and collection of receive power on the UL via selective process; hence the need for effective and efficient method for channel estimation (CE). Channel state estimation error affects MIMO system performance and the effect of imperfect channel knowledge has been studied in the literature [45]–[48]. The estimation of the CSI can be driven by training sequences [49] (also known as pilot), semi-blind [50] or blind [51], [52] based techniques. Training overheads for CSI and CSI feedback contributes to increased cost of CSI estimation and decreases multiple access channel efficiency [53]. In this section, we present different training methods with a focus on training sequence and semi-blind schemes which are based on the use of pilots. An overview of the CE operations in multi-cell massive MIMO systems under the TDD and FDD schemes is also presented.

A. Training Methods

Training-based (TB) CE has been widely adopted in MIMO systems and is well covered in the literature [49], [54]–[56]. In training-based estimation, pilot sequences known to the receiver are transmitted over the channel. This known pilot sequences are then used by the receiver to build estimates of the random MIMO channel [57]. Two different training schemes are identified in [55] which are conventional time-multiplexed pilot scheme (CP) and superimposed pilot scheme (SIP). In the CP, the pilot symbols are transmitted exclusively in dedicated time slots allocated for training while in the SIP, the pilot symbols are superimposed to the data and data are transmitted in all time slots. The performance analysis based on the maximum data rate for the CP and SIP scheme using different scenarios have been discussed in [55], [56]. Several works have considered different training schemes for both the flat-fading and frequency-selective MIMO cases which include design of estimator with both low complexity with good channel tracking ability and optimal placement of pilots (see, e.g. [48], [58]–[61] and reference therein). The criteria used to analyze the performance of training-based CE can be classified into two: 1) information theoretic (mutual information and channel capacity bounds, cut-off rate) and 2) signal processing (channel mean-square error (MSE), symbol MSE, bit error rate (BER)) [55].

In large MIMO systems, a key question of interest in TB as noted in [62] is - how much time should be spent in training, for a given number of transmit and receive antennas, length of coherence time (T) and average received SNR. The trade-off between the quality of channel estimate and information throughput plays an important role in selection of optimal training-based schemes. Hence, channel accesses employed for pilot transmission and for data transmission needs to be optimized for total throughput and fairness of the system [63], [64]. In massive MIMO, the large number of antennas necessitates the use of CE that have low computational complexity and high accuracy. Where the two constraints work against each other, a good trade-off is required for high data rates and low channel estimation errors. The minimum mean square error (MMSE) and minimum variance unbiased (MVU) CE have been shown to have high computational complexity for massive MIMO systems compared to the low complexity schemes developed in [65]–[67]. The different CE schemes for conventional MIMO [68] and massive MIMO have been studied [25], [65], [67], [69]–[72] and a summary of this is presented in Table II.

B. TDD Scheme

In TDD systems, BSs and UTs share the same frequency band for transmission; hence, it is considered an efficient way to obtain CSI for fast changing channels. A distinguishing feature of TDD systems as noted in [22], [27] is the notion of reciprocity, where it is assumed that the forward channel is equal to the transpose of the reverse channel. This eliminates the need for feedback and allows for acquisition of CSI through reciprocity of wireless medium with UL training signals [73]. The use of TDD scheme makes massive MIMO system scalable in the number of service antennas to a desired extent [44], although the constraint of coherence interval needs to be considered. The communication is divided into two phases: the UL phase and DL phase. In the UL phase, UTs transmit pilot signals to the BS, the BS uses these pilot signals for CE process and to form pre-coding matrices. The produced matrices are then used to transmit pre-coded data to the UT located in each BS cell in the DL phase. The TDD protocol is illustrated in Fig. 2. It is assumed that the channel is constant during the coherence interval (T) which is T symbols.

In a multi-cell scenario, non-orthogonal pilots across neighboring cells are utilized, as orthogonal pilots would need length of least $K \times L$ symbols (K= total number of UTs in a cell and L = total number of cell in the system) owing to frequency reuse factor of one. The use of $K \times L$ symbols training sequence is not feasible in practice for multi-cell as a result of short channel coherence times due to mobility of UTs. This causes a phenomenon known as pilot contamination and it has been considered as a major impairment in the performance of massive MIMO systems [16], [74], [75]. The phenomenon introduces a finite signal to interference ratio (SIR) to the network, which in turn, causes saturation effect i.e. the system throughput does not grow with the number of BS antennas. The effects of pilot contamination is discussed in detail in Section V.

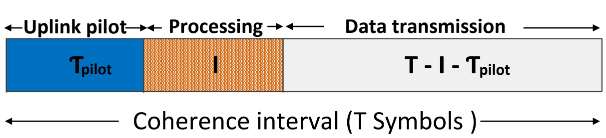


Fig. 2. TDD transmission protocol

C. FDD Scheme

The FDD scheme is briefly discussed here to highlight the cost of obtaining CSI through feedback in terms of system resources and bandwidth. Acquisition of CSI has been studied in great detail in literature under the frame work of FDD multi-user MIMO systems [76]–[78]. In FDD systems, the pre-coding in the DL and detection in the UL use different frequency bands, hence the use of feedback is required in obtaining CSI. In the process of obtaining the CSI in the DL, the BS first transmits pilot symbols to all UTs and then UTs feedback the estimated CSI (partial or complete) to the BS for the DL channels. The feedback resources used in FDD multi-user diversity system scales with the number of antennas and will thus grow large in a massive MIMO system with hundreds of antennas which leads to a loss of time-frequency resource [79]. As a result, the overhead in FDD becomes prohibitively large compared to TDD systems where the overhead scales only with the number of users. A study by [78], on feedback cost and spectral efficiency shows that data bandwidth decreases with the number of UTs. This presents a major challenge in the deployment of massive MIMO system using FDD mode and is also attracting lot of research interest because FDD system is popular among network providers [19]. Several schemes have been proposed for possible deployment of massive MIMO using FDD [16], [53], [71], [76], [79]–[82] but are not discussed here because it is outside the scope of this survey.

III. System Model

In this section we present two categories of system model for the massive MIMO system found in literature. The models from literature are based on independent channel [11], [22], [27], [32], [75], [83] and spatial correlated channel [18], [84], [85].

A. Independent channel

In this system, L number of cells are considered, where each cell contains one BS equipped with M antennas and K ($\leq M$) single-antenna UTs. The system models were represented using different symbols in reviewed papers but we have used symbols that we considered popular in the literature to allow ease of understanding.

For independent channel in [11], [22], [27], [32], [75], [83], the propagation factor from the kth UT of the jth cell to the mth antenna of the BS in the lth cell is denoted as $g_{l,k,j,m} = \sqrt{\beta_{l,k,j}} h_{l,k,j,m}$. Where $\{h_{l,k,j,m}\}$ is small scale fading factor which is independent and identically distributed (i.i.d) zero mean, circularly-symmetric complex Gaussian $\mathcal{CN}(0,1)$ random variables and $\{\beta_{l,k,j}\}$ is large scale fading coefficient, respectively [75]. See Fig. 3, the UTs from jth cell sends interfering signal to the BS in the lth cell. The small scale fading coefficients are assumed to be different for each user or different antennas while the large scale coefficients are assumed to be the same for different antennas at the BS but are user dependent. The $\{\beta_{l,k,j}\}$ accounts for path loss and shadow fading of each kth user that changes slowly and can remain constant over a coherence time interval and known prior. The channel matrix between all K UTs in the jth cell and BS in the lth cell can be expressed as:

$$\mathbf{G}_{l,j} = \mathbf{D}_{l,j}^{1/2}\mathbf{H}_{l,j} = \begin{pmatrix} g_{l,1,j,1} & \cdots & g_{l,K,j,1} \\ \vdots & \ddots & \vdots \\ g_{l,1,j,M} & \cdots & g_{l,K,j,M} \end{pmatrix}, \quad (1)$$

TABLE II
PERFORMANCE COMPARISON OF TRAINING METHODS

Reference	Pilot scheme	Channel Estimation Methods	Performance Analysis	Remarks
[65]	CP, Zhadoff Chu sequence based pilots	Improved MMSE joint channel estimation	Computational cost of improved MMSE versus conventional MMSE based on number of operations per iteration.	Improved algorithm saves 61.2% computational cost in million instructions per second (MIPS) compared to conventional MMSE
[66], [67]	CP	Low-complexity Bayesian channel estimators, coined Polynomial Expansion Channel (PEACH), Weighted-PEACH (W-PEACH) estimators, diagonalized estimator	MSE of different estimators for different polynomial order, computational complexity	PEACH estimator has the lowest computational complexity - $\mathcal{O}(LN_t^2N_r^2)^1$ compared to MMSE, MVU - $\mathcal{O}(N_t^3N_r^3)^1$ and W-PEACH - $\mathcal{O}(L^2N_t^2N_r^2+L^3)^1$
[72]	CP, Zhadoff Chu sequence based pilots	A pilot criterion was proposed, and Chu sequence was employed to design pilots with different phase shift for terminals in each cell. The best phase shift for the pilots were selected to meet the pilot criterion.	Signal to interference ratio (SIR)	Proposed pilots perform better than conventional pilots and eliminates partial interference
[69]	CP, semi-blind	LS, semi-blind maximum-a-posteriori (MAP)	Information throughput (downlink achievable rate), subspace estimation error	Semi-blind MAP estimation always outperforms the training based least-squares estimate and offers significant performance gains even for a small number of received UL data signals
[70]	CP, semi-orthogonal pilot design	Successive interference cancellation (SIC)	Mutual information throughput - UL and DL achievable rates	The achievable rate of proposed pilot design with concerned coherence interval (N_c) surpasses the conventional one at low and high SNRs no matter how long the coherence interval is

[1] N_t and N_r are number of transmit and receive antennas, respectively. L is order of polynomial expansion.

where

$$\mathbf{H}_{l,j} = \begin{pmatrix} h_{l,1,j,1} & \cdots & h_{l,K,j,1} \\ \vdots & \ddots & \vdots \\ h_{l,1,j,M} & \cdots & h_{l,K,j,M} \end{pmatrix}, \tag{2}$$

and

$$\mathbf{D}_{l,j} = \begin{pmatrix} \beta_{l,1,j} & & \\ & \ddots & \\ & & \beta_{l,K,j} \end{pmatrix}. \tag{3}$$

From the model, channel reciprocity is assumed for the forward and reverse links, i.e, the propagation factor $g_{l,k,j,m} = \sqrt{\beta_{l,k,j}}h_{l,k,j,m}$ is same for both forward and reverse links. The system equations describing the signals received at the BS and the UTs are given in Section IV. The average power at the BS is represented as p_d and the average power at each kth UT is p_u.

B. Spatial Correlated Channel

The spatial correlated channel model in [18], [84], [85], is characterized statistically in terms of the correlation between the entries of the channel matrix. In other words, spatial correlation between the antenna elements at the transmitter/receiver are factored in the the massive MIMO system where all the BS are equipped with M antennas in L number of cells. The K UTs are equipped with single antennas and pilot signal of length τ used by users in all the cells are assumed to be mutually orthogonal. Hence the intra-cell interference are considered negligible in the CE phase [84]. However, non-orthogonal pilots are reused from cell to cell, resulting in pilot contamination from $L-1$ interfering cells.

