Analysis and Design of Channel Estimation in Multicell Multiuser MIMO OFDM Systems

Peng Xu, Jiangzhou Wang, *Senior Member, IEEE*, Jinkuan Wang, *Member, IEEE*, and Feng Qi

***Abstract*—This paper investigates the uplink transmission in multicell multiuser multiple-input multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems. The system model considers imperfect channel estimation, pilot contamination (PC), and multicarrier and multipath channels. Analytical expressions are first presented on the mean square error (MSE) of two classical channel estimation algorithms [i.e., least squares (LS) and minimum mean square error (MMSE)] in the presence of PC. Then, a simple H-infinity (H-inf) channel estimation approach is proposed to have good suppression to PC. This approach exploits the space-alternating generalized expectation–maximization (SAGE) iterative process to decompose the multicell multiuser MIMO (MU-MIMO) problem into a series of single-cell single-user single-input single-output (SISO) problems, which reduces the complexity significantly. According to the analytic results given herein, increasing the number of pilot subcarriers cannot mitigate PC, and a clue for suppressing PC is obtained. It is shown from the results that the H-inf has better suppression capability to PC than classical estimation algorithms. Its performance is close to that of the optimal MMSE as the length of channel impulse response (CIR) is increased. By using the SAGE process, the performance of the H-inf does not degrade when the number of antennas is large at the base station (BS).**

***Index Terms*—Channel estimation, H-inf, multiple-input multiple-output (MIMO), multicell, multiuser, orthogonal frequency-division multiplex (OFDM), pilot contamination (PC).**

I. Introduction

FUTURE wireless communications require the outstanding capability to combat multipath fading and to offer high spectral efficiency. Multiple-input multiple-output (MIMO) combined with orthogonal frequency-division multiplexing (OFDM) has been widely considered to be a promising candidate [1], [2]. Unlike the point-to-point MIMO, a multiuser MIMO (MU-MIMO) system that has low cost in terminals and better tolerance to wireless propagation environment has been considered for future wireless communications [3]. In a multicell scenario, it is well known that accurate channel state information (CSI) is critical for achieving high system performance. Since the mobility of users and the limited bandwidth, it is not possible to allocate dedicated pilots for the users in each cell, and therefore, the reuse of pilots is a must for users in different cells.

One of the main consequences of pilot reuse is pilot contamination (PC), which is caused by using nonorthogonal pilots to the users in different cells. PC has a more severe impact on the system performance than channel noise. When the system is deployed with an increasing number of antennas at the base station (BS) and serves a multiplicity of single-antenna terminals, the effects of fast fading and uncorrelated interference will vanish [4]–[13]. However, PC due to the reuse of nonorthogonal pilots in other cells does not vanish. In such a multicell MU-MIMO system, it has been shown in [4] that, with perfect CSI at the BS, the potential benefits in throughput, reliability, and power efficiency will be obtained. These benefits are analyzed mainly based on single-carrier and flat-fading system model; however, a more realistic performance analysis that considers multicarrier and frequency-selective fading channels for future cellular mobile systems is important [14]–[18]. Since the BS cannot have perfect CSI in practice, it is crucial to consider the effect of PC on channel estimation based on a multicarrier multipath system model.

A. Related Work and the Contribution of This Paper

There are few researches specifically focused on channel estimation algorithms in the presence of PC in multicell MU-MIMO systems, although single-carrier and flat-fading transmission scenario has been considered [6]–[8]. In [6], a blind channel estimation algorithm based on eigenvalue decomposition was proposed; however, it requires a long-data record and employs the prior knowledge of stochastic information and high computational complexity. In [7], a coordinated channel estimation approach with correlated pilot sequences was developed to tackle the problem of PC; however, the complexity due to applying second-order statistical information is high. In [8], the asymptotic analysis on the impact of channel aging on both the uplink and the downlink achievable rates was provided, and a finite-impulse-response Wiener predictor was proposed to overcome channel aging effects.

For a multicarrier and multipath scenario, pilot-based channel estimation techniques in OFDM or MIMO-OFDM systems

Manuscript received July 18, 2013; revised December 13, 2013 and February 14, 2014; accepted April 10, 2014. Date of publication May 8, 2014; date of current version February 9, 2015. This work was supported by the National Natural Science Foundation of China under Grant 61374097, Grant 61104005, and Grant 61300195; by the Basic Scientific Research Business Expenses of China under Grant N120323003; and by the China Scholarship Council. The review of this paper was coordinated by Prof. X. Wang.

P. Xu is with the School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K., and also with the Department of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: p.xu@kent.ac.uk).

J. Wang is with the School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K. (e-mail: j.z.wang@kent.ac.uk).

J. Wang is with the Department of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: wjk@mail.neuq.edu.cn).

F. Qi is with the Advanced Science Institute, RIKEN, Sendai 980-0845, Japan (e-mail: qifeng@riken.jp).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2322654

have been studied extensively for many years by focusing on a single-cell single-user scenario [19]–[23]. Least squares (LS) channel estimation, by using many pilot subcarriers, is usually considered an initial estimator without requiring any prior knowledge [19]. When the transmit symbols are known, the LS algorithm becomes maximum likelihood (ML) estimation [20]. To obtain optimal performance, minimum mean square error (MMSE) estimation, using channel correlation matrix and transmit data, has been investigated in [21]. However, the complexity of the MMSE algorithm is high. By avoiding the drawbacks of conventional algorithms, the H-infinity (H-inf) method is introduced into MIMO-OFDM systems. It has been proven that this algorithm has almost the same performance as MMSE but is much less complex [22], [23].

Different from the existing researches, our system model considers imperfect channel estimation, PC, multicarrier, and multipath channels. In this paper, we first discuss the impact of the PC on two classical LS and MMSE algorithms. Analytical expressions on the mean square error (MSE) are derived. It is shown that MMSE is more resistant to the PC than the LS due to the use of prior information. Increasing the number of pilot subcarriers in both algorithms does not increase suppression capability to the PC. From the results given herein, a clue for mitigating PC can be obtained. The performance of both the algorithms in the presence of PC could be improved as the length of channel impulse response (CIR) or the number of OFDM subcarriers increases. Because of the difficulty of capturing prior information and high computational load, using MMSE is not realistic in practice. Taking the advantage of the H-inf into account [22], [23], the H-inf approach is introduced, and the effect of PC is analyzed. By applying the space-alternating generalized expectation–maximization (SAGE) iterative process, the complexity due to multicell MU-MIMO estimation problem can be simplified. Moreover, detailed analysis of its MSE in presence of PC is presented. According to the given expressions, we can conclude that the H-inf algorithm, by adjusting the scalar factor, is more resistant to PC than LS and ML. The performance of the H-inf is close to optimal MMSE when the length of CIR is large. Meanwhile, when the number of antennas at the BS is large, no performance degradation for H-inf is seen during the iterative process of SAGE.

B. Notations

Bold italic font variables denote matrices and vectors; $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and the Hermitian transpose, respectively; and $(\cdot)^{-1}$ and $(\cdot)^{\dagger}$ stand for the inverse and pseudoinverse operations, respectively. $\mathcal{CN}(\Gamma, \Upsilon)$ denotes complex Gaussian distribution with mean Γ and covariance matrix Υ. $\|\cdot\|$ and $\|\cdot\|_\infty$ denote the two-norm and infinite-norm operations, respectively. $E[\cdot]$ stands for expectation operation. $\boldsymbol{I}$ is an identity matrix, and $\mathbf{0}$ is a matrix or vector in which all elements are zero.

The remainder of this paper is organized as follows. In Section II, the system model is discussed. In Section III, detailed analysis is presented on MSE in the presence of PC for classical LS and MMSE algorithms in multicell MU-MIMO systems. Next, an H-inf estimator is introduced in Section IV, and the SAGE iterative process is designed for reducing the complexity. After the performance analysis of the proposed H-inf algorithm, the performances of the aforementioned channel estimation algorithms are evaluated via computer simulations in Section V. Finally, conclusions are drawn in Section VI.

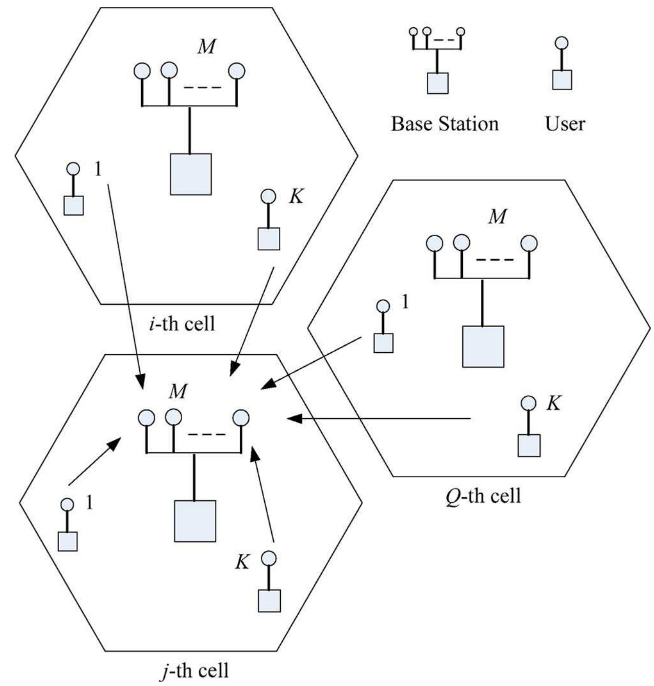


Fig. 1. Uplink transmission in multicell MU-MIMO systems.

