High-Dimensional CSI Acquisition in Massive MIMO: Sparsity-Inspired Approaches

Juei-Chin Shen, *Member, IEEE*, Jun Zhang, *Senior Member, IEEE*, Kwang-Cheng Chen, *Fellow, IEEE*, and Khaled B. Letaief, *Fellow, IEEE*

Abstract—**Massive multiple-input–multiple-output (MIMO) has been regarded as one of the key technologies for fifth-generation wireless networks, as it can significantly improve both the spectral efficiency and the energy efficiency. The availability of high-dimensional channel side information (CSI) is critical for its promised performance gains, but the overhead of acquiring CSI may potentially deplete the available radio resources. Fortunately, it has recently been discovered that harnessing various sparsity structures in massive MIMO channels can lead to significant overhead reduction, and thus improve the system performance. This paper presents and discusses the use of sparsity-inspired CSI acquisition techniques for massive MIMO, as well as the underlying mathematical theory. Sparsity-inspired approaches for both frequency-division duplexing and time-division duplexing massive MIMO systems will be examined and compared from an overall system perspective, including the design tradeoffs between the two duplexing modes, computational complexity of acquisition algorithms, and applicability of sparsity structures. Meanwhile, some future prospects for research on high-dimensional CSI acquisition to meet practical demands will be identified.**

Index Terms—**Channel estimation, compressed sensing, massive multiple-input–multiple-output (MIMO), pilot contamination, pilot sequences, sparsity, ℓ_1 minimization.**

I. INTRODUCTION

MASSIVE multiple-input–multiple-output (MIMO) systems promise to boost spectral efficiency by more than one order of magnitude [1], [2]. Full benefits of massive MIMO, however, will never come to fruition without the base stations (BSs) having adequate channel knowledge, which appears to be an extremely challenging task [3]. The challenges posed by MIMO channels of very high dimension are confronted in both frequency-division duplexing (FDD) and time-division duplexing (TDD) massive MIMO systems. In the FDD mode, both the pilot-aided training overhead and the feedback overhead for channel side information (CSI) acquisition grow proportionally with the BS antenna size. However, the proportion of radio resources allocated to CSI acquisition is severely restricted by

the channel coherence period. The situation is made worse in an environment with high user equipment (UE) mobility.

In view of this, a considerable research effort has been devoted to TDD massive MIMO by exploiting channel reciprocity. Although the training overhead for TDD operation becomes proportional to the number of active UEs rather than that of BS antennas, the inevitable reuse of the same pilot in neighboring cells can seriously degrade the quality of obtained channel knowledge. This is because the channels to UEs in adjacent cells who share the same pilot will be collectively acquired by the BS. In other words, the desired channel obtained by the BS will be contaminated by interference channels. Once this contaminated channel knowledge is utilized for transmitting or receiving data, intercell interference occurs immediately and hence limits the achievable performance. This problem, known as *pilot contamination*, cannot be circumvented simply by adding more BS antennas.

Several attempts have been made to tackle the challenges of acquiring high-dimensional CSI in massive MIMO. For instance, in [4], open-/closed-loop training that utilizes temporal and spatial channel statistics is proposed to reduce the amount of downlink training overhead. For mitigating pilot contamination, the optimal design of precoding matrices aimed at minimizing the square errors caused by pilot reuse has shown its superiority over linear precoding [5]. Due to the recent advances in compressed sensing [6], [7], sparse signal processing has attracted much attention in such high-dimensional settings, which has also demonstrated its power in CSI acquisition in terms of reconstructing CSI from a limited number of channel measurements. Various sparsity structures exhibited by massive MIMO channels have recently been identified, thereby motivating the development of new strategies for CSI acquisition. Surprisingly, not only can high training overhead be reduced, but pilot contamination can be also resolved by appealing to sparsity-inspired approaches.

In this paper, we provide a comprehensive overview of the state-of-the-art research on sparsity-inspired approaches for high-dimensional CSI acquisition. In Section II, the challenges in FDD and TDD massive MIMO are reviewed in detail, including a rarely mentioned issue of FDD pilot contamination. On the basis of different sparsity structures, a variety of methods for either achieving overhead reduction or alleviating the effects of pilot reuse are examined and compared in Section III. Finally, concluding remarks are made in Section IV.