The channel vector between the lth cell user and the target BS is $\mathbf{h}_l$. For instance, if $\mathbf{h}_l$ is the desired channel, then $\mathbf{h}_l$, $l > 1$ are interference channels. All channel vectors are assumed to be $M \times 1$ complex Gaussian, undergoing correlation due to the finite multi-path angle spread at the BS side:

$$\mathbf{h}_l = \mathbf{R}_l^{1/2}\mathbf{h}_{Wl}, \;\; l = 1, 2, \ldots, L, \tag{4}$$

where $\mathbf{h}_{Wl} \backsim \mathcal{CN}(0, \mathbf{I}_M)$ is the spatially white $M \times 1$ single-input multiple-output (SIMO) channel, and $\mathcal{CN}(0, \mathbf{I}_M)$ denotes zero-mean complex Gaussian distribution with covariance matrix $\mathbf{I}_M$. The covariance matrix $\mathbf{R}_l \triangleq \mathbb{E}\{\mathbf{h}_l\mathbf{h}_l^H\}$ contains the correlations of all the elements of the channel matrix and describes the spatial statistics.

IV. COMMUNICATION SCHEME

In this section, the UL training phase and DL transmission phase are presented.

A. UL Training

The channel information is acquired using the pilot signal transmitted from the UT in the UL as described in Section III. In the UL transmission, different scenarios can be considered which are the worst-case and average-case scenario. The worst-case is when all UTs are assumed to transmit synchronously and the average-case scenario is when arbitrary

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

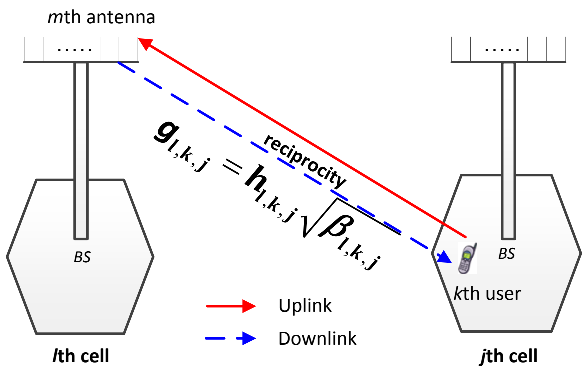


Fig. 3. System model showing the BS in lth cell and kth UT in jth cell

number of UTs at random positions transmit from different cells.

In dealing with pilot contamination, majority of research work in the literature considered the worst-case scenario and assumed that all UTs transmit synchronous pilot sequences of length τ symbols at the beginning of every coherence interval. The training vector transmitted by all cells is $\tau \times K$ orthogonal matrix $\mathbf{S}_j = (\mathbf{s}_{j1,} \ldots, \mathbf{s}_{jK})$ satisfying $\mathbf{S}_j^T \mathbf{S}_j = \tau \mathbf{I}$.

The received signal matrix at the lth BS, $\mathbf{Y}_l \in \mathbb{C}^{M \times \tau}$, can be expressed as

$$\mathbf{Y}_l = \sqrt{p_u} \sum_{j=1}^{L} \mathbf{D}_{l,j}^{1/2} \mathbf{H}_{l,j} \mathbf{S}_{\mathrm{j}}^{\mathrm{T}} + \mathbf{N}_l, \tag{5}$$

where $\mathbf{N}_l$ is the additive noise matrix with dimension ($M \times \tau$), whose elements are $\mathcal{CN}(0,1)$ random variables, p_u is the average transmit power at each user on the UL, $\mathbf{S}_j = [\mathbf{s}_{j,1}, \mathbf{s}_{j,2}, \ldots, \mathbf{s}_{\mathrm{j,K}}]$ is a $\tau \times K$ matrix, $\mathbf{D}_{l,j}$ and $\mathbf{H}_{l,j}$ are defined in (3) and (2) respectively.

B. DL Transmission

In the DL transmission, considering the BS of the lth cell, the information to be transmitted to K UTs in the same cell is expressed as $\mathbf{q}_l = [q_{l1} q_{l2}, \ldots, q_{lK}]^T$ and $M \times K$ pre-coding matrix used at the BS be $\mathbf{A}_l = f(\hat{\mathbf{H}}_l)$. Where the function $f(.)$ corresponds to the specific pre-coding method performed at the BS [27] and $\hat{\mathbf{H}}_l$ is the estimated channel. Hence, $\mathbf{A}_l \mathbf{q}_l$ signal vector is transmitted by the BS of the UTs in lth cell. To normalize the transmit power and satisfy average power constraint at the BS the transmission symbols and pre-coding methods are considered such that $\mathbb{E}[\mathbf{q}_l \mathbf{q}_l^H] = \mathbf{I}$, and $tr\{\mathbf{A}_l^H \mathbf{A}_l\} = 1$.

Consequently, the signals received in the jth cell which comprises of noisy $K \times 1$ signal vectors is expressed as

$$\mathbf{z}_j = \sum_{l=1}^{L} \sqrt{p_d} \mathbf{D}_{j,l}^{\frac{1}{2}} \mathbf{H}_{j,l} \mathbf{A}_l \mathbf{q}_l + \mathbf{n}_j, \tag{6}$$

where $\mathbf{n}_{j=}[n_{j1} n_{j2} \ldots n_{jK}]^T$ is the $K \times 1$ additive noise vector whose elements are $\mathcal{CN}(0,1)$ random variables. The signal received as expressed in [27] at the kth UT in jth cell is given as:

$$z_{j,k} = \sum_{l=1}^{L} \sum_{i=1}^{K} \sqrt{p_d \beta_{j,k,l}} [h_{j,k,l,1} h_{j,k,l,2} \ldots h_{j,k,l,M}] \mathbf{a}_{l,i} q_{l,i} + n_{j,k}, \tag{7}$$

where $\mathbf{a}_{l,i}$ is the ith column of the pre-coding matrix $\mathbf{A}_l$, $n_{j,k}$ is the kth element of $\mathbf{n}_j$ and p_d is the average transmit power form the BS on the DL.

V. Sources of Pilot Contamination, Effects and Achievable Rates

In this section, we provide the possible sources of pilot contamination in massive MIMO systems. Achievable rates using linear estimators in the presence of pilot contamination from different literature are presented.

A. Sources of Pilot Contamination

The possible causes of pilot contamination and basic theory is highlighted in this subsection. First, we consider the pilot contamination from the reuse of non-orthogonal training sequences across cells in a multi-cell massive MIMO systems due to limited coherence time, followed by the contamination of pilots resulting from hardware impairments and non-reciprocal transceivers.

1) Non-orthogonal Pilot Schemes : In a multi-cell system where the same frequency is shared by all L cells, the intra-cell interference is considered negligible since the pilots are assumed to be mutually orthogonal. When frequency reuse factor of 1 is used, the pilot signals are affected by inter-cell interference leading to pilot contamination from adjacent cells brought into the system. The UL transmission when same pilots are used in all cells at the lth BS, $\mathbf{Y}_l \in \mathbb{C}^{M \times \tau}$ similar to expression in (5) is

$$\mathbf{Y}_l = \sqrt{p_u} \sum_{j=1}^{L} \mathbf{G}_{l,j} \mathbf{S}_j^T + \mathbf{N}_l, \quad j = 1, \ldots, L, \tag{8}$$

where $\mathbf{G}_{l,j} \in \mathbb{C}^{M \times K}$ defined in (1), is the channel matrix from all K users in the jth cell to the lth BS. $\mathbf{N}_l \in \mathbb{C}^{M \times \tau}$ is the noise matrix at the lth BS during the pilot transmission phase. Each BS correlates its received pilot signals with its own orthogonal pilot signals while all terminals in the other cells contribute to the pilot contamination [11]. The resulting estimated channel matrix at the lth BS is

$$\hat{\mathbf{G}}_{l,l} = \sqrt{p_u} \mathbf{G}_{l,l} + \sqrt{p_u} \sum_{j \neq l} \mathbf{G}_{l,j} \mathbf{S}_j^T \mathbf{S}_l^* + \mathbf{N}_l. \tag{9}$$

2) Hardware Impairment : The impact of transceiver hardware impairments on massive MIMO system has been studied in [39], [41], [42]. The hardware components in radio frequency chain are prone to impairments such as phase noise, amplifier non-linearity, quadrature imbalance (I/Q) and quantization errors [86], [87]. Contrary to the view of majority of papers that assumed ideal transceiver hardware, Bjornson *et al.* [39] considered how hardware impairments create a mismatch between the intended transmit signal and actual signal generated, and distortions of received signal in the reception process. This impairment has been shown to affect the accuracy of the CE which can invariably lead to pilot contamination as well as impact the performance of massive MIMO system. To overcome the challenge of this impairment, non-ideal behavior of each component has been modeled so as to design compensation algorithm. However, modeling of aggregate residual transceiver impairments is considered more important to system performance rather than the modeling of individual components [41], [86], [88]. For ideal system model where the DL channel is used for data transmission and pilot-based CE, the received signal at the user is expressed as:

$$z = \mathbf{g}^T \mathbf{s} + n, \tag{10}$$

where $\mathbf{s} \in \mathbb{C}^{M\times 1}$ is either a deterministic pilot signal (during CE) or stochastic zero-mean data signal. n is the additive term $n = n_{noise} + n_{interf}$ is an ergodic stochastic process that consists of independent receiver noise $n_{noise} \sim \mathcal{CN}(0, \sigma^2_{UT})$ and interference n_{interf} from simultaneous transmission by other UTs. The covariance matrix of $\mathbf{s}$ is denoted by $\mathbf{W} = \mathbb{E}\{\mathbf{s}\mathbf{s}^H\}$ and average power is $p^{BS} = tr(\mathbf{W})$. The non-ideal hardware DL system model taking into consideration the distortion noise is expressed in (11) and (13) [39]:

$$z = \mathbf{g}^T(\mathbf{s} + \boldsymbol{\eta}^{BS}_t) + \eta^{UT}_r + n, \tag{11}$$

where $\boldsymbol{\eta}^{BS}_t \in \mathbb{C}^{M\times 1}$ and $\eta^{UT}_r \in \mathbb{C}$ are additive distortion noise terms which are stochastic processes that describe the residual transceiver impairments of the transmitter hardware at the BS and receiver hardware at UT, respectively. The distortion noise are assumed to be independent of the transmitted signal $\mathbf{s}$ but dependent on the channel realization $\mathcal{H}$. $\mathcal{H}$ denotes the set of channel realizations for all useful and interfering channels (i.e, $\mathbf{g} \in \mathcal{H}$). The conditional distributions for the given channel realization of $\mathcal{H}$ are $\boldsymbol{\eta}^{BS}_t \sim \mathcal{CN}(0, \boldsymbol{\Upsilon}^{BS}_t)$ and $\eta^{UT}_r \sim \mathcal{CN}(0, v^{UT}_r)$. The noise occurred at an antenna is proportional to the signal power at the antenna, thus $\boldsymbol{\Upsilon}^{BS}_t$ and v^{UT}_r are expressed as:

$$\boldsymbol{\Upsilon}^{BS}_t = k^{BS}_t diag(W_{11}, \ldots, W_{MM}),\ v^{UT}_r = k^{UT}_r \mathbf{g}^T \mathbf{W} \mathbf{g}^*, \tag{12}$$

where W_{ii} is the ith diagonal element of $\mathbf{W}$ and k^{BS}_t, $k^{UT}_r \geq 0$ are the proportionality coefficients which characterize the levels of impairments due to quantization errors in automatic-gain controlled analog-to-digital conversion, inter-carrier interference induced by phase noise, leakage from mirror subcarrier under I/Q imbalance, and amplitude non-linearity in the power amplifier [39].