II. System Model

We consider a multicell MU-MIMO system with Q cells, as shown in Fig. 1. Each cell includes one BS with M antennas and K single-antenna terminals. OFDM transmission with N subcarriers is considered. The frequency-selective fading channel is modeled as a finite-duration CIR with L taps. We assume that the uplink transmission from all users in the Q cells are synchronized, which constitutes a worst-case scenario from the standpoint of PC. Furthermore, the signals received for each antenna at the BS are assumed to experience independent fading.

The received $N \times 1$ signal vector on all N subcarriers at the rth antenna at the jth BS can be expressed as (for notational simplicity, we will neglect the antenna index r in this paper)

$$\boldsymbol{Y}_j = \boldsymbol{X}\boldsymbol{H}_j + \boldsymbol{Z}_j \tag{1}$$

where $\boldsymbol{Y}_j = [\boldsymbol{Y}_j(0), \ldots, \boldsymbol{Y}_j(N-1)]^T$, $\boldsymbol{X} = [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_Q]$, $\boldsymbol{X}_q$ is a diagonal matrix containing the transmit signal from the qth cell, and $\boldsymbol{Z}_j = [\boldsymbol{Z}_j(0), \ldots, \boldsymbol{Z}_j(N-1)]^T$ is a vector of independently and identically distributed (i.i.d.) complex zero-mean Gaussian noise variables with variance σ^2. $\boldsymbol{H}_j = [\boldsymbol{H}_{j1}^T, \ldots, \boldsymbol{H}_{jQ}^T]^T$, $\boldsymbol{H}_{jq}$ is the frequency response of the channel between the jth and qth cells, $\boldsymbol{H}_{jq} = [\boldsymbol{H}_{jq1}^T, \ldots, \boldsymbol{H}_{jqK}^T]^T$, $\mathbf{H}_{jqk} = \boldsymbol{F}_{N,L}\boldsymbol{C}_{jqk}$, $\boldsymbol{F}_{N,L}$ is $1/\sqrt{N}$ times the first L columns of discrete Fourier transform (DFT) matrix, $\boldsymbol{F}_{N,L}^H\boldsymbol{F}_{N,L} = \boldsymbol{I}_L$, and $\boldsymbol{C}_{jqk}$ is the $L \times 1$ propagation coefficients between the jth BS and the kth user in the qth cell and is given as follows:

$$\boldsymbol{C}_{jqk} = \boldsymbol{D}_{jqk}^{\frac{1}{2}} \boldsymbol{G}_{jqk} \tag{2}$$

where $\boldsymbol{G}_{jqk}$ denotes the $L \times 1$ fast-fading coefficient vector, which has an exponentially decaying multipath power-delay profile. The maximum tap delay is assumed shorter than the OFDM cyclic prefix (CP). Let g_{jqkl} be the lth element of $\boldsymbol{G}_{jqk}$, and it is normalized to unity, i.e., $\sum_{l=0}^{L-1} g_{jqkl}^2 = 1$. $\boldsymbol{D}_{jqk}$ is an $L \times L$ diagonal matrix whose diagonal elements d_{jqkl} denote path loss and shadow fading, which are assumed to be independent over l and k.[1] Since d_{jqkl} changes slowly, we rewrite it as d_{jq} for notational simplicity, and d_{jq} is assumed to be less than 1.

Rewrite the received vector at the jth BS as

$$\boldsymbol{Y}_j = \sum_{q=1}^{Q} \sum_{k=1}^{K} \boldsymbol{X}_{qk} \boldsymbol{H}_{jqk} + \boldsymbol{Z}_j. \tag{3}$$

Let $\boldsymbol{X}_{qk} = \boldsymbol{S}_{qk} + \boldsymbol{B}_{qk}$, where $\boldsymbol{S}_{qk}$ is an arbitrary $N \times N$ data diagonal matrix, and $\boldsymbol{B}_{qk}$ is an $N \times N$ pilot diagonal matrix. Model (3) could be further rewritten as

$$\boldsymbol{Y}_j = \sum_{q=1}^{Q} \boldsymbol{T}_q \boldsymbol{C}_{jq} + \sum_{q=1}^{Q} \boldsymbol{A}_q \boldsymbol{C}_{jq} + \boldsymbol{Z}_j \tag{4}$$

where $\boldsymbol{T}_q = [\boldsymbol{S}_{q1}\boldsymbol{F}_{N,L}, \ldots, \boldsymbol{S}_{qK}\boldsymbol{F}_{N,L}]$, $\boldsymbol{A}_q = [\boldsymbol{B}_{q1}\boldsymbol{F}_{N,L}, \ldots, \boldsymbol{B}_{qK}\boldsymbol{F}_{N,L}]$, and $\boldsymbol{C}_{jq} = [\boldsymbol{C}_{jq1}^T, \ldots, \boldsymbol{C}_{jqK}^T]^T$.

III. Impact of Pilot Contamination on Classical Least Squares and Minimum Mean Square Error Algorithms in Multicell Multiuser Multiple-Input Multiple-Output Systems

Here, we shall derive analytical MSE expressions and investigate the effect of the PC on LS and MMSE channel estimation algorithms in multicell MU-MIMO systems.

A. LS Channel Estimation

The following assumptions are made: 1) Each subcarrier has the same power; 2) for different users in each cell, phase-shift orthogonal pilot sequences are used [19]; and 3) the same pilot sequences are reused in other cells. Thus, we can easily get $\boldsymbol{A}_j^\dagger \boldsymbol{A}_q = \boldsymbol{I}_{LK} (\boldsymbol{A}^\dagger = (\boldsymbol{A}^H \boldsymbol{A})^{-1} \boldsymbol{A}^H)$ and $\boldsymbol{A}_j^\dagger \boldsymbol{T}_q = \boldsymbol{0}_{LK}$, $1 \leq j, q \leq Q$, and the channel vector between the jth BS and K users in the jth cell is obtained by multiplying $\boldsymbol{A}_j^\dagger$ in both sides of (4), as follows:

$$\hat{\boldsymbol{C}}_{jj}^{\mathrm{LS}} = \boldsymbol{A}_j^\dagger \boldsymbol{Y}_j. \tag{5}$$

The MSE of the LS is described in the following theorem.

[1] Here, we assume that the K users in each cell is located uniformly, and they have almost the same distance to the BS in the jth cell. This assumption can make the analysis simple. In general, d_{jqkl} is a function of k.

Theorem 1: The MSE expression of the LS algorithm for multicell MU-MIMO systems in the presence of PC is given as follows:

$$\mathrm{MSE}_{\mathrm{LS}} = \frac{1}{L} \sum_{q \neq j}^{Q} d_{jq} + \frac{N}{P} \sigma^2. \tag{6}$$

Proof: See Appendix A.

Remark: In Theorem 1, it can be seen that the MSE for multicell MU-MIMO systems is composed of two terms, namely, the term introduced by the PC and the term caused by the noise. While the first term becomes zero in the case of a single cell, the second term can be suppressed by employing more pilot subcarriers. The first term can be improved as the length L of CIR is large. It can be also seen that the value of the first term cannot be reduced by using more pilots. However, this expression also indicates that appropriate pilot reuse or allocation techniques could be developed to reduce the impact of cross gains.

B. MMSE Channel Estimation

By employing the channel characteristics, MMSE usually obtains optimal estimation performance. Due to the high computational complexity in MMSE for MIMO systems, we just consider a simplified version by using an expectation–maximization iterative process proposed in [21]. The channel frequency vector between the jth BS and the kth user in the jth cell is given as follows:

$$\hat{\boldsymbol{H}}_{jjk}^{\mathrm{MMSE}} = \boldsymbol{R}_{HH} \left(\boldsymbol{R}_{HH} + \sigma^2 \left(\boldsymbol{X}_{jjk}^H \boldsymbol{X}_{jjk} \right)^{-1} \right)^{-1} \hat{\boldsymbol{H}}_{jjk}^{\mathrm{LS}} \tag{7}$$

where $\boldsymbol{R} = E[\boldsymbol{H}_{jjk} \boldsymbol{H}_{jjk}^H]$ is the correlation matrix of $\boldsymbol{H}_{jjk}$. Assuming normalized constellation power and equally probable constellation points and independent data symbols, matrix $(\boldsymbol{X}_{jjk}^H \boldsymbol{X}_{jjk})^{-1}$ could be replaced by $E[(\boldsymbol{X}_{jjk}^H \boldsymbol{X}_{jjk})^{-1}] = \beta \boldsymbol{I}_N$, where β is a constant, depending on the signal constellation (e.g., β equals 1 for quadrature phase shift keying (QPSK).

To avoid the matrix inversion, singular value decomposition (SVD) is applied; then, one obtains

$$\boldsymbol{R}_{HH} = \boldsymbol{U} \wedge \boldsymbol{U}^H \tag{8}$$

where $\boldsymbol{U}$ is a unitary matrix containing the singular vectors, and $\wedge$ is a diagonal matrix containing the singular values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ on its diagonal. Equation (7) is rewritten as

$$\hat{\boldsymbol{H}}_{jjk}^{\mathrm{MMSE}} = \boldsymbol{U} \triangle_p \boldsymbol{U}^H \hat{\boldsymbol{H}}_{jjk}^{\mathrm{LS}} \tag{9}$$

where $\triangle_p$ is a diagonal matrix with entries

$$\delta_n = \begin{cases} \frac{\lambda_n}{\lambda_n + \frac{\beta}{\mathrm{SNR}}}, & n = 1, 2, \ldots, p \\ 0, & n = p+1, \ldots, N. \end{cases} \tag{10}$$

The MSE of the MMSE is presented in the following theorem.