Notations: $\mathbb{C}$ denotes complex number, $\Re$ denotes real part, $\|\cdot\|_p$ denotes p-norm, $(\cdot)'$ denotes transpose, $(\cdot)^H$ denotes Hermitian transpose, $\mathbf{I}_N$ denotes $N \times N$ identity matrix,

J.-C. Shen, J. Zhang, and K. B. Letaief are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: eejcshen@ust.hk; eejzhang@ust.hk; eekhaled@ust.hk).

K.-C. Chen was with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan. He is now with the Department of Electrical Engineering, University of South Florida, Tampa, FL 33620 USA (e-mail: kwangcheng@usf.edu).
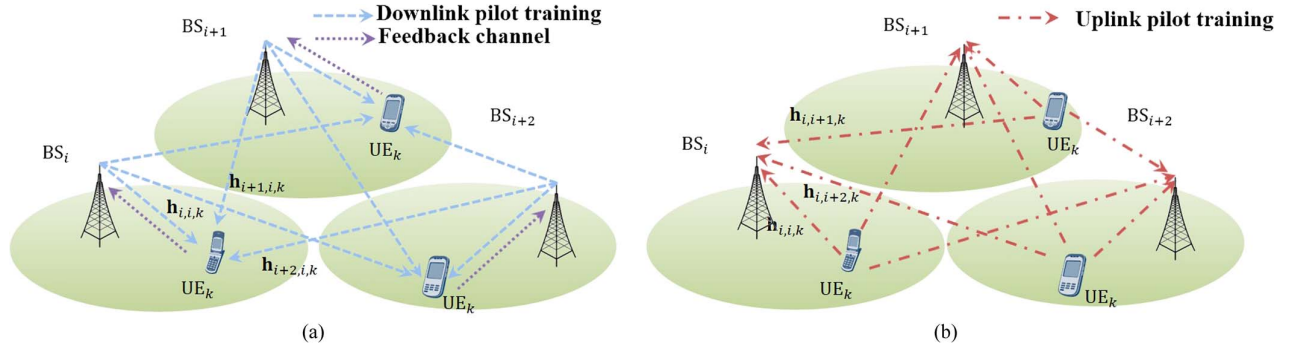
Fig. 1. Pilot reuse in multiple cells. (a) FDD downlink training. (b) TDD uplink training.

$\mathcal{N}(\cdot, \cdot)$ denotes normal distribution, $\mathbb{E}[\cdot]$ denotes expectation, $\mathbf{0}$ denotes zero vector, card$(\cdot)$ denotes cardinality, supp$(\cdot)$ denotes the set of indices of nonzero elements, Var$(\cdot)$ denotes variance, max$\{\cdot\}$ denotes the maximum element, Vec$(\cdot)$ denotes vectorization, $\otimes$ denotes Kronecker product, $\succeq$ denotes matrix inequality, and $(\cdot)^{\dagger}$ denotes pseudoinverse.

II. CHALLENGES OF HIGH-DIMENSIONAL CSI ACQUISITION

In massive MIMO systems with high-dimensional channels, CSI acquisition at BSs is a fundamentally challenging problem. In FDD massive MIMO, performing this task consumes a considerable amount of radio resources, which is proportional to the dimension of channels. On the other hand, in TDD-mode operation, it is hard to ensure the orthogonality of pilot sequences in the multicell scenario as the number of overall UEs becomes large. As a result, the inevitable reuse of correlated pilot sequences in different cells, known as pilot contamination, causes capacity-limiting intercell interference.

To illustrate these difficulties further, we will consider a massive MIMO network consisting of L hexagonal cells. In each cell, there is a BS equipped with an M-element linear array,[1] serving K single-antenna UEs. The channel between BS i and UE k in cell j is denoted by the $M \times 1$ vector $\mathbf{h}_{i,j,k}$. The BS antenna size is supposed to be greatly larger than the number of served UEs.

A. FDD Massive MIMO

In the FDD mode, obtaining CSI at BSs is normally performed in two steps. First, each BS sends a downlink training matrix to its served UEs. Second, each UE estimates the desired channel based on the downlink measurements and feeds back acquired CSI through dedicated uplink feedback channels.