Similarly, the UL non-ideal system model taking into consideration distortion noise for received signal $\mathbf{y} \in \mathbb{C}^M$ at the BS is expressed as:

$$\mathbf{y} = \mathbf{g}(d + \eta^{UT}_t) + \boldsymbol{\eta}^{BS}_r + \mathbf{v}, \tag{13}$$

where $d \in \mathbb{C}$ is either a deterministic pilot signal used for CE or a stochastic data signal. The average power is $p^{UT} = \mathbb{E}\{|d|^2\}$. The additive noise $\mathbf{v} = \mathbf{v}_{noise} + \mathbf{v}_{interf} \in \mathbb{C}^{M\times 1}$ is an ergodic process that consists of independent receiver noise $\mathbf{v}_{noise} \sim \mathcal{CN}(\mathbf{0}, \sigma^2_{BS}\mathbf{I})$ and potential interference $\mathbf{v}_{interf}$ from other simultaneous transmission. $\eta^{UT}_t \in \mathbb{C}$ and $\boldsymbol{\eta}^{BS}_r \in \mathbb{C}^{M\times 1}$ are additive distortion noise terms which are stochastic processes that describe the residual transceiver impairments of the transmitter hardware at the UT and receiver hardware at BS, respectively. According to Bjornson *et al.* the ergodic stochastic processes are independent of d but dependent on channel realization $\mathcal{H}$, where the conditional distribution for a given $\mathcal{H}$ are $\eta^{UT}_t \sim \mathcal{CN}(0, v^{UT}_t)$ and $\boldsymbol{\eta}^{BS}_r \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Upsilon}^{BS}_r)$. The model of the conditional covariance matrices are modeled as:

$$v^{UT}_t = k^{UT}_t p^{UT},\ \boldsymbol{\Upsilon}^{BS}_r = k^{BS}_r p^{UT} diag(|g_1|^2, \ldots, |g_M|^2), \tag{14}$$

where the hardware quality are characterized at the BS by k^{BS}_t, k^{BS}_r and at the UT by k^{UT}_t, k^{UT}_r. Thus far, the study from [39] shows that only the hardware impairment from UT limits the capacity in massive MIMO systems as M grows large and the hardware impairment BS antenna arrays are negligible. Hence more research work is needed in this area through practical investigation on the impact of hardware impairment to pilot and how it affects massive MIMO systems.

3) Non-Reciprocal Transceivers : In the TDD systems, the physical forward and the backward channel are considered reciprocal since they operate on the same carrier frequency. The models of point-to-point case TDD involving two devices adapted from [89], [90] is shown in Fig. 4. In ideal case, the power amplifiers (T1 and T2), low noise amplifiers (R1 and R2) and the effective electromagnetic channel (C(t)) are considered identical. In non-ideal case, it has been shown in [89], [91] that the presence of residual offset frequency does have a big impact on the exploitation of channel reciprocity. The smallest offset of a few Hz will accumulate and make the UL and DL channels non-reciprocal within a few seconds. The impulse response is modeled from device 1 to device 2 as [90]:

$$G(t, \iota) = R2(\iota) * C(t, \iota) * T1(\iota), \tag{15}$$

and in the reverse direction from device 2 to device 1 as:

$$H(t, \iota) = R1(\iota) * C(t, \iota) * T2. \tag{16}$$

The relationship between G and H is modeled as:

$$G(t, \iota) = H(t, \iota) * R(\iota). \tag{17}$$

Robust calibration techniques are required to provide accurate estimation of $R(\iota)$ from channel measurements where variable ι is the delay domain. Lack of accurate reciprocity-based CSI calibration scheme available at the transmitter in the presence of the residual offset frequency can be a source of pilot contamination during CE.

B. Impact of Pilot Contamination

The extent to which pilot contamination impacts system performance under different scenarios have been studied in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

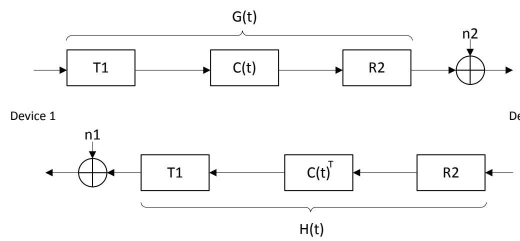


Fig. 4. Reciprocity model for point-to-point case adopted from [90]

the literature [11], [13], [26], [27], [32], [74], [75], [92], [93]. In analyzing the effects of pilot contamination, the worst-case scenario is assumed, where all transmissions and receptions are synchronous with a uniform array of UTs at the BS [27]. One of the key points being considered is that as the number of BS antennas M grows without limit, the effects of additive receiver noise, fast-fading and intra-cell interference disappears and the only challenge is the inter-cell interference from transmissions which are associated with reuse of the same pilot sequence [11]. The effects of pilot contamination are analyzed by first formulating expressions of signal-to-interference ratio (SIR). This SIR are then translated into achievable rates (throughput per cell and mean throughput per terminal) based on different modeling parameters such as log-normal shadow fading standard deviation, path loss exponent, radius of disk from which terminals are excluded to the radius of the cell, frequency reuse factors and pilot overhead [11]. The authors of [74] studied the rate of convergence of signal-to-interference-plus-noise ratio (SINR) in the presence of pilot contamination using a simplified system model. It was shown that in the absence of pilot contamination, SINR increases linearly with M and does not saturate as M increases to infinity. However, in the presence of pilot contamination SINR saturates due to corruption of channel estimate by interfering user. In [27], [75], the effect of pilot contamination was examined using sum-rate of UL transmission versus the cross gain at fixed SNR for different number of BS antennas. It was observed that regardless of the increasing number of BS antennas, the effect of pilot contamination is significant when the value of cross gain (across cell) due to interference is close to direct gain (within cell). In [13], it was shown that system performance is significantly degraded by pilot contamination when the inter-cell interference factor β (which accounts for path loss and shadow fading) increases. As pilot contamination increases with the increase of β, and considering the same average transmitted power from the UTs, there is a considerable reduction in the spectral efficiency and energy efficiency. The authors of [92] examined the asymptotic SINR of a matched filter with perfect estimate and with that of corrupted estimate for case of equal receive power. It was shown that using a pilot interference based estimate for matched filter causes a loss in SINR compared to perfect estimate in the matched filter. The effect of pilot contamination has also been illustrated using different frequency or pilot reuse factors in multi-cell systems. In [32], using frequency reuse factor of 1, the sum rate (bits/s/Hz) is much lower compared to higher reuse factors due to aggressive interference. An illustration of using different pilot reuse factor of 1, 2, 4 can be seen in Fig. 5 and Fig. 6 using the equations (55) and (20) from [26] for area throughput and energy efficiency (EE), respectively. In Fig. 5, the use of pilot reuse factor of 1 shows a significant loss in area throughput compared to reuse factors of 2 and 4 as a result of interference from all adjacent cells. Also, a reduction in the achievable EE can be seen in Fig. 6 for multi-cell scenario with imperfect CSI when pilot reuse factor of 1 is used compared to reuse factors of 2 and 4. This emphasizes the need to actively mitigate pilot contamination in multi-cell systems. The derivation of lower bound achievable rates in the presence of pilot contamination is discussed in Subsection V-C.

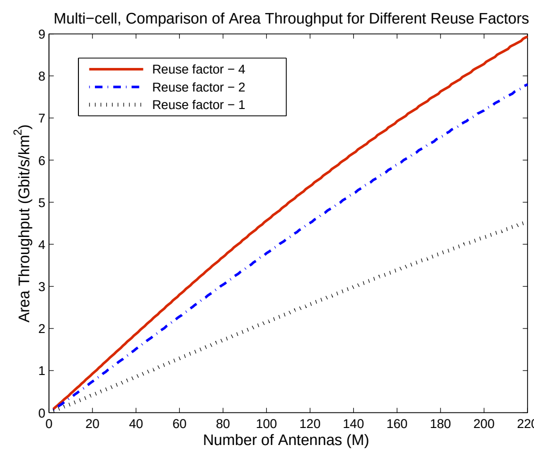


Fig. 5. Effect of pilot contamination on area throughput using different pilot reuse factors

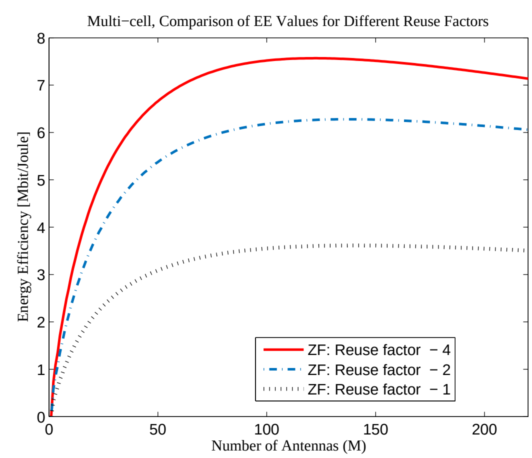


Fig. 6. Effect of pilot contamination on energy efficiency using different pilot reuse factors.

C. Capacity and Achievable Rate in TDD system

The reverse link and forward link achievable rate for non-cooperative multi-cell system was formulated in [11] based on the assumption that the terminals transmit Gaussian message-bearing symbols where the net sum throughput per cell is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

equal to the sum of net throughput per terminal in units of bits/sec/cell. The net throughput formula accounts for the total bandwidth, frequency reuse factor, the pilot overhead (the ratio of the time spent sending data to the total slot-length), and the overhead of the cyclic prefix. The reverse link capacity is given as:

$$C_{rsum} = \sum_{k=1}^{K} C_{rk}, \tag{18}$$

where C_{rk} is net throughput per terminal (bits/sec/terminal) and it is given as:

$$C_{rk} = \left(\frac{B}{\alpha}\right)\left(\frac{T_{slot} - T_{pilot}}{T_{slot}}\right)\left(\frac{T_u}{T_s}\right) log_2(1 + SIR_{rk}), \tag{19}$$

where B is the total bandwidth in Hz, α is the frequency reuse factor, T_{slot} is the slot length in symbols, T_{pilot} is the number of symbols used in transmitting reverse-link pilots, T_u is the useful symbol duration, and T_s is the OFDM symbol interval, SIR_{rk} is the signal-to-interference ratio.

On the other hand, the sum net throughput for the forward link capacity is given as:

$$C_{fsum} = \sum_{k=1}^{K} C_{fk}, \tag{20}$$

where C_{fk} is the net capacity per terminal (bits/sec/terminal) and it is given as:

$$C_{fk} = \left(\frac{B}{\alpha}\right)\left(\frac{T_{slot} - T_{pilot}}{T_{slot}}\right)\left(\frac{T_u}{T_s}\right) log_2(1 + SIR_{fk}). \tag{21}$$

The derived capacity in (18), (20) are random quantities and they depend on the terminal positions and shadow fading. The SIR_{fk} and SIR_{rk} play a major role in determining the achievable rates in the forward and reverse links, respectively. Different processing schemes, such as the maximum ratio combining (MRC), zero forcing (ZF) and MMSE have been used to determine the effective SIR in different scenarios, such as when the number of BS antennas tends towards infinity and also when $M \geqslant K$ under perfect CSI estimate and imperfect CSI estimate. The expression $\left(\frac{B}{\alpha}\right)\left(\frac{T_{slot}-T_{pilot}}{T_{slot}}\right)\left(\frac{T_u}{T_s}\right)$ from (19) and (21) is considered the pre-log factor. The expression $log_2(1 + SIR_{rk})$ is the achievable rate and denoted as R_{jk}. For the kth user in the jth cell the received signal can be expressed as:

$$z_{jk} = \sum_{l=1}^{M}\sum_{i=1}^{K} g_{jl}^{ki} q_{li} + x_{jk} = g_{jj}^{kk} q_{jk} + \sum_{(l,i)\neq(j,k)} g_{jl}^{ki} q_{li} + n_{jk}, \tag{22}$$

where the term $g_{jl}^{ki} = \sqrt{p_d \beta_{j,k,l}}[h_{j,k,l,1} h_{j,k,l,2} \dots h_{j,k,l,M}]\mathbf{a}_{l,i}$ and it depends on the system model, with or without BS cooperation. g_{jj}^{kk} is a complex term that is not known at the user, hence the received symbol expression is rewritten as:

$$z_{jk} = \mathbb{E}[g_{jj}^{kk}] q_{jk} + n_{jk}^{'}, \tag{23}$$

where $x_{jk}^{'}$ is the effective additive noise term and variance is given by $var\{n_{jk}^{'}\} = var\{g_{jj}^{kk}\} + \sum_{(l,i)\neq(j,k)} \mathbb{E}\left[\left|g_{jl}^{ki}\right|^2\right] + 1$. From (23), the rate R_{jk} suitable for any linear pre-coder is expressed in [62] and in [27] using methods suggested in [22] as:

$$R_{jk} = log_2\left(1 + \frac{\left|\mathbb{E}\left[g_{jj}^{kk}\right]\right|^2}{1 + var\left\{g_{jj}^{kk}\right\} + \sum_{(l,i)\neq(j,k)} \mathbb{E}\left[\left|g_{jl}^{ki}\right|^2\right]}\right). \tag{24}$$