Theorem 2: The MSE expression of the MMSE algorithm for multicell MU-MIMO systems in the presence of PC is given as follows:

$$\begin{aligned} \mathrm{MSE}_{\mathrm{MMSE}} &= \frac{1}{N}\sum_{q\neq j}^{Q}\sum_{n=1}^{p}\lambda_n\delta_n^2 + \frac{\sigma^2}{N}\sum_{n=1}^{p}\lambda_n \\ &+ \frac{1}{N}\left(\sum_{n=1}^{p}(\delta_n-1)^2\lambda_n + \sum_{n=p+1}^{N}\lambda_n\right) \\ &< \frac{1}{N}\sum_{q\neq j}^{Q} d_{jq} + \frac{\sigma^2+1}{N}. \end{aligned} \tag{11}$$

Proof: See Appendix B.

Remark: In Theorem 2, it can be seen that the MSE for MMSE is introduced by two factors. The values of both terms are much smaller than for LS. The first term can be suppressed as the number of subcarriers increases. However, the number of subcarriers could not be unlimitedly increased because OFDM that uses a large number of subcarriers will be very sensitive to the impact of the reuse of nonorthogonal pilots. Similar to the MSE expression of the LS algorithm, a large number of pilot subcarriers cannot decrease the MSE caused by PC. However, the clue for mitigating the impact of cross gains is to develop appropriate pilot reuse or allocation schemes.

IV. Design and Analysis of Space-Alternating Generalized Expectation–Maximization-Based H-inf Algorithm in Multicell Multiuser Multiple-Input Multiple-Output Systems

Earlier, we have shown that the MMSE algorithm can obtain optimal performance by using prior information and better suppression to PC. Although the use of SVD of channel correlation matrix is able to reduce the number of multiplications with negligible performance loss, its complexity is still quite high since obtaining the SVD itself has high computational complexity on the order of $\mathcal{O}(N^3)$. Here, we introduce the H-inf algorithm, which were proposed in [22] and [23], to multicell MU-MIMO systems.

A. H-inf Channel Estimation

As an alternative to the classical MMSE estimation, an H-inf filter can achieve an acceptable estimation performance without accurate knowledge of the statistical information of the involved signals. The idea of the H-inf filtering is to construct a filter that guarantees the H-inf norm of the estimation error is less than a prescribed positive value.

As for multicell MU-MIMO systems, the idea of the H-inf is to find an estimation method so that the ratio between the whole channel estimation error (between the jth BS and K users in each cell) and the input noise/interference is less than a prescribed threshold. Given a positive scalar factor s, the H-inf estimator for each received OFDM symbol needs to satisfy the following objective function [22], [23]:[2]

$$\sup_{\boldsymbol{Z}_j} = \frac{\|\hat{\boldsymbol{C}}_j - \boldsymbol{C}_j\|_{\boldsymbol{W}}^2}{\|\boldsymbol{Z}_j\|^2} < s \tag{12}$$

where $\|\hat{\boldsymbol{C}}_j - \boldsymbol{C}_j\|_{\boldsymbol{W}}^2 = (\hat{\boldsymbol{C}}_j - \boldsymbol{C}_j)^H\boldsymbol{W}(\hat{\boldsymbol{C}}_j - \boldsymbol{C}_j)$; $\hat{\boldsymbol{C}}_j$ is a $LQK \times 1$ vector, denoting the channel response vector to be estimated; $\boldsymbol{C}_j = [\boldsymbol{C}_{j1}^T, \ldots, \boldsymbol{C}_{jQ}^T]^T$; $\boldsymbol{C}_{jq} = [\boldsymbol{C}_{jq1}^T, \ldots, \boldsymbol{C}_{jqK}^T]^T$; and $\boldsymbol{W} > \boldsymbol{0}$ is a weighting matrix. The H-inf channel estimation in multicell MU-MIMO systems can be described as [22], [23]

$$\hat{\boldsymbol{C}}_j = \boldsymbol{\eta}_j\boldsymbol{\varepsilon}_j^{-1}\boldsymbol{T}^\dagger\boldsymbol{Y}_j \tag{13}$$

where $\boldsymbol{T} = [\boldsymbol{T}_1, \ldots, \boldsymbol{T}_Q]$, $\boldsymbol{T}_q = [\boldsymbol{T}_{q1}, \ldots, \boldsymbol{T}_{qK}]$, $\boldsymbol{T}_{qk} = \boldsymbol{X}_{qk}\boldsymbol{F}_{N,L}$, and $\boldsymbol{\varepsilon}_j = \boldsymbol{M}_{1,1} + \boldsymbol{M}_{1,2}\boldsymbol{\xi}_j$ and $\boldsymbol{\eta}_j = \boldsymbol{M}_{2,1} + \boldsymbol{M}_{2,2}\boldsymbol{\xi}_j$, are both $LQK \times LQK$ matrices. $\boldsymbol{\xi}_j$ is a $LQK \times 1$ vector, satisfying $\|\boldsymbol{\xi}_j\|_\infty = \max(|\xi_1|, \ldots, |\xi_{LQK}|) < 1$, and $\xi_1 = \cdots = \xi_{LQK}$. $\boldsymbol{M}_{1,1}$, $\boldsymbol{M}_{1,2}$, and $\boldsymbol{M}_{2,1}$, $\boldsymbol{M}_{2,2}$ can be expressed as

$$\begin{aligned} \boldsymbol{M}_{1,1} &= \boldsymbol{\Omega}\boldsymbol{R}^{\frac{1}{2}} + \boldsymbol{R}^{-\frac{1}{2}} \\ \boldsymbol{M}_{1,2} &= s^{-\frac{1}{2}}\boldsymbol{\Omega}\boldsymbol{W}^{\frac{1}{2}} \\ \boldsymbol{M}_{2,1} &= \boldsymbol{\Omega}\boldsymbol{R}^{\frac{1}{2}} \\ \boldsymbol{M}_{2,2} &= s^{-\frac{1}{2}}\boldsymbol{\Omega}\boldsymbol{W}^{\frac{1}{2}} - s^{\frac{1}{2}}\boldsymbol{W}^{\frac{1}{2}} \end{aligned} \tag{14}$$

where $\boldsymbol{R} = \boldsymbol{T}^\dagger\boldsymbol{T} = \boldsymbol{I}_{LQK}$ if QPSK is adopted, $\boldsymbol{\Omega} = \boldsymbol{\Omega}_1\boldsymbol{\Omega}_2^{1/2} - \boldsymbol{\Omega}_2$, $\boldsymbol{\Omega}_2 = (\boldsymbol{R} - s^{-1}\boldsymbol{W})^{-1}$, and $\boldsymbol{\Omega}_1$ can be easily obtained by the canonical factorization of $\boldsymbol{I}_{LQK} + \boldsymbol{\Omega}_2$.

B. H-inf Channel Estimation via SAGE Process

A direct solution to (13) will result from intense calculation of the matrix inversion and multiplication operations for each OFDM symbol of all users in Q cells over L paths, and the complexity is on the order of $\mathcal{O}(L^3Q^3K^3)$. In the case of large values of L, K, and Q, computational complexity load will be high.

In multicell MU-MIMO systems, propagation vectors between the BS antenna arrays and different terminals often could be considered uncorrelated [4]. Since the SAGE can decompose the spatially multiplexed channels, we can apply this iterative algorithm to deal with the problem of high complexity [20]. Generally, the SAGE process is developed to avoid matrix inversion of the ML estimator; therefore, we first assess the feasibility by applying SAGE. Equation (13) can be rewritten as follows:

$$\begin{aligned} \hat{\boldsymbol{C}}_j &= \boldsymbol{\eta}_j\boldsymbol{\varepsilon}_j^{-1}\boldsymbol{T}^\dagger\boldsymbol{Y}_j \\ &= \boldsymbol{\gamma}\hat{\boldsymbol{C}}_j^{\mathrm{ML}} \end{aligned} \tag{15}$$

[2]The numerator of (12) is considered to be the whole estimation error between the jth BS and K users in each cell. Thus, the denominator of (12) will be AWGN $\boldsymbol{Z}_j$. However, if the local estimation error is considered, (e.g., between the jth and K users in the qth cell), the signal, except for that from the qth cell, will be the interference, which will finally change the establishment of the objective function.

where $\gamma = \eta_j \varepsilon_j^{-1}$. Equation (15) can be interpreted as a filter matrix γ applied to the ML estimation, indicating some links between the H-inf and ML estimators. Thus, we can develop an H-inf estimator by combining the SAGE process. Instead of solving (13) directly, the SAGE algorithm converts a multicell MU-MIMO channel estimation problem into a series of single-cell single-user SISO channel estimation problems, making the dimensions of Ω, W, and R involved in the computation of ε_j, η_j much smaller. Thus, the calculation is simplified drastically.

The SAGE-based H-inf estimation can be iteratively implemented as follows.