During downlink training, UE k in cell i receives channel measurements

$$\mathbf{y}_{i,k}^{\mathrm{DL}} = \mathbf{S}_i^{\mathrm{DL}} \mathbf{h}_{i,i,k} + \sum_{l \neq i} \mathbf{S}_l^{\mathrm{DL}} \mathbf{h}_{l,i,k} + \mathbf{z}_{i,k}^{\mathrm{DL}} \qquad (1)$$

---

[1]For simplicity, the assumption of employing linear arrays is made. However, most of the results discussed in this paper can be generalized to include the cases of using planar or cylindrical arrays.

where $\mathbf{S}_l^{\mathrm{DL}}$ denotes the $N \times M$ pilot training matrix used in cell l, $\mathbf{z}_{i,k}^{\mathrm{DL}}$ is the additive noise, whereas the first term of the right-hand side (RHS) represents the desired channel measurements, and the next term results from intercell interference. Even without considering the impact of intercell interference, the required training overhead N for conventional least squares (LS) or minimum mean square error (MMSE) estimators to achieve a reasonable performance level still scales linearly with the BS antenna size. By taking intercell interference into account, a further increase in training overhead would occur. The explicit expressions of the optimal pilot training matrices ($N \geq M$) are provided in [8] for single-cell networks. In [9], the optimal design of training matrices for multicell MIMO-OFDM systems is considered.

What makes the situation worse is that typical feedback channels are finite rate. This implies that only quantized versions of channel estimates can be fed back to BSs. If there are predefined codebooks consisting of precoding vectors, then the index of the optimal codebook vector is required to be sent back [10], [11]. However, either the amount of quantized CSI or the size of codebooks increases in proportion to the number of BS antennas, and it, in turn, makes these two limited feedback techniques impractical in FDD massive MIMO.

Note that, when the same training matrix is repeatedly used in multiple cells, i.e., $\mathbf{S}_1^{\mathrm{DL}} = \cdots = \mathbf{S}_L^{\mathrm{DL}}$, this can be regarded as pilot contamination in FDD massive MIMO. As a result of such contamination, as shown in Fig. 1(a), BS i will acquire the composite channel $\sum_{l=1}^{L} \mathbf{h}_{l,i,k}$ rather than the desired channel $\mathbf{h}_{i,i,k}$, given the feedback channel being error free and the additive noise being ignored. Despite this fact, utilizing this composite CSI to form a precoding vector and transmit signals at BS i will not cause serious interference to UEs in the neighboring cells. For instance, given that maximum ratio transmission (MRT) precoding is employed, the transmitted signal from BS i can be expressed as $\mathbf{x}_i = \sum_{k=1}^{K} \mathbf{w}_{i,k}^{\mathrm{FDD}} x_{i,k}$, where $x_{i,k}$ is the signal intended for UE k within the cell, and $\mathbf{w}_{i,k}^{\mathrm{FDD}} = \sum_{l=1}^{L} (\mathbf{h}_{l,i,k}^H)'$ denotes the MRT precoding vector. During the downlink transmission phase, the received interference at UE m in cell j due to BS i is given by

$$I_{i,j,m} = \mathbf{h}_{i,j,m}' \mathbf{x}_i = \sum_{k=1}^{K} \sum_{l=1}^{L} \mathbf{h}_{l,i,k}^H \mathbf{h}_{i,j,m} x_{i,k}. \qquad (2)$$

When the number of BS antennas grows without limit, the channel vectors are asymptotically orthogonal. Thus, the channel products $\mathbf{h}_{l,i,k}^{H}\mathbf{h}_{i,j,m}$ approach zero and so does the interference $I_{i,j,m}$. In other words, intercell interference caused by pilot contamination diminishes asymptotically with increasing BS antenna size. This implies that there is no need to mitigate intercell interference by making training matrices distinct from each other in the asymptotic regime. Hence, the existing literature rarely addresses the issue of pilot contamination in FDD massive MIMO.

Note that uplink training in the FDD mode is not considered here. An explanation for this is provided as follows. The uplink CSI is mainly utilized for data acquisition in a multiple-access channel, instead of a broadcast channel. This means that more advanced signal processing techniques, such as blind multiuser detection, can be applied at the BS side. Thus, pilot-aided training may not be the best choice, and CSI acquisition is not necessarily separated from data acquisition.

B. TDD Massive MIMO

Making massive MIMO operate in the TDD mode is a promising way to circumvent the identified difficulties in the FDD mode. Owing to channel reciprocity in the TDD mode, the CSI obtained via uplink training can be utilized for downlink transmission. More important, the cost of uplink training now increases linearly with the number of active UEs rather than that of BS antennas. Typically, for obtaining accurate CSI, it requires that each UE transmits an orthogonal pilot sequence to its serving BS. However, the number of available orthogonal pilot sequences is limited by the ratio of the channel coherence interval to the channel delay spread [12], which may be small due to the mobility of UEs or adverse physical environments. When the number of overall UEs becomes large, the situation of using nonorthogonal pilot sequences, known as pilot contamination, inevitably arises. A consequence of pilot contamination is intra- and intercell interference.