From (24) different linear pre-coding methods can be applied by changing the variance and the expectation terms. In [11], using MRC, the achievable rates for a system with unlimited number of BS antennas for UL rate and DL rate were derived for simple cases where all pilots are being the same in all cells as shown in (25). That is:

$$SIR_{rk/fk} = \frac{\beta_{jkj}^2}{\sum_{l\neq j}^{L} \beta_{jkl}^2}, \tag{25}$$

where $SIR_{rk/fk}$ represents the SIR for reverse and forward links and β is the interference factor that accounts for path loss and shadow fading, which is defined as:

$$\beta_{jkl} = \frac{z_{jkl}}{r_{jkl}^{\gamma}}, \tag{26}$$

where r_{jkl} is the distance between the kth terminal in the lth cell and the BS in the jth cell, γ is the decay exponent and z_{jkl} is a log-normal random variable, i.e the quantity $10log_{10}(z_{jkl})$ is distributed zero-mean Gaussian with standard deviation of σ_{shad} [11]. However, (25) can be modified to account for normalized factor in forward link pre-coding as:

$$SIR_{fk} = \frac{\beta_{jkj}^2/\alpha_{kl}^2}{\sum_{l\neq j}^{L} \beta_{jkl}^2/\alpha_{kl}^2}, \tag{27}$$

where $\alpha_{kl}^2 = \sum_{l=1}^{L} \beta_{jkl} + \frac{1}{\tau p_u}$.

The set of lower bounds achievable rates of the UL transmission for the kth user when BS antennas is much greater than UTs, (i.e, $M \gg K$) in single- and multi-cell imperfect scenario, was derived in [13] and summarized in Table III. Noted from (25) and (27) , it can be seen that β and L play important factors in spectral efficiency. When the number of cell is one i.e L = 1, or β is reduced to zero, the equation of multi-cell reduces to the single-cell multi-user system.

VI. METHODS FOR MITIGATION OF PILOT CONTAMINATION

In this section, we review existing methods proposed to eliminate or reduce the effects of pilot contamination in multi-cell TDD systems considering full pilot reuse factor of 1. The proposed methods have been classified into two categories namely, pilot-based estimation approach and subspace-based estimation approach. In the pilot-based approach, channels of UTs are estimated using orthogonal pilots within the cell and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

TABLE III
SET OF COMPUTABLE ACHIEVABLE RATES FOR UL TRANSMISSION FOR THE kth USER IN MULTI-USER SYSTEM

Reference	Linear method	Single-cell imperfect CSI	Multi-cell imperfect CSI
[13]	MRC	$\frac{T-\tau}{T} K\ log_2\left(1+\frac{\tau(M-1)p_u^2}{\tau(K-1)p_u^2+(K+\tau)p_u+1}\right)$	$\frac{T-\tau}{T} Klog_2\left(1+\frac{\tau(M-1)p_u^2}{\tau(K\overline{L}^2-1+\beta(\bar{L}-1)(M-2))p_u^2+\bar{L}(K+\tau)p_u+1}\right)$
[13]	ZF	$\frac{T-\tau}{T} K\ log_2\left(1+\frac{\tau(M-K)p_u^2}{(K+\tau)p_u+1}\right)$	$\frac{T-\tau}{T} Klog_2\left(1+\frac{\tau(M-K)p_u^2}{\tau K(\overline{L}^2-\bar{L\beta}+\beta-1)p_u^2+\bar{L}(K+\tau)p_u+1}\right)$

non-orthogonal pilots across the cells while in the subspace-based estimation approach, the channels of UTs are estimated with or without limited pilots. The proposed methods under the two categories are discussed as follows.

A. Pilot-Based Estimation Approach

A time-shifted protocol for pilot transmission was considered in [38], [94] to reduce pilot contamination in multi-user TDD systems. The transmission of pilot signals in each cell is done by shifting the pilot locations in frames so that users in different cells transmit at non-overlapping times as shown in Fig. 7. It was shown that pilot contamination can be eliminated using the proposed scheme as long as pilots do not overlap in time. The use of power allocation algorithms in combination with the time-shifted protocol in [38] is shown to provide significant gains. Although the method looks promising, a major challenge in practice will be the control mechanism needed to dynamically synchronize the pilots across several cells so that they do not overlap. It is important to note that due to the emergence of multi-tier heterogeneous cellular networks and dynamic placement of small cells, there will always be overlap in time and frequency somewhere in the network. The authors in [84], [95] proposed a covariance aided CE method by exploiting the covariance information of both desired and interfering user channels. It was shown that in the ideal case where the desired and the interference covariance span distinct subspaces, the pilot contamination effect tends to vanish in the large antenna array case. As a result, users with mutually non-overlapping angle of arrival (AoA) hardly contaminate each other. Based on the finding, a coordinated pilot assignment strategy was proposed based on assigning carefully selected groups of users to identical pilot sequences. Similar approach was used in [85] for a cognitive massive MIMO system. Although this method shows a significant reduction in inter-cell interference and a corresponding increase in UL and DL SINRs, in practice it might be difficult to implement because it requires second order statistics of all the UL channels [96].

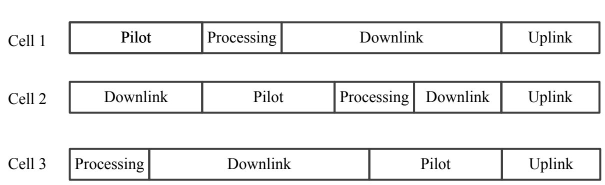


Fig. 7. Time shifted pilot scheme adapted from [38]

Wang *et al.* [97] proposed a spatial domain based method with the main idea that spatial and/or temporal characteristics of channel coefficient of different UTs are distinguishable - for example AoA or direction of arrival (DoA). In addition, the channel coefficient of strongest UL path is chosen as the DL beamformer. Based on similar assumptions made in [95] that the AoA of UTs are non-overlapping, an angular tunable predetermined scheme or offline generated codebook is used to match UL path and then used as DL beamvector with the aim of avoiding leaked signals to UTs in adjacent cells. However, in this method the coherence interval needs to be considered while searching for the optimal steering vector to be used for DL beamforming.

As a result of constraints from [95], Wang *et al.* [96] proposed a pilot contamination elimination scheme, which relies on two processing stages namely DL training and scheduled UL training. In the DL stage, UT supported by each BS estimates their specific DL frequency-domain channel transfer functions (DL-FDCHTFs) from the DL pilots of the BS. In the scheduled UL training stage, from each cell at a time, the UTs use the estimated DL-FDCHTFs to pre-distort their UL pilot symbols in which the uncontaminated DL-FDCHTFs are 'encapsulated' in the UL pilot symbols for exploitation at the BS. Thereafter, the BS extracts all the DL-FDCHTFs of its UTs from the received UL signals by eliminating UL pilot signals of UTs from all other cells, hence, eliminating pilot contamination. A major drawback of this scheme, as noted, is the cost of overhead used in training, which in reality can increase infinitely. In [27], the authors developed a multi-cell MMSE-based pre-coding scheme, which accounts for the training sequence allocated to all UTs in order to mitigate pilot contamination problem. As a result, pre-coding matrix at each BS is designed to minimize the sum of the mean-square error of signals received at the UTS in the same cell and the mean-square interference occurred at the UTs in other cells. It was shown that the proposed method offers significant performance gains and reduces the inter-cell and intra-cell interference compared to conventional single-cell pre-coding method. However, the proposed method assumes all the UTs are the same without differentiating them based on channels.

Ashikhmin and Marzetta [98] proposed a pilot contamination pre-coding (PCP) method, which involves limited collaboration between BSs. In the PCP method, the lth BS shares the slow-fading coefficient estimate with the other BS or to a network hub, which computes the PCP pre-coding matrices. The computed lth pre-coded matrices are then forwarded to each corresponding BS for computation of the transmitted signal vectors through its M antennas. This process is performed in the UL and DL. The effectiveness of this method lies in the accuracy of the shared information from each BS and the computation of PCP by the network hub. Further work on the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

PCP method [98] has been extended in [31] by proposing an outer multi-cellular pre-coding called large-scale fading pre-coding (LSFP) and large scale fading decoding (LSFD) in the regime of a finite number of BS antennas. In [99], the proposed methods are designed to maximize the minimum rate with individual BS power constraints which delivers significant improvement on the 5% outage rate compared to existing methods in [31]. The summary of the pilot-based methods and assumptions are shown in Table IV.

B. Subspace-Based Estimation Approach

The use of subspace-based CE techniques has been studied in the literature and it is seen as a promising approach for increased spectral efficiency because it requires no or very limited number of pilot symbols for operation [100]. In this approach, signal properties, such as finite alphabet structure, fixed symbol rate, constant modulus, independence, and higher order statistical properties can be used for CE. This approach has been extended to CE in multi-cell TDD systems with the aim of eliminating pilot contamination. CSI is obtained in [101] by applying subspace estimation technique using eigenvalue decomposition (EVD) on the covariance matrix of the received samples but up to a scalar ambiguity. To overcome this ambiguity, short training orthogonal pilots were introduced in all the L cells. The EVD-based method is prone to error due to the assumption that channel vectors between the users and the BS become pair-wisely orthogonal when the number of BS antennas M tends towards infinity. However, in practice M is large but finite. To reduce these errors, the EVD algorithm is combined with the iterative least-square with projection algorithm of [102]. Results obtained shows that the EVD method is not affected by pilot contamination and performs better than conventional pilot-based techniques but its accuracy depends on large number of BS antennas and increased sampling data within the coherence time.

Further studies by [103]–[106] proposed a blind method for CE in a cellular systems with power control and power-controlled hand-off. The main idea is to find the singular value decomposition of the received signal matrix and to determine which system parameters in the subspace of the signal of interest can be identified blindly using approximate analysis from random matrix theory. Results from the algorithm show that it is sufficient to know the subspace which the channel vectors of interest span, in order to acquire accurate channel estimates for the projected channel. However, the limitation of this approach in practice is that the assumption that all desired channels are stronger than all interfering channels does not always hold. To overcome this limitation, Neumann *et al.* [107] proposed a maximum a-posteriori (MAP) criterion for subspace CE. The MAP method is shown to be more robust and offers better performance than the blind method proposed in [103]–[106] but with increased complexity.

Diagonal jacket-based estimation method with iterative least-square projection was proposed in [108] for fast CE and reduction of pilot contamination problem. The BS correlates the received pilot transmissions which are corrupted by pilot transmissions from other cells to produce its channel estimates. It is shown that as the geometric attenuation from neighboring cells increases, the system performance of conventional pilot based system degrades due to pilot contamination whereas the diagonal jacket matrix is not affected. The summary of the subspace-based methods and assumptions are shown in Table V.