- Initialization:
  For $q = 1, \ldots, Q$,
  For $k = 1, \ldots, K$

$$\hat{\boldsymbol{Y}}_{jqk}^{(0)} = \boldsymbol{T}_{qk} \varepsilon_{jqk} \eta_{jqk}^{-1} \hat{\boldsymbol{C}}_{jqk}^{(0)} \tag{16}$$

  where ε_{jqk} and η_{jqk} of dimension $L \times L$ are the simplified versions of ε_j and η_j, respectively. The initial value of channel estimation $\hat{\boldsymbol{C}}_{jqk}^{(0)}$ is 1_L, where 1_L is an $L \times 1$ vector whose elements are all 1.
- At the ith iteration $(i = 0, 1, 2, \ldots)$:
  For $k = 1 + [i \mod K]$, we have

$$\hat{\boldsymbol{\Pi}}_{jqk}^{(i)} = \hat{\boldsymbol{Y}}_{jqk}^{(i)} + \left[\boldsymbol{Y}_j - \sum_{k=1}^{K} \hat{\boldsymbol{Y}}_{jqk}^{(i)} \right] \tag{17}$$

$$\hat{\boldsymbol{C}}_{jqk}^{(i+1)} = \arg\min_{\boldsymbol{C}_{jqk}} \left\{ \left\| \hat{\boldsymbol{\Pi}}_{jqk}^{(i)} - \boldsymbol{T}_{qk} \varepsilon_{jqk} \eta_{jqk}^{-1} \boldsymbol{C}_{jqk} \right\|^2 \right\}. \tag{18}$$

  By solving (18), we can obtain

$$\hat{\boldsymbol{C}}_{jqk}^{(i+1)} = \eta_{jqk} \varepsilon_{jqk}^{-1} \boldsymbol{T}_{qk}^{\dagger} \hat{\boldsymbol{\Pi}}_{jqk}^{(i)} \tag{19}$$

$$\hat{\boldsymbol{Y}}_{jqk}^{(i+1)} = \boldsymbol{T}_{qk} \varepsilon_{jqk} \eta_{jqk}^{-1} \hat{\boldsymbol{C}}_{jqk}^{(i+1)} \tag{20}$$

  while for $1 \le k' \le K$ and $k' \ne k$

$$\hat{\boldsymbol{Y}}_{jqk}^{(i+1)} = \hat{\boldsymbol{Y}}_{jqk}^{(i)}. \tag{21}$$

C. Performance Analysis

1) Analysis of Matrix γ: To find a solution for the H-inf, we assume $\boldsymbol{R} - s^{-1}\boldsymbol{W} > 0$ [22], [23], where $\boldsymbol{R}$ is an identity matrix because QPSK is adopted,[3] s is a positive scalar factor, and $\boldsymbol{W}$ is also a diagonal matrix that have equal dimensions. Thus, $\boldsymbol{M}_{1,1}$, $\boldsymbol{M}_{1,2}$, $\boldsymbol{M}_{2,1}$, and $\boldsymbol{M}_{2,2}$ are all diagonal matrices, respectively. Finally, matrix γ is a real diagonal matrix with equal diagonal elements.

Since the diagonal matrix γ is needed to estimate the performance of the H-inf, we will find the relation between γ and the identity matrix. First, it is assumed that

$$\gamma < \boldsymbol{I}_{LQK}. \tag{22}$$

[3]Note that $\boldsymbol{R}$ will not be an identity matrix if 16-QAM, 64-QAM, or other modulations are adopted. However, γ is always a diagonal matrix. The proposed algorithm is valid for the different modulations.

To satisfy (22), one has $\varepsilon - \eta > 0$. By applying (14), we can get

$$\begin{aligned} \varepsilon - \eta &= (\boldsymbol{M}_{1,1} + \boldsymbol{M}_{1,2}\xi_j) - (\boldsymbol{M}_{2,1} + \boldsymbol{M}_{2,2}\xi_j) \\ &= \boldsymbol{R}^{-\frac{1}{2}} + s^{\frac{1}{2}} \boldsymbol{W}^{-\frac{1}{2}} \xi_j > 0. \end{aligned} \tag{23}$$

Therefore, our hypothesis is valid. Intuitively, when $\boldsymbol{W}$ is fixed, a smaller s is made, a smaller γ is obtained, and a better performance is achieved, which is the intrinsic characteristic of the H-inf algorithm, as will be discussed in the following.

2) Impact of PC on H-inf: Since the estimation errors in cells are independent of each other, we analyze the channels from the K users in the jth cells. The following assumptions are made: 1) All subcarriers have equal power; 2) phase-shift orthogonal pilot sequences are used for different users within each cell; and 3) the same pilot sequences are reused in other cells.

The channel estimation of the H-inf can be rewritten as

$$\begin{aligned} \hat{\boldsymbol{C}}_{jj}^{\mathrm{H-inf}} &= \gamma \boldsymbol{T}_j^{\dagger} \boldsymbol{Y}_j \\ &= \gamma \boldsymbol{T}_j^{\dagger} \sum_{q \ne j}^{Q} \boldsymbol{T}_q \boldsymbol{C}_{jq} + \gamma \boldsymbol{C}_{jj} + \gamma \boldsymbol{T}_j^{\dagger} \boldsymbol{Z}_j. \end{aligned} \tag{24}$$

The MSE of the H-inf is given in the follow theorem.

Theorem 3: The MSE expression of the H-inf algorithm for multicell MU-MIMO systems in the presence of PC is given as follows:

$$\mathrm{MSE}_{H-\mathrm{inf}} = \underbrace{\frac{1}{L} r_{nn}^2 \sum_{q \ne j}^{Q} d_{jq}}_{\mathrm{PC}} + \underbrace{\frac{1}{L} r_{nn}^2 \sigma^2 + \frac{1}{L}(1 - r_{nn})^2}_{\mathrm{noise}}. \tag{25}$$

Proof: See Appendix C.

Remark: In Theorem 3, it can be seen that the MSE includes two terms. The MSE caused by PC or noise cannot be decreased by increasing the number of pilots. The MSE of the first term will be decreased as the length of CIR increases; however, this is limited by the length of CP of OFDM. The first term can be improved by decreasing the value of r_{nn} (or s). The appropriate pilot reuse or allocation techniques should be developed to improve the impact of cross gains.

To make a comparison of the MSE between the H-inf and the ML, the following analysis is presented. If $\gamma = \boldsymbol{I}_{LQK}$, the H-inf channel estimation will be converted into ML. Similarly, the MSE expression of the ML algorithm can be obtained in multicell MU-MIMO systems in presence of PC, as follows.

Theorem 4: The MSE expression of the ML algorithm for multicell MU-MIMO systems in the presence of PC is given as follows:

$$\mathrm{MSE}_{\mathrm{ML}} = \frac{1}{L} \sum_{q \ne j}^{Q} d_{jq} + \sigma^2. \tag{26}$$

Proof: See Appendix D.

TABLE I
COMPLEXITY OF CHANNEL ESTIMATION ALGORITHMS

Algorithm	Number of operations per OFDM symbol
LS	$(2(LK)^2 + (LK))N + (LK)^3$
MMSE	$(2(LK)^2 + (LK))N + (LK)^3 + 2N^3 + LN^2 + LN$
H-inf	$(3(LKQ)^2 + LKQ)N + 3(LKQ)^3$
SAGE-based	$K_{it}((3L^2 + L)N + 3L^3)$

Remark: According to Theorem 4, the MSE caused by PC for H-inf is obviously less than for ML. For the MSE caused by the noise, since $0 < r_{nn} < 1$, and σ^2 is a set unit variance, d_{jj} denotes direct gains within the jth cell, and it is assigned to be unity; therefore, the upper bound of the MSE caused by the noise can be given as follows:

$$\mathrm{MSE}_{N,\mathrm{H-inf}} < \frac{1}{L}\sigma^2. \tag{27}$$

Referring to Theorems 1, 3, and 4, the MSE comparison between LS, ML, and H-inf algorithms can be summarized as follows:

$$\mathrm{MSE}_{\mathrm{H-inf}} < \mathrm{MSE}_{\mathrm{ML}} < \mathrm{MSE}_{\mathrm{LS}}. \tag{28}$$

3) Complexity Analysis: Considering the number of complex multiplications for each OFDM symbol as a complexity metric, the inversion of an $n \times n$ matrix requires n^3 operations, the pseudoinverse of an $n \times r$ matrix requires $2r^2n + r^3$ operations, and the product of an $m \times r$ matrix with an $r \times n$ matrix requires mrn operations. Let K_{it} denote the number of iterations that should not be too large due to the superior convergence property of SAGE [20]. A comparison of complexity between the LS, MMSE, and proposed H-inf algorithms is given in Table I. As expected, the H-inf estimation has less complexity than the MMSE algorithm, and the complexity can be further reduced by using the SAGE iterative process.

V. SIMULATION RESULTS

We consider a multicell MU-MIMO system with M antennas at each BS to investigate the impact of PC on the MSE of channel estimation algorithms. It is assumed that $Q = 3$ cells, and $K = 10$ users in each cell. The phase-shifted orthogonal pilot sequences used in the first cell is reused in the second and third cells. Thus, we consider a scenario where pilot sequences are reused. Furthermore, for all k, $d_{jqk} = 1$ (direct gain) if $j = q$, and $d_{jqk} = a$ (cross gain) if $j \neq q$. Since the cell layout and shadowing are captured by using the constant d_{jqk}, PC is handled by adjusting the cross gains. For OFDM symbols with N subcarriers, the length of CP is 16, and QPSK is used. For the number of iterations in SAGE iterative process, we choose $K_{\mathrm{it}} = 3$.

A. Performance of Two Classical Estimation Algorithms

Here, the representative performances for classical LS and MMSE algorithms in multicell MU-MIMO systems are shown.

Fig. 2 shows the MSE performance of LS and MMSE algorithms versus the SNR for different values of L at $M = 50$, $a = 0.6$, and $N = 128$. It is shown in the figure that MMSE is more resistant to PC than LS. This is because LS just utilizes few pilot subcarriers, whereas MMSE makes use of more prior information. The performance of LS can be improved by increasing the length of CIR. In addition, the performance of the MMSE is not related to the length of CIR. The analytical results given by (6) and (11) are very close to the simulation results.

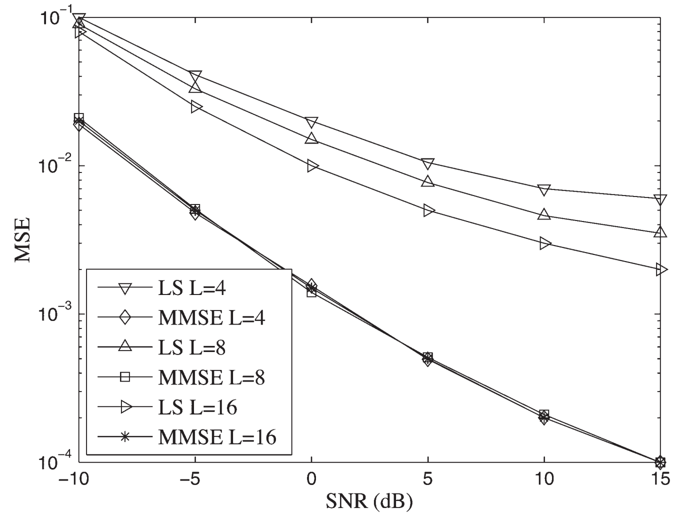


Fig. 2. MSE versus SNR with different L at $M = 50$, $a = 0.6$, and $N = 128$.