During the uplink training phase, the received signal at the ith BS is given by

$$\mathbf{Y}_{i}^{\mathrm{UL}} = \sum_{l=1}^{L} \mathbf{S}_{l}^{\mathrm{UL}}\mathbf{H}_{i,l} + \mathbf{Z}_{i}^{\mathrm{UL}} \qquad (3)$$

where $\mathbf{H}_{i,l} = [\mathbf{h}_{i,l,1}, \ldots, \mathbf{h}_{i,l,K}]'$ consists of channel vectors from UEs in the lth cell to the ith BS, the columns of $\mathbf{S}_{l}^{\mathrm{UL}}$ form a set of $\tau \times 1$ pilot sequences $\{\mathbf{s}_{l,k}\}_{k=1}^{K}$, and $\mathbf{Z}_{i}^{\mathrm{UL}}$ denotes an additive noise matrix. To illustrate the case of intercell interference, assume that the same set of orthogonal pilot sequences is reused in each cell, i.e., $\mathbf{S}_{1}^{\mathrm{UL}} = \cdots = \mathbf{S}_{L}^{\mathrm{UL}}$ and $\mathbf{s}_{l,k_1}'\mathbf{s}_{l,k_2} = 0$ for $k_1 \neq k_2$, as shown in Fig. 1(b). Employing the LS estimator yields the channel estimate

$$\widehat{\mathbf{H}}_{i,i} = \left[\left(\mathbf{S}_{i}^{\mathrm{UL}}\right)^{H}\mathbf{S}_{i}^{\mathrm{UL}} \right] \left(\mathbf{S}_{i}^{\mathrm{UL}}\right)^{H}\mathbf{Y}_{i}^{\mathrm{UL}}$$

$$= \mathbf{H}_{i,i} + \sum_{l \neq i} \mathbf{H}_{i,l} + \left[\left(\mathbf{S}_{i}^{\mathrm{UL}}\right)^{H}\mathbf{S}_{i}^{\mathrm{UL}} \right] \left(\mathbf{S}_{i}^{\mathrm{UL}}\right)^{H}\mathbf{Z}_{i}^{\mathrm{UL}} \quad (4)$$
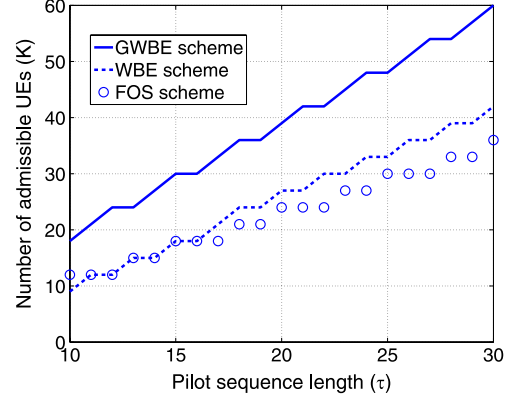

Fig. 2. Number ($K = 3l$) of admissible UEs versus pilot sequence length for the GWBE, WBE, and FOS schemes, given a fixed SINR requirement pattern, that is, $\{\gamma_{1 \sim l} = (1/3), \gamma_{(l+1) \sim 2l} = 1, \gamma_{(2l+1) \sim 3l} = 3\}$ (from [13]).

where the rows of $\widehat{\mathbf{H}}_{i,i}$ are given by $\hat{\mathbf{h}}_{i,i,k} = \sum_{l=1}^{L} \mathbf{h}_{i,l,k}$ when ignoring the noise. During downlink transmission, using estimates $\hat{\mathbf{h}}_{i,i,k}$ to form the transmit signal $\mathbf{x}_{i} = \sum_{k=1}^{K} \mathbf{w}_{i,k}^{\mathrm{TDD}} x_{i,k}$, where $\mathbf{w}_{i,k}^{\mathrm{TDD}} = \sum_{l=1}^{L} \left(\mathbf{h}_{i,l,k}^{H}\right)'$ are MRT precoding vectors, will cause interference

$$I_{i,j,m} = \mathbf{h}_{i,j,m}'\mathbf{x}_{i}$$
$$= \|\mathbf{h}_{i,j,m}\|_{2}^{2}x_{i,m} + \sum_{\substack{k \neq m \\ \text{or } l \neq j}} \mathbf{h}_{i,l,k}^{H}\mathbf{h}_{i,j,m}x_{i,k} \quad (5)$$

to UE m in cell j. Although the second term on the RHS of (5) decreases with the increasing BS antenna size, the first term, which does not vanish, makes the received signal-to-interference-plus-noise ratio (SINR) at UE m in cell j converge to a limit and becomes the performance limiting factor.