VII. Survey Summary

In this section, we summarize the study on pilot contamination analysis in massive MIMO TDD systems, common techniques in elimination of pilot contamination, highlight the open issues and possible future direction of research in pilot contamination in massive MIMO systems. In addition, we also discuss briefly the effect of pilot contamination in other related technologies and identify research opportunities.

A. Pilot Contamination Analysis

The use of TDD scheme in massive MIMO is seen as a better option when channel reciprocity is considered for CE. However, the major constraint is coherence interval. The use of non-orthogonal pilot signals in multi-cell introduces the problem of pilot contamination. The effect of pilot contamination is shown to be profound in multi-cell scenario with high frequency reuse factor of 1. Some researchers advocate the use of fractional frequency reuse to mitigate the effect of pilot contamination. However, several other researchers have proposed different techniques in reducing the effects of pilot contamination for a reuse factor of 1. Majority of the reviewed papers consider the worst-case scenario and uniform distribution when analyzing the effects of pilot contamination. Achievable lower bound rates have been derived based on perfect and imperfect CE and for uniformly distributed UTs in L number of cells with finite number of BS antennas. Other possible causes of pilot contamination, such as hardware impairment and non-reciprocal are also pointed out in existing literature.

B. Common Techniques

Two mitigation techniques for pilot contamination are identified, which are pilot-based and subspace-based approaches. In the pilot-based approach, coordination or protocol techniques, such as time-shifted protocols for pilot transmission [38], [94], and DL and UL protocol for pilot transmission [96] are used. Other pilot-based approaches include AoA technique, which exploits the covariance information of UTs for CE [84], [95], [97] and pre-coding technique, used in [27], [98] to minimize the interference between cells. In the subspace-based approaches, blind and semi-blind approaches, with single value decomposition of channels combined with iterative least-square algorithm [101], [103]–[106] and MAP [109] were proposed in the literature.

C. Open Issues

So far, we have discussed the effect of pilot contamination and reviewed the proposed techniques in mitigating the problem of pilot contamination. In the analysis of pilot

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

TABLE IV
SUMMARY OF PROPOSALS ON PILOT AIDED ESTIMATION APPROACH

Reference	Proposed methods	Topology/scenario	Assumptions
[38], [94]	Time shifted pilot scheme with non-overlapping pilots across each cell	Hexagonal cells with radius 1.6km, UTs uniformly distributed in each cell 100 meters away from BS.	UTs in one cell transmit pilot sequences simultaneously while UTs in other cells receive DL data
[84], [85], [95]	Covariance-aided channel estimation method with pilot assignment strategy	Hexagonal cells with radius 1km, UTs distributed uniformly at the cell-edge 800 meters away from BS	Individual covariance matrices can be estimated separately, multipath AoA of desired UTs do not overlap with that of interfering UTs
[96]	DL training and scheduled UL training scheme	Hexagonal cells L=7, 4 UTs per cell and each BS with 50 antennas.	The pilot signals used by UTs in a given cell are mutually orthogonal
[27]	Multi-cell MMSE-based pre-coding method	Hexagonal cells L = 4, K = 2 users per cell, M = 8 antennas at BS	The propagation factor $\{\beta_{jkl}\}$ is non-negative and constant and known at the BS and UT while $\{h_{j,k,l,m}\}$ is random variable and unknown
[98]	Pilot contamination pre-coding (PCP) with limited collaboration between BS	Hexagonal L- cells, with large K - UTs, M antennas at BS, network hub that computes the PCP pre-coding matrices	All signals transmitted to the kth UT of the lth cell are accessible to BS across the entire network, slow-fading coefficients $\{\beta_{jkl}\}$ can be accurately estimated and made available to all BS or alternatively to network hub
[97]	Spatial domain based method	Hexagonal cells L = 7 with radius 1km, UTs distributed uniformly at the cell-edge 800 meters away from BS	DoA, AoA of UTs are non-overlapping

TABLE V
SUMMARY OF PROPOSALS ON BLIND BASED ESTIMATION APPROACH

Reference	Proposal	Topology	Assumptions
[101]	EVD-based Method for channel estimation with iterative least-square with projection (ILSP)	Hexagonal cells L = 3, each containing 3 UTs.	The channel vectors from different users are orthogonal
[103]–[106]	SVD-based method - subspace projection	Different number of cells (L), coherence time (C), receive antenna (R) and transmit UT antennas are used across the papers.	The signal subspace of each cell is orthogonal to the interference subspace when M grows large
[108]	Diagonal Jacket Matrix with iterative least-square with projection ILSP	Hexagonal cells L = 5, UTs (K=2) per cell and BS antennas (M= 10)	Propagation matrix between the interfering cell and kth UT in target cell is i.i.d
[109]	Maximum a-posteriori estimation (MAP)	Hexagonal cells L = 7, BS antennas (M = 150, 500) and UTs (K=2,3) per cell	Channel coefficient are mutually independent for different UTs

contamination from reviewed literature, several assumptions are made which are shown in Table IV and Table V. These assumptions enable simplified models but may not accurately model the impact of pilot contamination in real life scenario. A good number of works have utilized certain special properties of the channels, examples of which are, the channel covariance matrices having a certain structure such as non-overlapping AoA and DoA [84], [95], [97]. Such special properties can lead to good performance in an ideal system where it is assumed that the system exhibits the same channel structure. However, in reality the channel structure may appear different. For practical systems, mitigation techniques that take into consideration the dynamic properties of the channel, need to be exploited and not just limited to certain structures. Moreover, the amount of pilot symbols for channel estimation is limited by the coherence time of the channel and it becomes very challenging to obtain good CSI quality when BSs are equipped with massive MIMO antenna arrays. This poses a practical challenge to cooperative and coordinated schemes due to backhaul latency. Hence, the performance of these schemes can be very sensitive to CSI errors due to outdatedness.

1) Training Overhead : The trade-off between the number of pilot signals and spectral efficiency in any proposed method is crucial in massive MIMO systems. In the pilot based methods, the use of pilot signals for CE could result in waste of channel bandwidth due to the duration occupied by training symbols [45]. As an example, using techniques suggested in [96] will lead to high training overhead. Although the formula (18) and (20) helps determine the net throughput for the system, however, there is a need to account for the percentage of bandwidth consumed in existing techniques used for pilot contamination. The implementation of subspace-based approach suggests a better spectral efficiency since less or no training overheads are incurred. However, there is a need to determine if the subspace-based approach can achieve sufficient CSI quality required for same target SINRs in the pilot-based approach. Study from [110] has shown that the use of EVD based CE under the subspace-based approach is prone to error from imperfect orthogonalized channel resulting from limited number of BS antennas. Moreover, the subspace methods require additional techniques or information to identify which eigenvector corresponds to which UT and the assumption that all desired channels are stronger than all interfering channels does not always hold in practical systems.

2) Deployment Scenario : The study and effect of pilot contamination has been examined using the worst-case scenario in which UTs are considered to be symmetrically distributed and uniform in all cells. In addition, uncorrelated Rayleigh fading channel model has been mostly used for performance evaluation especially in deriving achievable rates of the massive MIMO systems in the presence of pilot contamination. Using such assumptions leads to model simplicity and formation of simplified closed-form analysis for the achievable rates which is probably optimistic. Findings from measurement results in [12] have shown that channels in massive MIMO are not i.i.d and thus there is a performance loss compared to ideal channels. Also majority of the work assumed that the large-scale fading coefficients do not depend on the antenna index m of a given BS because typically, the distance between a UT and BS is significantly larger than the distance between BS antennas [31].

There is a need to investigate the effect of pilot contamination using a more realistic channel model for large MIMO systems by considering statistical channel properties, in particular the spatial-temporal-frequency correlation properties and the role of large-scale fading. Recent works have considered the role of large-scale fading in massive MIMO systems. As noted by Gao *et al.* [111], system performance can be affected when large arrays experience large scale fading such as shadowing effects. This makes antenna elements contribute unequally to the overall system performance. Hence, the Kronecker and Toeplitz assumptions on the fading and spatial correlation properties of large arrays cannot be well characterized. The variations in large-scale fading coefficients over the massive MIMO antenna array on sum-capacity of massive MIMO systems are taken into consideration in [112], [113]. The work carried out by Yang *et al.* [113] suggests that large-scale fading will have a significant effect on the capacity of the system under both perfect and imperfect CSI especially in the urban scenario. Channel measurement should be carried out to capture the properties of the real channel in order to come out with a more realistic model [114]. Furthermore, real-world cellular environment is extremely challenging where UTs have widely varying gains with cell-edge UTs experiencing the largest amount of other-cell interference. Hence using the worst-case scenario in pilot contamination only represents a degree to which the system is affected and does not give a true picture of a dynamic system as presented in [15], [115].

3) Computational Complexity and Cost : The cost of both end-user mobile phone and the BS as noted in [116] are very sensitive. The added cost and complexity for advanced signal processing algorithms for elimination of pilot contamination in massive MIMO needs to be considered carefully. It is expected that massive MIMO will be more demanding in terms of signal processing than conventional systems although UTs can be as simple as today or even more simpler using the key properties of maximum ratio transmission and combining (MRT)/MRC. As noted in [65], [117], linear pre-coding techniques and CE methods such as LS and MMSE involve high computational complexity owing to the fact that complicated matrix inversion of a large dimension is required. Hence, there is a need for less complex pre-coding techniques and CE methods. While some of the proposed methods to mitigate pilot contamination sound promising theoretically, there is a need to evaluate their performance by considering the trade-off between its accuracy and complexity. This will require further research in implementing and testing of the proposed models in realistically deployed massive MIMO systems.

4) The Use of Channel Reciprocity : Majority of the mitigation techniques for pilot contamination in a multi-cell system reviewed in this paper are based on the assumption that the physical forward and reverse channels are reciprocal. However, as noted in [23], [118] the DL channel is not reciprocal to the UL channel due to random phase and amplitude difference in the RF hardware. In view of this, the proposed mitigation techniques, either pilot-based or subspace-based approach, need to consider calibration schemes such as an internal or relative calibration scheme as proposed in [119], [120]. Considering the calibration technique in mitigation of pilot contamination will provide a more realistic approach to pilot mitigation in massive MIMO system.

D. Broader Perspectives

Since CSI is an important topic in wireless communication and pilot training remains one of the most widely used methods, we survey the effect of pilot contamination in other technologies such as heterogeneous network (HetNet) and in-band full-duplex (IBFD) MIMO systems. Furthermore, we also discuss the effect of pilot contamination in the coexistence of massive MIMO in a multi-tier network [121] and IBFD technology with massive MIMO [122]–[124].

1) Heterogeneous Network : Multi-tier HetNet has been identified as a key technology in the architecture for 5G cellular networks [5], [125], [126]. According to Hossain and Hasan [5], 5G cellular will be a multi-tier HetNet consisting of macro-cells along with a large number of low power nodes such as small cells access points (SCAs), relays, remote radio heads along with the provisioning for point-to-point (such as device-to-device and machine-to-machine) communication. The low power nodes are primarily added to increase capacity in hot spots (indoors and outdoors), especially in areas which are not covered by the marco-cells. They also provide improved network performance by off-loading data from the macro-cells. The major issues identified in multi-tier network are complexity of network planning and co-channel interference (CCI).