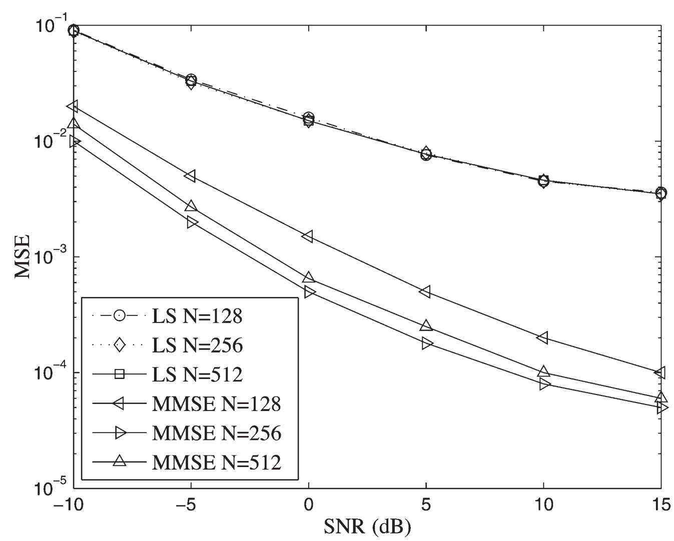


Fig. 3. MSE versus the SNR with different N at $M = 50$ and $a = 0.6$.

The MSE performance of LS and MMSE algorithms is shown in Fig. 3 for different values of N in the case of $a = 0.6$ as a function of SNR. It is shown in the figure that the performance of the LS is independent of the number of subcarriers, whereas the MSE of MMSE can be improved by increasing the number of subcarriers. Since the number of subcarriers is larger, the subcarrier spacing becomes smaller; the systems with a larger number of subcarriers are more sensitive to PC.

Fig. 4 shows that the MSE performance of the LS and MMSE algorithms as a function of a for different values of L and N at SNR = 5 dB. It is shown that the performance of LS and MMSE generally degrades due to increasing cross gain a, implying that the interference caused by PC is more serious. Meanwhile, the performance improves when $a < 0.7$.

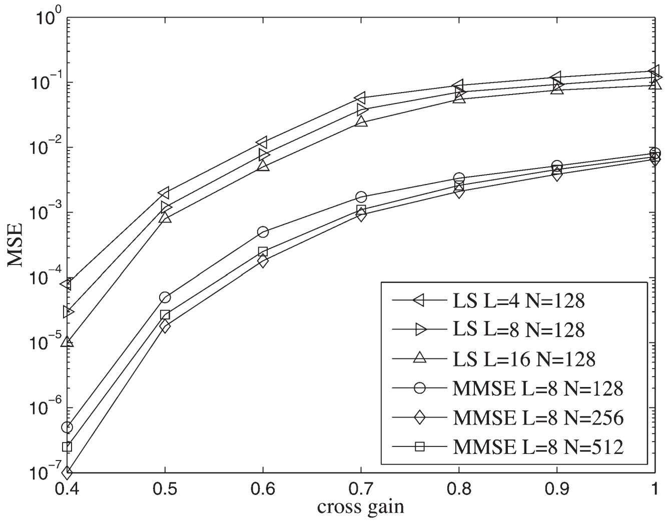


Fig. 4. MSE versus a with different L and N at $M = 50$, and SNR $= 5$ dB.

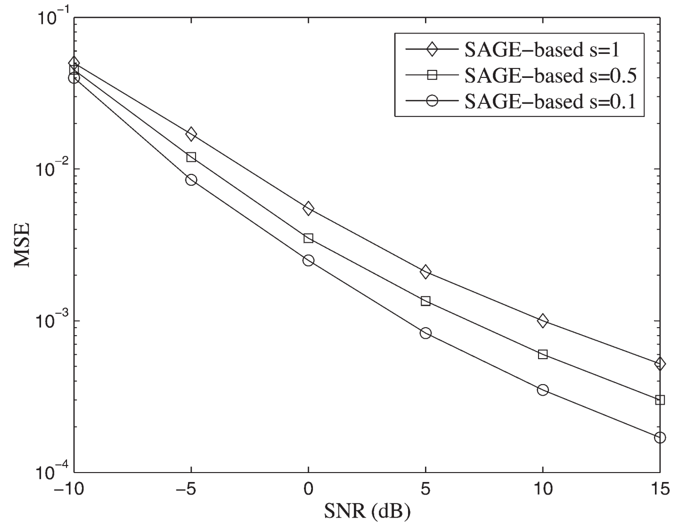


Fig. 6. MSE versus SNR with different s at $a = 0.6$, $M = 50$, and $L = 8$.

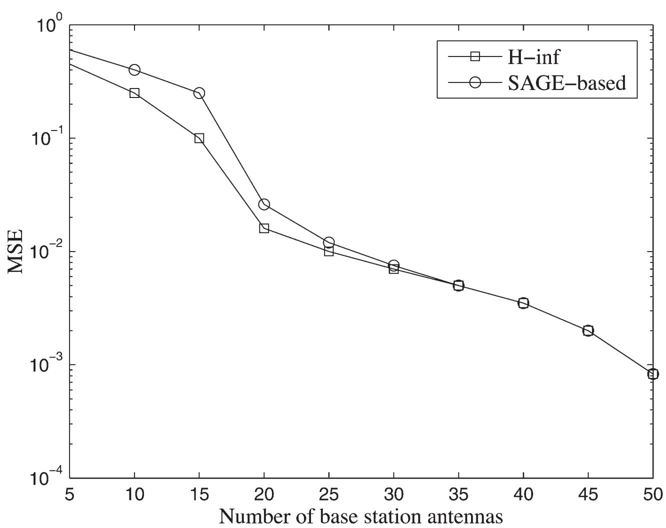


Fig. 5. MSE versus M at $a = 0.6$, $L = 8$, and $s = 0.1$.

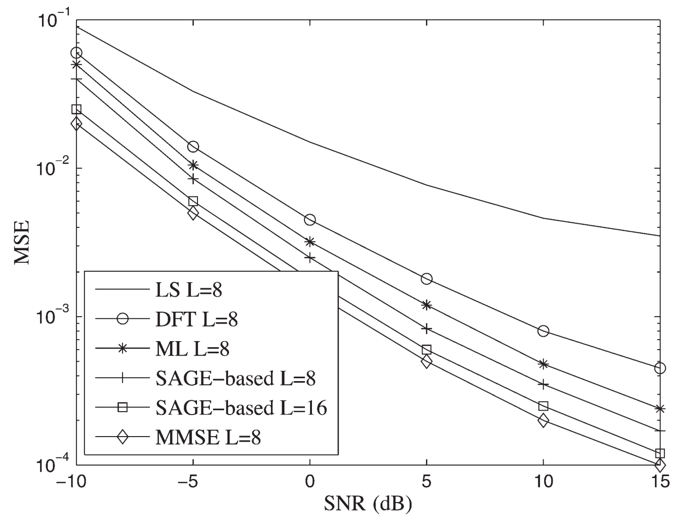


Fig. 7. MSE versus SNR with different L at $a = 0.6$, $M = 50$, and $s = 0.1$.

It is worth noting that the performance of the MMSE with 512 subcarriers is worse than with 256 subcarriers since OFDM is vulnerable to PC, particularly in the case of a large number of subcarriers.

B. Performance of the Proposed H-inf Estimation Algorithm

Here, the performance of the proposed H-inf algorithm in multicell MU-MIMO systems is plotted.

Fig. 5 shows the MSE performance of H-inf- and SAGE-based algorithms versus M for $a = 0.6$, $L = 8$, and $s = 0.1$. In this figure, it is shown that the performance of the SAGE-based algorithm is almost the same as that of H-inf when $M > 30$, due to the fact that, when the number of antennas at each BS is large, the estimation error caused by local convergence of the SAGE iterative process will vanish.

Fig. 6 shows the MSE performance of the SAGE-based algorithm versus SNR for different values of s for $a = 0.6$, $M = 50$, and $L = 8$. It is shown that the MSE performance is gradually enhanced when s decreases. According to (12) in Section IV, the intrinsic characteristics of the H-inf algorithm can make the estimation error less than a prescribed bound. As expected, the performance is much improved by decreasing s.

In the following, the DFT (or called transform domain) algorithm will first be introduced as a reference because it can be used to reduce the noise component effectively and has been widely used in OFDM-based systems. Furthermore, ML algorithm will also be considered to make a comparison with the proposed algorithm.