The current investigation into TDD pilot contamination focuses on its impact on the received SINR or the sum rate when linear precoders/detectors are applied. However, very little is known about its impact on the system equipped with nonlinear precoders/detectors. A recent work [13], [14] has provided an interesting perspective on the user capacity of pilot-contaminated massive MIMO, which quantifies the maximum number of admissible UEs given their own SINR requirements. As shown in Fig. 2, the user capacity of three schemes[2] of joint pilot design and transmit power allocation is fundamentally limited by the length of pilot sequences. For further details about pilot contamination in TDD massive MIMO, the studies [15] and [16] and the references therein should be consulted.

III. SPARSITY-INSPIRED CSI ACQUISITION

Despite the challenges imposed by the high dimensionality of channel matrices, a number of research efforts have sought to address them and have achieved reasonably efficient CSI acquisition. In particular, sparsity-inspired approaches have been proved to be powerful tools, as presented in the following.

---

[2]The pilot sequences employed in the GWBE, WBE, and FOS schemes are generalized Welch bound equality (GWBE) sequences, WBE sequences, and finite orthogonal sequences (FOS), whose correlation among sequences is either 1 or 0, respectively. The same downlink power allocation, i.e., $P_{i} \propto \gamma_{i}/(1 + \gamma_{i})$, is used in the three schemes.

A. FDD Massive MIMO

1) Joint CSI Recovery Method: Authors of [17] proposed a method for low-overhead pilot training in the single-cell scenario, taking advantage of channel sparsity. Provided that a uniform linear array with critically spaced antennas is employed at the BS, the channel $\mathbf{h}_k$, where indices of BSs are discarded in the single-cell scenario, exhibits a sparse representation $\mathbf{h}_k^{\mathrm{a}}$ in the angular domain, i.e.,

$$\mathbf{h}_k = \mathbf{U}\mathbf{h}_k^{\mathrm{a}} \tag{6}$$

where $\mathbf{U}$ is a discrete Fourier transform (DFT) matrix whose columns form an angular basis. The cardinality of $\mathrm{supp}(\mathbf{h}_k^{\mathrm{a}})$ can be reasonably assumed to be greatly less than M because of limited local scattering at the BS whose antenna array mounted higher than surrounding scatterers. Additionally, based on the results in [18], it has been argued that the channels to UEs are likely to share a partially common support in the angular domain, i.e., $\cap_{k=1}^K \mathrm{supp}(\mathbf{h}_k^{\mathrm{a}}) = \Omega_c$. In order to utilize the channel sparsity and common support property simultaneously, channel measurements acquired at UEs are fed back to the serving BS via error-free feedback channels. Hence, a joint channel recovery problem can be formulated as follows:

$$\min_{\{\mathbf{h}_k, \forall k\}} \sum_{k=1}^K \left\| \mathbf{y}_k^{\mathrm{DL}} - \mathbf{S}^{\mathrm{DL}}\mathbf{h}_k \right\|_1^2$$
$$\text{s.t. } \cap_{k=1}^K \mathrm{supp}\left(\mathbf{h}_k^{\mathrm{a}}\right) = \Omega_c. \tag{7}$$

Using orthogonal matching pursuit (OMP) as a basis, a greedy algorithm has been proposed to efficiently solve this problem. The simulation results show that the required training overhead for this recovery algorithm can be significantly less than that for the conventional LS estimator. Moreover, the mean square error performance improves with the increasing cardinality of Ω_c.

One major concern about this joint recovery approach is the underlying assumption of perfect channel measurements being fed back. As practical feedback channels are rate limited, it is more reasonable to assume quantized measurements at the BS. The impact of quantization on the channel recovery performance requires further investigation. On the other hand, it has been suggested that the amount of channel measurements that is needed at the BS should be adaptively adjusted according to the sensitivity of the system performance to the CSI inaccuracy [19]. Furthermore, there has been little quantitative analysis of the required training overhead against the channel sparsity level. This quantification is greatly needed as it will help us measure the actual training overhead reduction that can be achieved without relying on time-consuming simulations.