The deployment of the small nodes are dynamic and unpredictable compared to the macro/micro cells and it introduces CCI from overlapping network tiers which in turn reduces the system performance across tiers. Channel estimation in the presence of CCI is bound to affect the system performance. Two channel assignment mechanisms for multi-tier network: dedicated channel and shared channel have been identified [127]. In the dedicated channel, the overlay network occupies a dedicated frequency band while in the shared channel configuration, the overlay tier of the network shares the entire frequency band with the underlying infrastructure. In the shared channel configuration, the users of both macro and low power nodes use the same frequency bands for UL and DL communication. The shared channel mechanism is considered as a better way of sharing resource in the HetNet [128]. However, the shared channel can suffer from significant interference compared to the dedicated channel. The pilot signals used for channel estimation in the TDD mode between target UT and BS is not only contaminated by the neighboring macro cells, but also by the surrounding low power nodes which are within the coverage area.

The coexistence of massive MIMO with small cells access point is seen as a promising solution in a two-tier HetNet to improve network performance [7], [121]. In [121], the SCAs serve the static and low mobility UTs while the excess antennas at the BS are used to serve the macro UTs (MUTs) and at the same time to play the role of wireless backhaul to the SCAs. Interference from the UTs as they try to connect to the SCAs or macro BS is bound to cause pilot contamination during the training phase. To overcome this problem, a reverse TDD mode can be used, i.e. the BS is in DL mode when the SCAs operate in UL, and vice versa. While majority of the studies have focused on pilot contamination in homogeneous networks, there is a need for more research work on pilot contamination in HetNets. Novel pilot mitigation techniques and resource allocation will be needed in multi-tier network compared to existing methods such as the fractional pilot reuse and coordination among the multi-cell and multi-tier network due to varying transmit power and traffic load imbalance.

2) In-Band Full-Duplex MIMO systems : The concept of IBFD in which terminals (eg. BS, relays, or mobiles) transmit and receive simultaneously over same frequency band at the same time has the potential of doubling the spectral efficiency. The IBFD takes advantage of full-duplex transmission either in bidirectional or unidirectional communication and the advantages it offers over the TDD and FDD half-duplex schemes have been discussed in [129]. The history of IBFD dates back to the 1940s as discussed in [130] and recently, it has garnered lot of research interest because it is considered as a key technology for future wireless networks [131] especially in the shift towards small-cell systems [130]. Although most research work focused on full-duplex relay, the concept has also been extended to other areas such as cellular networks, access points, cognitive radio, HetNet and wireless network virtualization. IBFD with large MIMO has been studied in [122]–[124].

To achieve full-duplex, the terminal has to sufficiently suppress the significant self-interference (SI) that results from a power leakage between the transmit and receive antennas to below the receiver noise floor [132]. The SI enters the receiver through direct crosstalk, limited antenna isolation and reflections through the environment [133]. Studies from [130], [131], [134] show that SI cancellation can be categorized in three domains: analog-circuit domain, digital-domain and propagation-domain. The mitigation techniques can further be divided into two categories: passive suppression and active cancellation. Each cancellation technique can either operate in channel-aware or channel-unaware manner based on how well they adapt to environmental effect. Several mitigation techniques for SI have been proposed by various researchers and they are well documented in [122]–[124]. The cancellation techniques proposed in the digital-domain requires accurate knowledge of the SI channels and delays in order to cancel out the direct and reflected SI. We discuss the issue of pilot contamination in the digital-domain cancellation because known pilot signals are used to eliminate the residual linear and non-linear components of the SI [132], [135].

The digital-domain cancellation is considered as the last line of defense, where it is expected to cancel the SI left over from the propagation-domain and analog-circuit-domain cancellation approaches [130]. In the digital-domain cancellation, two processes are involved: estimating the SI channel; and using the channel estimate on the known transmit signal to generate digital samples to subtract from the received signal [135]. The CE process as described by authors in [135], [136] requires training symbols to be sent by the transmitter to estimate the SI. The self-interference training symbol is being received after transmission from the same radio. The received training symbol can be contaminated either by another interfering transmitter present during the CE phase, non-linear distortions in the RF circuit, coherence time of the SI channel or I/Q imbalance [135], [137]. The pilot contamination affects the accuracy of the CE and the effects of CE errors on self-interference mitigation performance in IBFD have been studied in [136] where a lower bound on the self-interference mitigation residual caused by channel estimation errors was formulated.

The management of interference in IBFD massive MIMO systems is expected to become significantly more complex due to interference that exists in a massive MIMO system when a pilot reuse factor of 1 is adopted and the SI that results from the IBFD. Hence a robust algorithm and mitigation technique for pilot contamination will be required in IBFD massive MIMO systems. This could lead to high complexity and computational cost.

E. Future Trends

A lot of research activities are on going in 5G network, the choice of TDD mode or FDD mode is still debatable. TDD systems seem promising because they operate on the same carrier frequency and the physical forward and backward channels are considered reciprocal [138]. Evaluation of proposed methods based on complexity and cost is likely to pave way for new optimized methods in mitigating pilot contamination in TDD systems. Although the worst-case scenario have been extensively studied, the effect of pilot contamination in a dynamic or real-world system is likely to attract further studies by considering more deployment scenarios. Possible extension is to investigate the effects of pilot contamination in a multi-cell network by providing performance analysis based on Poison Point Process and stochastic geometry. New methods in mitigating pilot contamination in massive MIMO systems with heterogeneous UTs requires more research work (see e.g, [139], [140]).

Significant amount of work focused on the impact of pilot contamination on system performance based on achievable rates, spectrum efficiency and EE. More work is expected on how pilot contamination influences the performance of the system based on quality-of-service. Furthermore, novel methods on mitigation of pilot contamination are expected for multi-cell systems that offer higher performance per users in terms of bit/s/Hz when a frequency reuse factor of 1 is employed. In addition, such methods are expected to take into consideration all sources of pilot contamination and not just limiting it to the non-orthogonal reuse training sequence. This would be beneficial to modeling massive MIMO systems for real-life scenarios. More research effort is required in mitigation of pilot contamination in the coexistence of massive MIMO in multi-tier heterogeneous network and practical implementation of IBFD massive MIMO systems.

VIII. Conclusions

In this paper, an overview of TDD and FDD schemes in massive MIMO systems is presented. We have analyzed the impact of pilot contamination in TDD massive MIMO systems through published work. Different methods for mitigating pilot contamination due to limited coherence time proposed by various authors are reviewed. We have also categorized the proposed methods based on the approach used in estimating channel information, i.e, pilot-based and subspace-based approaches. It is observed that majority of the proposed mitigation methods assume channel reciprocity, whereas in a practical system hardware impairments and non-reciprocal transceivers can be sources of pilot contamination. In addition we have explored some broader perspectives, such as pilot contamination in in-band full-duplex system and HetNet. We have also discussed the open issues in pilot contamination and possible future direction of research in pilot contamination in TDD massive MIMO systems.

Appendix

ACRONYMS and **TERMS**

3GPP - 3rd generation partnership project
AoA - angle-of-arrival
BS - base station
CP - conventional time-multiplexed pilot scheme
CCI - co-channel interference
CE - channel estimation
CSI - channel state information
DoA - direction-of-arrival
DL - downlink
EE - energy efficiency
EVD - eigenvalue decomposition
FDD - frequency division duplex
HetNet - heterogeneous network
IBFD - in-band full-duplex
I/Q - quadrature imbalance
i.i.d - independent and identically distributed
LSFD - large scale fading decoding
LSFP - large-scale fading pre-coding
M2M - machine-to-machine
MAP - maximum a-posteriori
MIMO - multiple input multiple output
MU-MIMO - multi-user MIMO
MMSE - minimum mean square error
MRC - maximum ratio combining
MRT - maximum ratio transmission
MVU - minimum variance unbiased
RF - radio frequency
SI - self-interference
SIC - Successive interference cancellation
SCAs - small cells access points
SIMO - single-input multiple-output
SIR - signal to interference ratio
SIP - superimposed pilot scheme
SINR - signal-to-interference-plus-noise ratio
SVD - singular value decomposition
TB - training-based
TDD - time division duplex
UL - uplink
UTs/MS - user terminal or mobile station
ZF - zero forcing
kth - the kth UT in K UTs
mth - the mth antenna in M
lth cell - one of the cells in the L cells
p_d - average power during transmission at the BS
p_u - average power during transmission at the UT
β - inter-cell interference factor
τ - pilot training length
K- total number of UTs
L- total number of cells
M- total number of BS antennas

Acknowledgment

The authors wish to acknowledge the feedback and contribution of Dr. Emil Björnson an Assistant Professor at Linköping University, Sweden and Dr. Thomas Peter of Service Communication International Pte Ltd, Singapore. We say thank you to the anonymous reviewers for their valuable feedback, which helped in improving the presentation of the paper.

References

[1] O. Elijah, C. Y. Leow, A. Tharek, S. Nunoo, and S. Iliya, "Mitigating pilot contamination in massive MIMO system - 5G: An overview," in *Control Conference (ASCC), 2015 10th Asian*, May 2015, pp. 1–6.

[2] Q. Li, H. Niu, A. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *Vehicular Technology Magazine, IEEE*, vol. 9, no. 1, pp. 71–78, March 2014.

[3] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2014-2019," Feb 2015.

[4] N. Prasad, H. Zhang, H. Zhu, and S. Rangarajan, "Multi-user MIMO scheduling in the fourth generation cellular uplink," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, Nov 2013, pp. 1855–1859.

[5] E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," *Instrumentation & Measurement Magazine, IEEE*, vol. 18, no. 3, pp. 11–21, 2015.

[6] R. Chávez-Santiago, M. Szydełko, A. Kliks, F. Foukalas, Y. Haddad, K. E. Nolan, M. Y. Kelly, M. T. Masonta, and I. Balasingham, "5G: The convergence of wireless communications," *Wireless Personal Communications*, pp. 1–26.

[7] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[8] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the metis project," *Communications Magazine, IEEE*, vol. 52, no. 5, pp. 26–35, May 2014.

[9] ICT-317669METISproject, "Positioning of multi-node/multi-antenna transmission technologies," July 2013. [Online]. Available: https://www.metis2020.com/documents/deliverables/.

[10] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[11] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 11, pp. 3590–3600, 2010.

[12] 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," *Signal Processing Magazine, IEEE*, vol. 30, no. 1, pp. 40–60, 2013.

[13] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *Communications, IEEE Transactions on*, vol. 61, no. 4, pp. 1436–1449, April 2013.

[14] H. Q. Ngo, E. Larsson, and T. Marzettat, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, Sept 2011, pp. 1272–1279.

[15] E. Bjornson, E. G. Larsson, and M. Debbah, "Optimizing multi-cell massive MIMO for spectral efficiency: How many users should be scheduled?" in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*. IEEE, 2014, pp. 612–616.

[16] J. Hoydis, S. Ten Brink, M. Debbah *et al.*, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, 2013.

[17] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 742–758, 2014.

[18] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing - The large-scale array regime," *Information Theory, IEEE Transactions on*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.

[19] J. Nam, J.-Y. Ahn, A. Adhikary, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*. IEEE, 2012, pp. 1–6.

[20] J. Choi, D. Love, and T. Kim, "Trellis-extended codebooks and successive phase adjustment: A path from LTE-advanced to FDD massive MIMO systems," 2014.

[21] S. L. H. Nguyen and A. Ghrayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. IEEE, 2013, pp. 2890–2895.

[22] T. Marzetta, "How much training is required for multiuser MIMO?" in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, Oct 2006, pp. 359–363.

[23] X. Hou, A. Harada, and H. Suda, "Experimental study of advanced MU-MIMO scheme with antenna calibration for the evolving LTE TDD system," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, Sept 2012, pp. 2443–2448.