The MSE performance of LS, DFT, ML, MMSE, and SAGE-based algorithms is shown in Fig. 7 versus SNR for different values of L at $a = 0.6$, $M = 50$, and $s = 0.1$. It is shown that the SAGE-based algorithm is more resistant to PC than LS, DFT, and ML. This is because the SAGE-based algorithm utilizes other information, such as, transmitted data and scalar factor s. It is also shown that the performance of SAGE-based algorithm can be improved by increasing the length of CIR, as shown in the analytical expression in (25). Furthermore, when $L = 16$, its performance is close to the optimal MMSE, which can be considered a lower bound of the MSE. In terms of

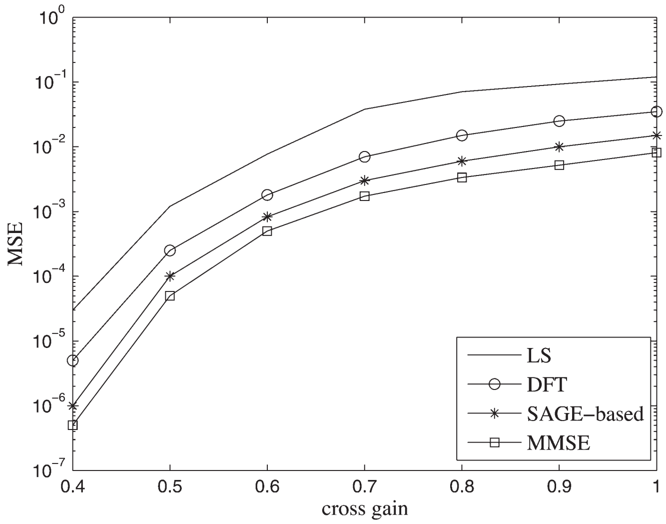


Fig. 8. MSE versus a at SNR $= 5$ dB, $M = 50$, $s = 0.1$, and $L = 8$.

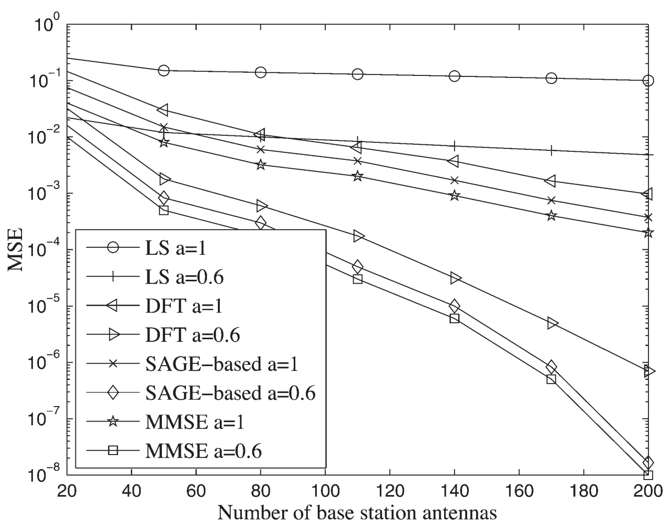


Fig. 9. MSE versus M with different a at $L = 8$ and SNR $= 5$ dB.

the complexity presented in Section IV-C, the SAGE-based algorithm can be considered a substitute for MMSE in multicell MU-MIMO systems.

Fig. 8 shows the MSE performance of LS, DFT, MMSE, and SAGE-based algorithms for SNR $= 5$ dB, $M = 50$, $s = 0.1$, and $L = 8$ as a function of a. It is shown that the performance of all the algorithms degrades much when cross gain a is large. Furthermore, their performances have obvious improvement when the value of a decreases. Finally, the MSE performance of LS, DFT, MMSE, and SAGE-based algorithms is shown in Fig. 9 as a function of M with different values of a for SNR $= 5$ dB and $L = 8$. It is shown that the performance of LS is vulnerable to the impact of PC no matter what the value of M is. However, the performance of DFT, MMSE, and SAGE-based algorithms improve significantly as M increases when $a = 0.6$. The impact of PC is big when cross gains $a = 1$.

VI. Conclusion

In this paper, we have analytically investigated the impact of PC on the several pilot-based channel estimation algorithms, including classical LS, MMSE algorithms, and our proposed H-inf algorithms in multicell MU-MIMO systems under a realistic system model that considers imperfect channel estimation, PC, multicarrier, and multipath channels. Analytical expressions were derived, and comparisons were made. It has been shown that, of all the algorithms, the optimal MMSE is most resistant to PC with high complexity. By slightly increasing the number of OFDM subcarriers, PC suppression can be achieved in the MMSE. In addition, by increasing the number of pilot subcarriers for all channel estimation algorithms, PC cannot be mitigated. For the proposed H-inf algorithms, proper length increment of CIR is helpful for the suppression of PC. Simulation results have shown that the proposed H-inf algorithm has almost the same performance as MMSE, and it leads to better suppression to PC than LS, DFT, and ML. In addition, the H-inf via the SAGE iterative process does not introduce any performance loss when the number of antennas is large at each BS in multicell MU-MIMO systems.

Appendix A
Mean Square Error of Least Squares

According to (5), the channel estimation error of LS is expressed as

$$\begin{aligned} \boldsymbol{S}_{\mathrm{LS}} &= \hat{\boldsymbol{C}}_{jj}^{\mathrm{LS}} - \boldsymbol{C}_{jj} \\ &= \sum_{q\neq j}^{Q} \boldsymbol{C}_{jq} + \boldsymbol{A}_j^{\dagger}\boldsymbol{Z}_j \\ &= \boldsymbol{S}_{\mathrm{PC,\,LS}} + \boldsymbol{S}_{N,\,\mathrm{LS}} \end{aligned} \tag{A.1}$$

where $\boldsymbol{S}_{\mathrm{PC,LS}}$ and $\boldsymbol{S}_{N,\,\mathrm{LS}}$ denote the errors introduced by the PC and the noise, respectively, and they are independent of each other.

For the PC term, it is assumed that the channel vectors of different users in different cells have the same characteristics. Furthermore, channel vectors between BS antenna arrays and different terminals are uncorrelated. Thus, the MSE of PC for LS is given as follows:

$$\begin{aligned} \mathrm{MSE}_{\mathrm{LS}} &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{\mathrm{PC,\,LS}}\boldsymbol{S}_{\mathrm{PC,\,LS}}^{H}\right]\right\} \\ &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\sum_{q\neq j}^{Q}\boldsymbol{C}_{jq}\left(\sum_{q\neq j}^{Q}\boldsymbol{C}_{jq}\right)^{H}\right]\right\} \\ &= \frac{1}{LK}\sum_{q\neq j}^{Q}\sum_{k=1}^{K}\mathrm{tr}\{\boldsymbol{R}_{CC}\} \end{aligned} \tag{A.2}$$

where $\boldsymbol{R}_{CC} = E[\boldsymbol{C}_{jqk}\boldsymbol{C}_{jqk}^{H}] = \boldsymbol{D}_{jqk}^{1/2}E[\boldsymbol{G}_{jqk}\boldsymbol{G}_{jqk}^{H}]\boldsymbol{D}_{jqk}^{1/2}$, and the MSE of PC for LS can be simplified as

$$\begin{aligned} \mathrm{MSE}_{\mathrm{LS}} &= \frac{1}{LK}\sum_{q\neq j}^{Q} d_{jq}\left(\sum_{k=1}^{K}\sum_{l=0}^{L-1} g_{jqkl}^{2}\right) \\ &= \frac{1}{L}\sum_{q\neq j}^{Q} d_{jq}. \end{aligned} \tag{A.3}$$

It can be seen that the impact of PC can be significant if d_{jq} (cross gain) between cells are of the same order in terms of d_{jj} (direct gain) within the same cell.

For the noise term, the MSE is given as follows:

$$\begin{aligned}\mathrm{MSE}_{\mathrm{LS}} &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{N,\mathrm{LS}}\boldsymbol{S}_{N,\mathrm{LS}}^H\right]\right\}\\ &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{A}_j^\dagger \boldsymbol{Z}_j\left(\boldsymbol{A}_j^\dagger \boldsymbol{Z}_j\right)^H\right]\right\}\\ &= \frac{1}{LK}\mathrm{tr}\left\{\boldsymbol{A}_j^\dagger \boldsymbol{R}_{ZZ}\left(\boldsymbol{A}_j^\dagger\right)^H\right\} \qquad \text{(A.4)}\end{aligned}$$

where $\boldsymbol{R}_{ZZ} = E[\boldsymbol{Z}_j\boldsymbol{Z}_j^H] = \sigma^2\boldsymbol{I}_{LK}$. According to $\boldsymbol{A}_j^H\boldsymbol{A}_j = (P/N)\boldsymbol{I}_{LK}$, the MSE of the noise term for LS can be rewritten as

$$\begin{aligned}\mathrm{MSE}_{\mathrm{LS}} &= \frac{\sigma^2}{LK}\mathrm{tr}\left\{\left(\boldsymbol{A}_j^H\boldsymbol{A}_j\right)^{-1}\right\}\\ &= \frac{N}{P}\sigma^2 \qquad \text{(A.5)}\end{aligned}$$

where P is the number of pilot subcarriers in each OFDM symbol.

Appendix B
Mean Square Error of Minimum Mean Square Error

According to (9), the channel estimation error of MMSE is given as follows:

$$\begin{aligned}\boldsymbol{S}_{\mathrm{MMSE}} &= \hat{\boldsymbol{H}}_{jjk}^{\mathrm{MMSE}} - \boldsymbol{H}_{jjk}\\ &= \boldsymbol{U}\triangle_p\boldsymbol{U}^H\sum_{q\neq j}^{Q}\boldsymbol{H}_{jqk} + \boldsymbol{U}\triangle_p\boldsymbol{U}^H\boldsymbol{F}_{N,L}\boldsymbol{A}_{jk}^\dagger\boldsymbol{Z}_j\\ &\quad + \boldsymbol{U}\left(\begin{bmatrix}\triangle_p & 0\\ 0 & 0\end{bmatrix} - \boldsymbol{I}_N\right)\boldsymbol{U}^H\boldsymbol{H}_{jjk}\\ &= \boldsymbol{S}_{\mathrm{PC,\,MMSE}} + \boldsymbol{S}_{N,\,\mathrm{MMSE}} \qquad \text{(B.1)}\end{aligned}$$

where $\boldsymbol{S}_{\mathrm{PC,\,MMSE}}$ and $\boldsymbol{S}_{N,\mathrm{MMSE}}$ denote the errors introduced by the PC and the noise, respectively.