2) Weighted ℓ_1 Minimization Method: Considering a similar single-cell scenario, the study in [20] has drawn attention to utilizing partial support information of sparse massive MIMO channels, which is a collection of indices of significant entries of channel vectors in the angular domain. The main advantage of using partial support information is the possibility of achieving a remarkable training overhead reduction. Specifically, the order of the required overhead decreases from $\mathcal{O}(s \log M)$ to $\mathcal{O}(s)$, where $s = \mathrm{card}[\mathrm{supp}(\mathbf{h}_k^{\mathrm{a}})]$ is the channel sparsity level.

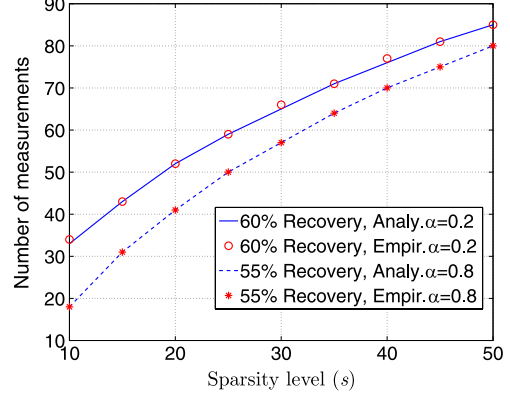

Fig. 3. Phase transition curves of (8) over different values of α given $M = 100$, $\hat{s} = 10$, $\mathbf{z}_k^{\mathrm{DL}} = \mathbf{0}$, and $\epsilon = 0$ (from [20]).

Assume that the partial support information $\widehat{T}_k$ of channel $\mathbf{h}_k^{\mathrm{a}}$ is available at UE k, where $\mathrm{card}(\widehat{T}_k) = \hat{s}$, and $\mathrm{card}[\mathrm{supp}(\mathbf{h}_k^{\mathrm{a}}) \cap \widehat{T}_k]$ is given by $\lfloor \alpha\hat{s} \rfloor$. The higher the factor α, the higher is the accuracy level of partial support information. Based on a weighted ℓ_1 minimization framework, the channel recovery is performed as follows:

$$\min_{\hat{\mathbf{h}}_k^{\mathrm{a}} \in \mathbb{C}^M} \left\| \hat{\mathbf{h}}_k^{\mathrm{a}} \right\|_{1,\mathbf{w}}$$
$$\text{subject to } \left\| \mathbf{S}^{\mathrm{DL}}\mathbf{U}\hat{\mathbf{h}}_k^{\mathrm{a}} - \mathbf{y}_k^{\mathrm{DL}} \right\|_2 \leq \epsilon$$
$$\text{with} \quad w_i = \begin{cases} 1, & i \notin \widehat{T}_k \\ 0, & i \in \widehat{T}_k \end{cases} \tag{8}$$

where $\mathbf{S}^{\mathrm{DL}} \in \mathbb{C}^{N \times M}$ is designed to be a Gaussian random matrix of independent complex normal entries, the noise $\mathbf{z}_k^{\mathrm{DL}}$ is assumed to be upper bounded, i.e., $\|\mathbf{z}_k^{\mathrm{DL}}\|_2 \leq \epsilon$, and $\|\hat{\mathbf{h}}_k^{\mathrm{a}}\|_{1,\mathbf{w}} = \sum_{i=1}^M w_i|\hat{h}_k^{\mathrm{a}}[i]|$. In the objective function, the entries that are expected to be zero are weighted more heavily than others. The results show a significant improvement over the method without using partial support information when the accuracy level α exceeds a certain threshold. Moreover, taking a convex geometry approach, the authors have successfully and precisely quantified the required training overhead for achieving a certain percentage of exact recovery. The exact recovery is declared if $\|\hat{\mathbf{h}}_k^{\mathrm{a}} - \mathbf{h}_k^{\mathrm{a}}\|_2 \leq 10^{-4}$. As shown in Fig. 3, the analytical curves of $\alpha = 0.2$ and $\alpha = 0.8$ can accurately depict the empirical phase transition curves of 60% exact recovery and 55% exact recovery, respectively.

Unlike the previous method, here, channel measurements are not fed back to the BS. In other words, it avoids the assumption of error-free feedback channels. However, it raises another issue of storing random matrices at UEs with limited memory. In addition, performing convex optimization can impose a stringent computation requirement on UEs without seeking for low-complexity solutions. Several attempts have been made to design practical training matrices. In [21], Toeplitz-structured training matrices, suggested for the realistic implementation, are shown to perform comparably with Gaussian random matrices and require generating less independent random variables.