[24] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *Information Theory, IEEE Transactions on*, vol. 49, no. 4, pp. 951–963, April 2003.

[25] E. Bjornson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated rician MIMO channels with rician disturbance," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1807–1820, 2010.

[26] E. Bjornson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *Wireless Communications, IEEE Transactions on*, vol. 14, no. 6, pp. 3059–3075, June 2015.

[27] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 8, pp. 2640–2651, August 2011.

[28] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *Vehicular Technology, IEEE Transactions on*, vol. 48, no. 4, pp. 1047–1066, 1999.

[29] C. ZHANG and G. ZENG, "Pilot contamination reduction scheme in massive MIMO multi-cell TDD systems," *Journal of Computational Information Systems*, vol. 10, no. 15, pp. 6721–6729, 2014.

[30] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 74–80, February 2014.

[31] A. Ashikhmin, T. L. Marzetta, and L. Li, "Interference reduction in multi-cell massive mimo systems i: Large-scale fading precoding and decoding," *arXiv preprint arXiv:1411.4182*, 2014.

[32] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination problem in multi-cell TDD systems," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 2184–2188.

[33] H. Huh, A. Tulino, and G. Caire, "Network MIMO with linear zero-forcing beamforming: Large system analysis, impact of channel estimation, and reduced-complexity scheduling," *Information Theory, IEEE Transactions on*, vol. 58, no. 5, pp. 2911–2934, May 2012.

[34] H. Huh, S.-H. Moon, Y.-T. Kim, I. Lee, and G. Caire, "Multi-cell MIMO downlink with cell cooperation and fair scheduling: A large-system limit analysis," *Information Theory, IEEE Transactions on*, vol. 57, no. 12, pp. 7771–7786, Dec 2011.

[35] T. Zhou, M. Peng, W. Wang, and H.-H. Chen, "Low-complexity coordinated beamforming for downlink multicell SDMA/OFDM systems," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 1, pp. 247–255, Jan 2013.

[36] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[37] F. Fernandez, A. Ashikhmin, and T. Marzetta, "Interference reduction on cellular networks with large antenna arrays," in *Proc. of IEEE International Conference on Communications (ICC)*, 2012.

[38] F. Fernandes, A. Ashikhmin, and T. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 192–201, February 2013.

[39] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *Information Theory, IEEE Transactions on*, vol. 60, no. 11, pp. 7112–7139, Nov 2014.

[40] B. Maham and O. Tirkkonen, "Transmit antenna selection OFDM systems with transceiver I/Q imbalance," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 2, pp. 865–871, 2012.

[41] E. Bjornson, P. Zetterberg, and M. Bengtsson, "Optimal coordinated beamforming in the multicell downlink with transceiver impairments," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 4775–4780.

[42] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *arXiv preprint arXiv:1306.4495*, 2013.

[43] N. Moghadam, P. Zetterberg, P. Handel, and H. Hjalmarsson, "Correlation of distortion noise between the branches of MIMO transmit antennas," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, Sept 2012, pp. 2079–2084.

[44] T. L. Marzetta, G. Caire, M. Debbah, I. Chih-Lin, and S. K. Mohammed, "Special issue on massive MIMO," *Communications and Networks, Journal of*, vol. 15, no. 4, pp. 333–337, 2013.

[45] L. Berriche, K. Abed-Meraim, and J.-C. Belfiore, "Investigation of the channel estimation error on MIMO system performance," *Third European Signal Processing (EUSIPCO2005), Antalya*, 2005.

[46] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *Information Theory, IEEE Transactions on*, vol. 46, no. 3, pp. 933–946, 2000.

[47] T. Yoo and A. Goldsmith, "Capacity of fading MIMO channels with channel estimation error," in *Communications, 2004 IEEE International Conference on*, vol. 2. IEEE, 2004, pp. 808–813.

[48] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 12, pp. 3055–3069, 2002.

[49] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 884–893, 2006.

[50] T. Wo and P. A. Hoeher, "Semi-blind channel estimation for frequency-selective MIMO systems," in *IST Mobile & Wireless Communications Summit*. Citeseer, 2005.

[51] H. Bolcskei, R. W. Heath, and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna systems," *Signal Processing, IEEE Transactions on*, vol. 50, no. 1, pp. 96–109, 2002.

[52] B. Muquet, M. De Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *Signal Processing, IEEE Transactions on*, vol. 50, no. 7, pp. 1699–1712, 2002.

[53] R. Kudo, S. Armour, J. McGeehan, and M. Mizoguchi, "A channel state information feedback method for massive MIMO-OFDM," *Communications and Networks, Journal of*, vol. 15, no. 4, pp. 352–361, Aug 2013.

[54] L. Tong, B. M. Sadler, and M. Dong, “Pilot-assisted wireless transmissions: general model, design criteria, and signal processing,” *Signal Processing Magazine, IEEE*, vol. 21, no. 6, pp. 12–25, 2004.

[55] M. Coldrey and P. Bohlin, “Training-based MIMO systems–part I : Performance comparison,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 11, pp. 5464–5476, 2007.

[56] ——, “Training-based MIMO systems : Part II–Improvements using detected symbol information,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 296–303, 2008.

[57] G. Taricco and E. Biglieri, “Space-time decoding with imperfect channel estimation,” *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1874–1888, 2005.

[58] C. Tellambura, M. G. Parker, Y. J. Guo, S. J. Shepherd, and S. K. Barton, “Optimal sequences for channel estimation using discrete fourier transform techniques,” *Communications, IEEE Transactions on*, vol. 47, no. 2, pp. 230–238, 1999.

[59] S. Ohno and G. B. Giannakis, “Optimal training and redundant precoding for block transmissions with application to wireless OFDM,” *Communications, IEEE Transactions on*, vol. 50, no. 12, pp. 2113–2123, 2002.

[60] X. Ma, L. Yang, and G. B. Giannakis, “Optimal training for MIMO frequency-selective fading channels,” *Wireless Communications, IEEE Transactions on*, vol. 4, no. 2, pp. 453–466, 2005.

[61] Y. Shen and E. Martinez, “Channel estimation in OFDM systems,” *Application note, Freescale semiconductor*, 2006.

[62] A. Chockalingam and B. S. Rajan, *Large MIMO Systems*. Cambridge University Press, 2014.

[63] J. Hoydis, M. Kobayashi, and M. Debbah, “Optimal channel training in uplink network MIMO systems,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 6, pp. 2824–2833, June 2011.

[64] D. Neumann, A. Grundinger, M. Joham, and W. Utschick, “On the amount of training in coordinated massive MIMO systems,” in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2014 IEEE 8th*. IEEE, 2014, pp. 293–296.

[65] K. Li, X. Song, M. O. Ahmad, and M. Swamy, “An improved multicell MMSE channel estimation in a massive MIMO system,” *International Journal of Antennas and Propagation*, vol. 2014, 2014.

[66] N. Shariati, E. Bjornson, M. Bengtsson, and M. Debbah, “Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.

[67] ——, “Low-complexity channel estimation in large-scale MIMO using polynomial expansion,” in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 1157–1162.

[68] M. Biguesh and A. B. Gershman, “Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 884–893, 2006.

[69] D. Neumann, M. Joham, and W. Utschick, “Channel estimation in massive MIMO systems,” *arXiv preprint arXiv:1503.08691*, 2015.

[70] H. Zhang, X. Zheng, W. Xu, and X. You, “On massive MIMO performance with semi-orthogonal pilot-assisted channel estimation,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2014, no. 1, pp. 1–14, 2014.

[71] S. Noh, M. D. Zoltowski, Y. Sung, and D. J. Love, “Pilot beam pattern design for channel estimation in massive MIMO systems,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 787–801, 2014.

[72] A. Hu, T. Lv, H. Gao, Y. Lu, and E. Liu, “Pilot design for large-scale multi-cell multiuser MIMO systems,” in *Communications (ICC), 2013 IEEE International Conference on*, June 2013, pp. 5381–5385.

[73] T. Marzetta and B. Hochwald, “Fast transfer of channel state information in wireless systems,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 4, pp. 1268–1278, April 2006.

[74] B. Gopalakrishnan and N. Jindal, “An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas,” in *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*, June 2011, pp. 381–385.

[75] 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 *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 3464–3467.

[76] Y. Yu and D. Gu, “Enhanced MU-MIMO downlink transmission in the FDD-based distributed antennas system,” *Communications Letters, IEEE*, vol. 16, no. 1, pp. 37–39, January 2012.

[77] J. Chang, I. Lu, and Y. Li, “Adaptive codebook-based channel prediction and interpolation for multiuser multiple-input multiple-output - orthogonal frequency division multiplexing systems,” *Communications, IET*, vol. 6, no. 3, pp. 281–288, February 2012.

[78] A. Rajanna and N. Jindal, “Multiuser diversity in downlink channels: When does the feedback cost outweigh the spectral efficiency benefit?” *Wireless Communications, IEEE Transactions on*, vol. 11, no. 1, pp. 408–418, January 2012.

[79] Z. Jiang, A. Molisch, G. Caire, and Z. Niu, “Achievable rates of FDD massive MIMO systems with spatial channel correlation,” *Wireless Communications, IEEE Transactions on*, vol. 14, no. 5, pp. 2868–2882, May 2015.

[80] J. Choi, D. Love, and P. Bidigare, “Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.

[81] J. Choi, Z. Chance, D. Love, and U. Madhow, “Noncoherent trellis coded quantization: A practical limited feedback technique for massive MIMO systems,” *Communications, IEEE Transactions on*, vol. 61, no. 12, pp. 5016–5029, December 2013.

[82] X. Rao and V. K. Lau, “Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems,” *Signal Processing, IEEE Transactions on*, vol. 62, no. 12, pp. 3261–3271, 2014.

[83] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, “Scheduling and pre-conditioning in multi-user MIMO TDD systems,” in *Communications, 2008. ICC '08. IEEE International Conference on*, May 2008, pp. 4100–4105.

[84] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 264–273, February 2013.

[85] M. Filippou, D. Gesbert, and H. Yin, “Decontaminating pilots in cognitive massive MIMO networks,” in *Wireless Communication Systems (ISWCS), 2012 International Symposium on*, Aug 2012, pp. 816–820.

[86] P. Zetterberg, “Experimental investigation of TDD reciprocity-based zero-forcing transmit precoding,” *EURASIP Journal on Advances in Signal Processing*, vol. 2011, p. 5, 2011.

[87] B. Goransson, S. Grant, E. Larsson, and Z. Feng, “Effect of transmitter and receiver impairments on the performance of MIMO in HSDPA,” in *Signal Processing Advances in Wireless Communications, 2008. SPAWC 2008. IEEE 9th Workshop on*. IEEE, 2008, pp. 496–500.

[88] C. Studer, M. Wenk, and A. Burg, “MIMO transmission with residual transmit-RF impairments,” in *Smart Antennas (WSA), 2010 International ITG Workshop on*. IEEE, 2010, pp. 189–196.

[89] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, “Relative channel reciprocity calibration in MIMO/TDD systems,” in *Future Network and Mobile Summit, 2010*. IEEE, 2010, pp. 1–10.

[90] M. Guillaud and F. Kaltenberger, “Towards practical channel reciprocity exploitation: Relative calibration in the presence of frequency offset,” in *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*. IEEE, 2013, pp. 2525–2530.

[91] E. Björnson and E. Jorswieck, “Optimal resource allocation in coordinated multi-cell systems,” *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2-3, pp. 381–p, 2013.

[92] N. Krishnan, R. Yates, and N. Mandayam, “Cellular systems with many antennas: Large system analysis under pilot contamination,” in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, Oct 2012, pp. 1220–1224.