For the PC term, the channel estimation error for the kth user in the jth BS is

$$\begin{aligned}&\mathrm{MSE}_{\mathrm{PC,\,MMSE}}\\ &= \frac{1}{N}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{\mathrm{PC,\,MMSE}}\boldsymbol{S}_{\mathrm{PC,\,MMSE}}^H\right]\right\}\\ &= \frac{1}{N}\mathrm{tr}\left\{E\left[\boldsymbol{U}\triangle_p\boldsymbol{U}^H\sum_{q\neq j}^{Q}\boldsymbol{H}_{jqk}\left(\boldsymbol{U}\triangle_p\boldsymbol{U}^H\sum_{q\neq j}^{Q}\boldsymbol{H}_{jqk}\right)^H\right]\right\}\\ &= \frac{1}{N}\sum_{q\neq j}^{Q}\mathrm{tr}\left\{\boldsymbol{U}\triangle_p\boldsymbol{U}^H\boldsymbol{R}_{HH}(\boldsymbol{U}\triangle_p\boldsymbol{U}^H)^H\right\}. \qquad \text{(B.2)}\end{aligned}$$

To simplify the expression, we use the following properties of the trace operator $\mathrm{tr}\{\boldsymbol{DAD}\} = \sum_k a_{kk}d_k^2$ in which $\boldsymbol{D}$ is a diagonal matrix with the elements d_k on its diagonal and $\boldsymbol{A}$ (not necessarily a diagonal elements) has diagonal elements a_{kk}.

According to (B.2), the MSE becomes

$$\mathrm{MSE}_{\mathrm{PC,\,MMSE}} = \frac{1}{N}\sum_{q\neq j}^{Q}\sum_{n=1}^{p}\lambda_n\delta_n^2. \qquad \text{(B.3)}$$

Since the channel power (variance) is normalized in both frequency domain and time domain, (B.3) can be limited as

$$\begin{aligned}\mathrm{MSE}_{\mathrm{PC,\,MMSE}} &< \frac{1}{N}\sum_{q\neq j}^{Q}\sum_{n=1}^{p}\lambda_n = \frac{1}{N}\sum_{q\neq j}^{Q}\mathrm{tr}\{\boldsymbol{R}_{CC}\}\\ &= \frac{1}{N}\sum_{q\neq j}^{Q}\sum_{l=0}^{L-1}d_{jq}g_{jql}^2 = \frac{1}{N}\sum_{q\neq j}^{Q}d_{jq}. \qquad \text{(B.4)}\end{aligned}$$

Equation (B.4) is similar to (A.2), in which the impact of PC also depends on the value of d_{jq}.

For the noise term, according to (B.1), the MSE can be given as follows:

$$\begin{aligned}&\mathrm{MSE}_{N,\,\mathrm{MMSE}}\\ &= \frac{1}{N}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{N,\,\mathrm{MMSE}}\boldsymbol{S}_{N,\,\mathrm{MMSE}}^H\right]\right\}\\ &= \frac{1}{N}\mathrm{tr}\left\{\triangle_p\boldsymbol{U}^H\boldsymbol{F}_{N,L}\boldsymbol{R}_{ZZ}\boldsymbol{F}_{N,L}^H\boldsymbol{U}\triangle_p^H\right\}\\ &\quad + \frac{1}{N}\mathrm{tr}\left\{\left(\begin{bmatrix}\triangle_p & 0\\ 0 & 0\end{bmatrix} - \boldsymbol{I}_N\right)\Lambda\left(\begin{bmatrix}\triangle_p & 0\\ 0 & 0\end{bmatrix} - \boldsymbol{I}_N\right)^H\right\}\\ &= \frac{\sigma^2}{N}\sum_{n=1}^{p}\lambda_n + \frac{1}{N}\left(\sum_{n=1}^{p}(1-\delta)^2\lambda_n + \sum_{n=p+1}^{N}\lambda_n\right). \qquad \text{(B.5)}\end{aligned}$$

Equation (B.5) can be also constrained as follows:

$$\begin{aligned}\mathrm{MSE}_{\mathrm{MMSE}} &< \frac{\sigma^2}{N}\sum_{n=1}^{p}\lambda_n + \frac{1}{N}\sum_{n=1}^{p}\lambda_n\\ &= \frac{\sigma^2+1}{N}\sum_{n=1}^{p}\lambda_n = \frac{\sigma^2+1}{N}\mathrm{tr}\{\boldsymbol{R}_{CC}\}\\ &< \frac{\sigma^2+1}{N}d_{jq} < \frac{\sigma^2+1}{N}. \qquad \text{(B.6)}\end{aligned}$$

Appendix C
Mean Square Error of H-inf

Given (24), the channel estimation error in the jth cell can be given as follows:

$$\begin{aligned}\boldsymbol{S}_{\mathrm{H-inf}} &= \hat{\boldsymbol{H}}_{jj}^{\mathrm{H-inf}} - \boldsymbol{H}_{jj}\\ &= \gamma\boldsymbol{T}_j^\dagger\sum_{q\neq j}^{Q}\boldsymbol{T}_q\boldsymbol{C}_{jq} + \gamma\boldsymbol{T}_j^\dagger\boldsymbol{Z}_j - (\boldsymbol{I}_{LQK} - \gamma)\boldsymbol{C}_{jj}\\ &= \boldsymbol{S}_{\mathrm{PC,\,H-inf}} + \boldsymbol{S}_{N,\,\mathrm{H-inf}} \qquad \text{(C.1)}\end{aligned}$$

where $\boldsymbol{S}_{\mathrm{PC,\,H-inf}}$ and $\boldsymbol{S}_{N,\,\mathrm{H-inf}}$ denote the errors introduced by the PC and the noise, respectively.

By taking the PC term into account, the channel estimation error is

$$\begin{aligned}
&\mathrm{MSE}_{\mathrm{PC,\,H-inf}} \\
&= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{\mathrm{PC,\,H-inf}}\boldsymbol{S}^H_{\mathrm{PC,\,H-inf}}\right]\right\} \\
&= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{\gamma}\boldsymbol{T}_j^{\dagger}\sum_{q\neq j}^{Q}\boldsymbol{T}_q\boldsymbol{C}_{jq}\left(\boldsymbol{\gamma}\boldsymbol{T}_j^{\dagger}\sum_{q\neq j}^{Q}\boldsymbol{T}_q\boldsymbol{C}_{jq}\right)^H\right]\right\} \\
&= \frac{1}{LK}\sum_{q\neq j}^{Q}\sum_{k=1}^{K}\mathrm{tr}\left\{\boldsymbol{\gamma}\boldsymbol{\Theta}\boldsymbol{F}_{N,\,L}\boldsymbol{R}_{CC}\boldsymbol{F}^H_{N,\,L}\boldsymbol{\Theta}^H\boldsymbol{\gamma}^H\right\} \qquad \text{(C.2)}
\end{aligned}$$

where $\boldsymbol{\Theta} = \boldsymbol{F}^H_{N,\,L}\Psi_{jqk}$, and $\Psi_{jqk} = \boldsymbol{X}^{\dagger}_{jk}\boldsymbol{X}_{qk}$. Notice that $\boldsymbol{\Theta}$ is a unitary matrix, whereas $\boldsymbol{\gamma}$ is a diagonal matrix; therefore, (C.2) can be simplified as

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{PC,\,H-inf}} &= \frac{1}{LK}\sum_{q\neq j}^{Q}\sum_{k=1}^{K}\mathrm{tr}\left\{\boldsymbol{\gamma}\boldsymbol{R}_{CC}\boldsymbol{\gamma}^H\right\} \\
&= \frac{1}{LK}\sum_{q\neq j}^{Q}r^2_{nn}d_{jq}\left(\sum_{k=1}^{K}\sum_{l=0}^{L-1}g^2_{jqkl}\right) \\
&= \frac{1}{L}\sum_{q\neq j}^{Q}r^2_{nn}d_{jq} \qquad \text{(C.3)}
\end{aligned}$$

where r_{nn} is the diagonal element of $\boldsymbol{\gamma}$.

For the noise term, we have $\boldsymbol{T}^H_{jk}\boldsymbol{T}_{jk} = \boldsymbol{I}_L$. Thus, the MSE is given as follows:

$$\begin{aligned}
\mathrm{MSE}_{N,\,\mathrm{H-inf}} &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{N,\,\mathrm{H-inf}}\boldsymbol{S}^H_{N,\,\mathrm{H-inf}}\right]\right\} \\
&= \frac{1}{LK}\sum_{k=1}^{K}\mathrm{tr}\left\{\boldsymbol{\gamma}\boldsymbol{T}^{\dagger}_{jk}\boldsymbol{R}_{ZZ}\left(\boldsymbol{\gamma}\boldsymbol{T}^{\dagger}_{jk}\right)^H\right\} \\
&\quad + \frac{1}{LK}\sum_{k=1}^{K}\mathrm{tr}\left\{\left(\boldsymbol{I}_L-\boldsymbol{\gamma}\right)\boldsymbol{R}_{CC}\left(\boldsymbol{I}_L-\boldsymbol{\gamma}\right)^H\right\} \\
&= \frac{1}{L}r^2_{nn}\sigma^2 + \frac{1}{L}d_{jj}(1-r_{nn})^2. \qquad \text{(C.4)}
\end{aligned}$$