A deterministic approach to the training matrix design is first considered by appealing to matrix properties such as mutual coherence [22]. More advanced deterministic training matrices are developed in [23] to yield higher recovery accuracy. In the context of FDD massive MIMO, it would be interesting to invent structurally random or deterministic training matrices that take partial support information of channels to multiple UEs into consideration. In addition, the similar concepts of using prior channel knowledge to lower training overhead can be found in [4], where spatial and temporal correlations are harnessed. More study is needed to better understand how to integrate all the relevant prior knowledge into efficient CSI acquisition.

B. TDD Massive MIMO

As mentioned in Section II-B, employing uplink training to obtain high-dimensional downlink CSI results in undesired pilot contamination, and the following are some efforts to address this issue.

1) Coordinated MMSE Method: Contradicting conventional wisdom, it has been shown that it is possible to mitigate pilot contamination using the linear MMSE estimator [24]. The key factor in determining the success of MMSE estimation is that each channel to the UE can be regarded as a linear combination of finite steering vectors

$$\mathbf{h}_{i,j,k} = \frac{1}{\sqrt{P}} \sum_{p=1}^{P} \alpha_{i,j,k}(p) \mathbf{a} \left[\theta_{i,j,k}(p) \right] \qquad (9)$$

where P is the number of paths, $\alpha_{i,j,k}(p)$ are zero-mean path gains, and $\mathbf{a}[\theta_{i,j,k}(p)]$ denote the steering vectors due to angle of arrivals (AoAs) $\theta_{i,j,k}(p)$. Consequently, the rank of the channel covariance matrix $\mathbf{R}_{i,j,k} \triangleq \mathbb{E}\{\mathbf{h}_{i,j,k}\mathbf{h}_{i,j,k}^H\}$ depends on the range $[\theta_{i,j,k}^{\min}, \theta_{i,j,k}^{\max}]$ in which AoAs $\theta_{i,j,k}(p)$ lie, which typically turns out to be low. Let us focus on the kth row of (4), i.e., $\hat{\mathbf{h}}_{i,i,k} = \sum_{l=1}^{L} \mathbf{h}_{i,l,k} + \mathbf{z}_k$. Based on it, the desired channel $\mathbf{h}_{i,i,k}$ can be further extracted by the MMSE estimator, i.e.,

$$\hat{\hat{\mathbf{h}}}_{i,i,k} = \mathbf{R}_{i,i,k} \left(\sigma_z^2 \mathbf{I}_M + \sum_{l=1}^{L} \mathbf{R}_{i,l,k} \right)^{-1} \hat{\mathbf{h}}_{i,i,k} \qquad (10)$$

where the covariance matrix of $\mathbf{z}_{i,k}$ is assumed to be $\sigma_z^2 \mathbf{I}_M$. When the range of AoAs due to interfering UEs that use the same pilot sequence does not overlap with the AoA range due to the desired UE, the estimate $\hat{\hat{\mathbf{h}}}_{i,i,k}$ approaches the desired $\mathbf{h}_{i,i,k}$ as the BS antenna size grows to infinity. This feature is highly attractive because the dimension of the BS antennas can be made as large as desired in massive MIMO. Moreover, the condition of nonoverlapping AoA ranges can be satisfied if the reused pilot sequence is properly allocated to UEs in neighboring cells. A heuristic algorithm has been developed to perform pilot allocation in a coordinated manner. Another favorable feature of this method recently demonstrated in [25] is that the asymptotically optimal estimate is obtainable whether uniform or nonuniform arrays are employed. As a

result, BS antenna arrays are exempt from the requirement of high calibration accuracy.

The second-order statistics of high-dimensional channels have been successfully utilized to facilitate robust MMSE channel estimation under pilot contamination. However, obtaining channel covariance matrices of high dimension imposes another challenge to the massive MIMO system. It is interesting to know if the low rankness can help speed up the acquisition of channel covariance matrices. Furthermore, it is still unknown if this covariance-matrix-aware method is sensitive to the inaccuracy of the second-order statistics. On the other hand, the information about AoAs actually can be extracted from statistical channel knowledge prior to commencing the instantaneous CSI acquisition [26]. In this case, the dimension of the parameter space of each channel shrinks to P, which can be significantly less than the original. Most important, this information could aid BSs in distinguishing between training signals from UEs using the same pilot.