[93] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel,” *Communications, IEEE Transactions on*, vol. 61, no. 6, pp. 2350–2361, 2013.

[94] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, “Pilot contamination reduction in multi-user TDD systems,” in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[95] H. Yin, D. Gesbert, M. C. Filippou, and Y. Liu, “Decontaminating pilots in massive MIMO systems,” in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 3170–3175.

[96] J. Zhang, B. Zhang, S. Chen, X. Mu, M. El-Hajjar, and L. Hanzo, “Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.

[97] H. Wang, Z. Pan, J. Ni, and C.-L. I, “A spatial domain based method against pilot contamination for multi-cell massive MIMO systems,” in *Communications and Networking in China (CHINACOM), 2013 8th International ICST Conference on*, Aug 2013, pp. 218–222.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

[98] A. Ashikhmin and T. Marzetta, "Pilot contamination precoding in multi-cell large scale antenna systems," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 1137–1141.

[99] L. Li, A. Ashikhmin, and T. Marzetta, "Interference reduction in multi-cell massive MIMO Systems II: Downlink analysis for a finite number of antennas," *arXiv preprint arXiv:1411.4183*, 2014.

[100] A.-J. van der Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *Signal Processing, IEEE Transactions on*, vol. 45, no. 1, pp. 173–190, Jan 1997.

[101] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3249–3252.

[102] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array. I. Algorithms," *Signal Processing, IEEE Transactions on*, vol. 44, no. 5, pp. 1184–1197, May 1996.

[103] R. R. Müller, M. Vehkaperä, and L. Cottatellucci, "Blind pilot decontamination," in *Smart Antennas (WSA), 2013 17th International ITG Workshop on*. VDE, 2013, pp. 1–6.

[104] L. Cottatellucci, R. Muller, and M. Vehkapera, "Analysis of pilot decontamination based on power control," in *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*, June 2013, pp. 1–5.

[105] R. Mueller, L. Cottatellucci, and M. Vehkapera, "Blind pilot decontamination," *Selected Topics in Signal Processing, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2014.

[106] R. R. Muller, M. Vehkapera, and L. Cottatellucci, "Analysis of blind pilot decontamination," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, Nov 2013, pp. 1016–1020.

[107] D. Neumann, A. Gruendinger, M. Joham, and W. Utschick, "Pilot coordination for large-scale multi-cell TDD systems," in *Smart Antennas (WSA), 2014 18th International ITG Workshop on*, March 2014, pp. 1–6.

[108] M. Latif Sarker and M. H. Lee, "A fast channel estimation and the reduction of pilot contamination problem for massive MIMO based on a diagonal jacket matrix," in *Fiber Optics in Access Network (FOAN), 2013 4th International Workshop on*, Sept 2013, pp. 26–30.

[109] D. Neumann, M. Joham, and W. Utschick, "Suppression of pilot-contamination in massive MIMO systems," in *Signal Processing Advances in Wireless Communications (SPAWC), 2014 IEEE 15th International Workshop on*, June 2014, pp. 11–15.

[110] K. Guo, Y. Guo, and G. Ascheid, "On the performance of EVD-based channel estimations in MU-Massive-MIMO systems," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. IEEE, 2013, pp. 1376–1380.

[111] X. Gao, M. Zhu, F. Rusek, F. Tufvesson, and O. Edfors, "Large antenna array and propagation environment interaction," in *The 48th Annual Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2014.

[112] M. Kazemi and H. Aghaeinia, "Approximate ergodic capacity of multiuser massive multiple input multiple output in a rayleigh fading uplink channel with variance profile," *IET Communications*, vol. 9, no. 6, pp. 844–852, 2015.

[113] A. Yang, Z. He, C. Xing, Z. Fei, and J.-M. Kuang, "The role of large-scale fading in uplink massive MIMO systems," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[114] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *Wireless Communications, IEEE Transactions on*, vol. 14, no. 7, pp. 3899–3911, July 2015.

[115] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *arXiv preprint arXiv:1412.7102*, 2014.

[116] J. G. Andrews, "Interference cancellation for cellular systems: a contemporary overview," *Wireless Communications, IEEE*, vol. 12, no. 2, pp. 19–29, 2005.

[117] L. Dai, X. Gao, S. Han, Z. Wang *et al.*, "Near-optimal linear precoding with low complexity for massive MIMO," *arXiv preprint arXiv:1411.4141*, 2014.

[118] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 53–64.

[119] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Future Network and Mobile Summit, 2010*. IEEE, 2010, pp. 1–10.

[120] M. Guillaud, D. T. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration." in *ISSPA*. Citeseer, 2005, pp. 403–406.

[121] L. Sanguinetti, A. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD hetnets with wireless backhaul: A large system analysis," *Selected Areas in Communications, IEEE Journal on*, vol. 33, no. 6, pp. 1187–1200, June 2015.

[122] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and J. Lilleberg, "Full-duplex in large-scale wireless systems," in *Signals, Systems and Computers, 2013 Asilomar Conference on*. IEEE, 2013, pp. 1623–1627.

[123] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU: Enabling MIMO full duplex," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 257–268.

[124] X. Du, J. Tadrous, C. Dick, and A. Sabharwal, "Multiuser MIMO beamforming with full-duplex open-loop training," *arXiv preprint arXiv:1505.01746*, 2015.

[125] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 122–130, 2014.

[126] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *Access, IEEE*, vol. 3, pp. 1206–1232, 2015.

[127] H. Mahmoud, I. Guvenc, F. Watanabe *et al.*, "Performance of open access femtocell networks with different cell-selection methods," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. IEEE, 2010, pp. 1–5.

[128] X. An and F. Pianese, "Understanding co-channel interference in LTE-based multi-tier cellular networks," in *Proceedings of the 9th ACM symposium on Performance evaluation of wireless ad hoc, sensor, and ubiquitous networks*. ACM, 2012, pp. 107–112.

[129] X. ZHANG, W. CHENG, and H. ZHANG, "Full-duplex transmission in PHY and MAC layers for 5G mobile wireless networks," *IEEE Wireless Communications*, p. 3, 2015.

[130] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 9, pp. 1637–1652, 2014.

[131] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1–1, 2015.

[132] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4. ACM, 2013, pp. 375–386.

[133] B. Debaillie, D.-J. van den Broek, C. Lavin, B. van Liempd, E. Klumperink, C. Palacios, J. Craninckx, B. Nauta, A. Parssinen *et al.*, "Analog/RF solutions enabling compact full-duplex radios," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 9, pp. 1662–1673, 2014.

[134] G. Liu, F. Yu, H. Ji, V. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *Communications Surveys Tutorials, IEEE*, vol. 17, no. 2, pp. 500–524, Secondquarter 2015.

[135] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM, 2011, pp. 301–312.

[136] D. W. Bliss and Y. Rong, "Effects of channel estimation errors on in-band full-duplex MIMO radios using adaptive transmit spatial mitigation," in *Signals, Systems and Computers, 2013 Asilomar Conference on*. IEEE, 2013, pp. 9–13.

[137] Z. Zhan, G. Villemaud, F. Hutu, and J.-M. Gorce, "Digital estimation and compensation of I/Q imbalance for full-duplex dual-band OFDM radio," in *PIMRC 2014*.

[138] G. S. Smith, "A direct derivation of a single-antenna reciprocity relation for the time domain," *Antennas and Propagation, IEEE Transactions on*, vol. 52, no. 6, pp. 1568–1577, 2004.

[139] R. Mochaourab, E. Björnson, and M. Bengtsson, "Adaptive pilot clustering in heterogeneous massive MIMO networks," *arXiv preprint arXiv:1507.04869*, 2015.

[140] S. L. H. Nguyen, T. Le-Ngoc, and A. Ghrayeb, "Robust channel estimation and scheduling for heterogeneous multiuser massive MIMO systems," in *European Wireless 2014; 20th European Wireless Conference; Proceedings of*. VDE, 2014, pp. 1–6.

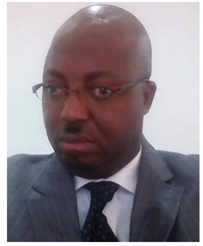
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2504379, IEEE Communications Surveys & Tutorials

Olakunle Elijah received a B.Eng. from Federal University of Technology Minna, Nigeria in 2003, MSc in Micro-Electronics & Computing from Bournemouth University, (UK) in 2008, and Postgraduate Certificate in Advance Microelectronics from Bolton University (UK) in 2010. He is currently a Ph.D. student at the Universiti Teknologi Malaysia and is conducting research in the field of wireless communications. He worked as a field engineer for Kuyet Nigeria Ltd in 2006. From 2011 to 2013, he was the MD/CEO at Mircoscale Embedded Limited, Nigeria. His research interests are in embedded systems, wireless communication, massive MIMO, 5G, interference mitigation and heterogeneous network.

Chee Yen Leow obtained his B.Eng. degree in Computer Engineering from Universiti Teknologi Malaysia (UTM) in 2007. Since July 2007, he has been an academic staff in the Faculty of Electrical Engineering, UTM. In 2011, he obtained a Ph.D. degree from Imperial College London. He is currently a senior lecturer in the faculty and a member of the Wireless Communication Center (WCC), UTM. His research interest includes but not limited to wireless relaying, MIMO, physical layer security, convex optimization, communications theory, near field wireless charging and 5G. He may be contacted at bruceleow@fke.utm.my

Abdul Rahman Tharek currently is a professor in wireless communication at faculty of electrical engineering, Universiti Teknologi Malaysia. He obtained his BSc (Hons) (Electrical Engineering) from University of Strathclyde, UK , MSc in Communication Engineering from UMIST, Manchester, UK and Ph.D. in Mobile Communication from University of Bristol, UK. He is the Director of Wireless Communication Centre (WCC), Faculty of Electrical Engineering, Universiti Teknologi Malaysia and currently conducting research related to mobile communications, antenna and propagation. He has also conducted various short courses related to mobile and satellite communication to the telecommunication industry and government agencies since 1988. Prof. Tharek has published more than 300 scientific papers in journals and conferences and obtained many national and international awards. He is also a consultant for many communication companies and an active member in several research academic entities.

Solomon Nunoo received the B.S. degree in electrical/electronic engineering from Western University College of KNUST, now University of Mines and Technology (UMaT), Tarkwa in 2004, the M.Phil. degree in electrical and electronic engineering from UMaT, Tarkwa in 2007 and a Ph.D. degree in electrical engineering at Universiti Teknologi Malaysia (UTM), Johor in 2015. He is currently a lecturer at the Department of Electrical and Electronic Engineering, University of Mines and Technology, where he also served as the head of the Electrical and Electronic Engineering laboratories. His research interests lie in the area of energy management, and statistical signal processing with emphasis on adaptive filtering, compressive sampling, and wireless communications. He is a co-recipient for the Best Paper Award at the 2013 IEEE Malaysia International Conference on Communications (MICC). He is a member of the IEEE and IAENG.



Solomon Zakwoi Iliya obtained a Ph.D. degree from Wireless Communication Center (WCC), Universiti Teknologi Malaysia (UTM). His research interests include Microwave, Antenna and Satellite Subsystems Designs, among others. He has over 12 years experience in his field of study. He attended Beijing Institute of Tracking and Telecommunications Technology (BITTT) and the Federal University of Technology Minna for his postgraduate and undergraduate studies respectively. He is a registered and active member of the Nigerian Society of Engineers (NSE) with Reg no.(18515) as well as the Council for the Regulation of Engineering in Nigeria (COREN) with Reg no. RC.16003. Currently, he is a Deputy Director Engineer with the Center for Satellite Technology Development (CSTD); an activity arm of the National Space Research and Development Agency (NASRDA).