Appendix D
Mean Square Error of Maximum Likelihood

For the ML algorithm, channel estimation between the jth BS and K users in the jth cell is performed by directly minimizing the following cost function [20]:

$$\hat{\boldsymbol{C}}^{\mathrm{ML}}_{jj} = \arg\min_{\boldsymbol{C}_{jj}}\left\{\|\boldsymbol{Y}_j - \boldsymbol{\Omega}_j\boldsymbol{C}_j\|^2\right\} \qquad \text{(D.1)}$$

where $\boldsymbol{\Omega}_j = [\boldsymbol{\Omega}_{j1},\ldots,\boldsymbol{\Omega}_{jK}]$, and $\boldsymbol{\Omega}_{jk} = \boldsymbol{X}_{qk}\boldsymbol{F}_{N,\,L}$. The channel estimation is given as follows:

$$\begin{aligned}
\hat{\boldsymbol{C}}^{\mathrm{ML}}_{jj} &= \boldsymbol{\Omega}^{\dagger}_j\boldsymbol{Y}_j \\
&= \boldsymbol{C}_{jj} + \boldsymbol{\Omega}^{\dagger}_j\sum_{q\neq j}^{Q}\boldsymbol{\Omega}_q\boldsymbol{C}_{jq} + \boldsymbol{\Omega}^{\dagger}_j\boldsymbol{Z}_j. \qquad \text{(D.2)}
\end{aligned}$$

Given (D.2), the channel estimation error is given as follows:

$$\begin{aligned}
\boldsymbol{S}_{\mathrm{ML}} &= \hat{\boldsymbol{C}}^{\mathrm{ML}}_{jj} - \boldsymbol{C}_{jj} \\
&= \boldsymbol{\Omega}^{\dagger}_j\sum_{q\neq j}^{Q}\boldsymbol{\Omega}_q\boldsymbol{C}_{jq} + \boldsymbol{\Omega}^{\dagger}_j\boldsymbol{Z}_j \\
&= \boldsymbol{S}_{\mathrm{PC,\,ML}} + \boldsymbol{S}_{N,\,\mathrm{ML}} \qquad \text{(D.3)}
\end{aligned}$$

where $\boldsymbol{S}_{\mathrm{PC,\,ML}}$ and $\boldsymbol{S}_{N,\,\mathrm{ML}}$ denote the errors introduced by the PC and the noise, respectively.

For the PC term, the channel estimation error for the kth user in the jth BS is

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{PC,\,ML}} &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{\mathrm{PC,\,ML}}\boldsymbol{S}^H_{\mathrm{PC,\,ML}}\right]\right\} \\
&= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{\Omega}^{\dagger}_j\sum_{q\neq j}^{Q}\boldsymbol{\Omega}_q\boldsymbol{C}_{jq}\left(\boldsymbol{\Omega}^{\dagger}_j\sum_{q\neq j}^{Q}\boldsymbol{\Omega}_q\boldsymbol{C}_{jq}\right)^H\right]\right\} \\
&= \frac{1}{LK}\sum_{q\neq j}^{Q}\sum_{k=1}^{K}\mathrm{tr}\left\{\boldsymbol{\Theta}\boldsymbol{F}_{N,\,L}\boldsymbol{R}_{CC}\boldsymbol{F}^H_{N,\,L}\boldsymbol{\Theta}^H\right\}. \qquad \text{(D.4)}
\end{aligned}$$

Therefore, (D.4) can be simplified as

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{PC,\,ML}} &= \frac{1}{LK}\sum_{q\neq j}^{Q}\sum_{k=1}^{K}\mathrm{tr}\left\{\boldsymbol{F}_{N,\,L}\boldsymbol{R}_{CC}\boldsymbol{F}^H_{N,\,L}\right\} \\
&= \frac{1}{LK}\sum_{q\neq j}^{Q}d_{jq}\left(\sum_{k=1}^{K}\sum_{l=0}^{L-1}g^2_{jqkl}\right) \\
&= \frac{1}{L}\sum_{q\neq j}^{Q}d_{jq}. \qquad \text{(D.5)}
\end{aligned}$$

For the noise term, we have $\boldsymbol{\Omega}^H_{jk}\boldsymbol{\Omega}_{jk} = \boldsymbol{I}_{LK}$, and the MSE is given as follows:

$$\begin{aligned}
\mathrm{MSE}_{N,\,\mathrm{ML}} &= \frac{1}{LK}\mathrm{tr}\left\{E\left[\boldsymbol{S}_{N,\,\mathrm{ML}}\boldsymbol{S}^H_{N,\,\mathrm{ML}}\right]\right\} \\
&= \frac{1}{LK}\sum_{k=1}^{K}\mathrm{tr}\left\{E\left[\boldsymbol{\Omega}^{\dagger}_j\boldsymbol{Z}_j\left(\boldsymbol{\Omega}^{\dagger}_j\boldsymbol{Z}_j\right)^H\right]\right\} \\
&= \frac{\sigma^2}{LK}\mathrm{tr}\left\{\left(\boldsymbol{\Omega}^{\dagger}_{jk}\boldsymbol{\Omega}_{jk}\right)^{-1}\right\} \\
&= \sigma^2. \qquad \text{(D.6)}
\end{aligned}$$

References

[1] D. Gesbert, M. Shafi, D. Shiu, and P. J. Smith, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications-a key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[3] D. Gesbert, M. Kountouris, R. W. Heath, Jr., C. B. Chae, and T. Sälzer, "From single user to multiuser communications: Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Oct. 2007.

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

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

[6] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna array," in *Proc. ICASSP*, Kyoto, Japan, Mar. 2012, pp. 3249–3252.

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

[8] K. T. Truong and R. W. Heath, Jr., "Effects of channel aging in massive MIMO systems," *J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[9] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Proces. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[11] H. Ngo, T. L. Marzetta, and E. G. Larsson, "Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models," in *Proc. ICASSP*, Prague, Czech Republic, May 2011, pp. 3464–3467.

[12] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "On the optimality of single-carrier transmission in large-scale antenna systems," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 276–279, Aug. 2012.

[13] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.

[14] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems-Part I: Chunk allocation," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, Sep. 2009.

[15] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems-Part II: Joint chunk, power and bit allocation," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 499–509, Feb. 2012.

[16] H. Zhu, "Performance comparison between microcellular and distributed antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1151–1163, Jun. 2011.

[17] J. Wang, H. Zhu, and N. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 675–683, May 2012.

[18] H. Zhu, "Radio resource allocation for OFDMA systems in high speed environments," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 748–759, May 2012.

[19] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[20] Y. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106–115, Jan. 2003.

[21] J. Gao and H. Liu, "Low-complexity MAP channel estimation for mobile MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 774–780, Mar. 2008.

[22] P. Xu, J. K. Wang, F. Qi, and X. Song, "Space-alternating generalised expectationCmaximisation-based H-infinity channel estimator for multiple-input multiple-output-orthogonal frequency division multiplexing systems," *IET Commun.*, vol. 5, no. 14, pp. 2068–2074, Sep. 2011.

[23] P. Xu, J. K. Wang, and F. Qi, "EM-based H-inf channel estimation in MIMO-OFDM systems," in *Proc. ICASSP*, Kyoto, Japan, Mar. 2012, pp. 3189–3192.

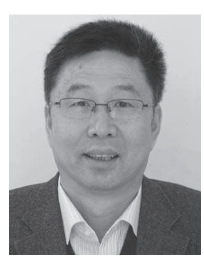
Peng Xu received the M.S. and Ph.D. degrees in communication and information systems from Northeastern University, Shenyang, China, in 2008 and 2011, respectively.

Since 2011, he has been with the Department of Information Science and Engineering, Northeastern University. He is also currently working as a Visiting Researcher with the Broadband and Wireless Communications Research Group, School of Engineering and Digital Arts, University of Kent, Kent, U.K. His research interests include wireless communications and signal processing, particularly channel identification and channel precoding.

Jiangzhou Wang (M'91–SM'94) received the B.S. and M.S. degrees from Xidian University, Xian, China, in 1983 and 1985, respectively, and the Ph.D. degree (with Greatest Distinction) from the University of Ghent, Belgium, in 1990, all in communication engineering.

He is currently the Chair of telecommunications and the Head of the Broadband and Wireless Communications Research Group with the School of Engineering and Digital Arts, University of Kent, Kent, U.K. He has published over 200 papers in international journals and conferences in the areas of wireless mobile communications and has written/edited three books. His research interests include fifth-generation mobile and small-cell technologies, device-to-device communications in cellular networks, massive multiple-input multiple-output and beamforming technologies, distributed antenna systems, and cooperative communications.

Mr. Wang is a Fellow of the Institution of Engineering and Technology and an IEEE Distinguished Lecturer. He serves/served as an Editor or Guest Editor for a number of international journals, such as the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He served/will serve as the Technical Program Chair for the 2013 IEEE Wireless Communications and Networking Conference in Shanghai, China, and as the Executive Chair of the 2015 IEEE International Conference on Communications in London, U.K. He received the Best Paper Award at the 2012 IEEE Global Communications Conference.



Jinkuan Wang (M'11) received the M.S. degree from Northeastern University, Shenyang, China, in 1985 and the Ph.D. degree from The University of Electro-Communications, Tokyo, Japan, in 1993.

In 1990, he joined the Institute of Space Astronautical Science, Japan, as a Special Member. In 1994, he was an Engineer with the Research Department, COSEL, Japan. Since 1998, he has been a Professor with the Department of Information Science and Engineering, Northeastern University. His main research interests include adaptive signal processing, mobile communications, and intelligent control.

Feng Qi received the Ph.D. degree from Katholieke Universiteit Leuven, Leuven, Belgium, in 2011.

He worked on active millimeter-wave imaging for five years, from system modeling to design, including both focal plane imaging and synthetic-aperture-radar imaging. He studied intensively on imaging noise issues and denoising methods during image acquisition. After graduation, he transferred to nonlinear optics and worked on terahertz (THz) generation and detection by using laser techniques. Recently, he began work on THz waveguide and THz frequency upconversion imaging. His research interests include antennas, planar circuits, computational electromagnetics, and scientific imaging and sensing systems.