2) Quadratic SDP Method: It is suggested that a BS should collect CSI of both the desired links within the cell and interference links from its neighboring cells [27]. In other words, the CSI of interference links should not be regarded as irrelevant information. From this new angle, the expression (3) can be recast as

$$\mathbf{Y}_i^{\mathrm{UL}} = \underline{\mathbf{S}}^{\mathrm{UL}} \mathbf{H}_i + \mathbf{Z}_i^{\mathrm{UL}} \qquad (11)$$

where $\underline{\mathbf{S}}^{\mathrm{UL}} \triangleq [\mathbf{S}_1^{\mathrm{UL}}, \dots, \mathbf{S}_L^{\mathrm{UL}}]$ and $\mathbf{H}_i \triangleq [\mathbf{H}_{i,1}', \dots, \mathbf{H}_{i,L}']'$ are the full CSI of wireless links that should be recovered. Thus, the currently challenging issue is similar to that in FDD massive MIMO, i.e., how to reduce the required training overhead.

In the undesirable scattering propagation environments, the rank of the channel matrix is equal to the number r of the feasible AoAs $\theta_{i,j,k}(p)$ in (9), which is greatly less than $\max\{M, K \cdot L\}$. Based on this observation, an unclear norm regularized problem can be formulated as

$$\min_{\mathbf{H}_i} \quad \frac{1}{2} \left\| \mathrm{vec} \left(\mathbf{Y}_i^{\mathrm{UL}} \right) - \Psi \mathrm{vec} \left(\mathbf{H}_i \right) \right\|_2^2 + \gamma \| \mathbf{H}_i \|_F \qquad (12)$$

where $\Psi = \underline{\mathbf{S}}^{\mathrm{UL}} \otimes \mathbf{I}_M$, and γ is a regularization factor. The sole purpose of adopting unclear norm regulation is to minimize the sum of the matrix's singular values, thereby achieving rank minimization. The preceding problem has been further recast as a quadratic semidefinite programming (SDP) problem, i.e.,

$$\begin{aligned} \min_{\mathbf{v}} \quad & \frac{1}{2}\mathbf{v}^H\mathbf{v} - \Re\left\{ \left[\mathrm{vec}\left(\mathbf{Y}_i^{\mathrm{UL}}\right) \right]^H \mathbf{v} \right\} \\ \mathrm{s.t.} \quad & \begin{bmatrix} \gamma\mathbf{I}_{KL} & \mathrm{vec}_{KL,M}^{-1}(\Psi^H\mathbf{v}) \\ \left[\mathrm{vec}_{KL,M}^{-1}(\Psi^H\mathbf{v})\right]^H & \gamma\mathbf{I}_M \end{bmatrix} \succeq 0. \end{aligned} \qquad (13)$$

The solution $\mathbf{v}^*$ to this SDP problem determines the estimate of the channel matrix

$$\mathbf{H}_i^* = \mathrm{vec}_{KL,M}^{-1} \left\{ \Psi^\dagger \left[\mathrm{vec}\left(\mathbf{Y}_i^{\mathrm{UL}}\right) - \mathbf{v}^* \right] \right\} \qquad (14)$$

which can now be obtained efficiently, due to the readily available polynomial-time SDP solvers.

TABLE I
COMPARISON OF SPARSITY-INSPIRED CSI ACQUISITION METHODS

Methods	Sparsity Types	Pros	Cons
Joint CSI Recovery (FDD)	Sparse channel vectors & Common supports	• Jointly exploit sparsity & common-support property • Perform channel recovery at the BS	• UEs need to feed back perfect channel measurements
Weighted ℓ_1 Minimization (FDD)	Sparse channel vectors & partial support information	• Sharp estimate of the required training overhead • Lower training overhead	• Need to obtain partial support information
Coordinated MMSE (TDD)	Low-rank channel covariance matrices	• Performance improves with increasing antenna size • Lower training overhead	• Need to obtain second-order channel statistics
Quadratic SDP (TDD)	Low-rank channel matrices	• No need for knowledge of second-order channel statistics	• Only suitable for poor scattering propagation environments • Higher training overhead
Sparse Bayesian Learning (TDD)	Sparse channel vectors in the UE domain	• No need for knowledge of second-order channel statistics	• Channels are not jointly recovered • Higher training overhead

In the commencing study of massive MIMO [28], the CSI of interference links at BSs is viewed as nonessential. This is because desired links and interference links are asymptotically orthogonal, and more important, intercell interference can be proved manageable with the CSI of desired links only. Here, we offer an explanation why there is a need for acquiring the CSI of interference links in the poor scattering environments. Consider that $\mathbf{H}_i = \mathbf{G}_i \mathbf{A}$, where $\mathbf{A} = [\mathbf{a}(\phi_1), \ldots, \mathbf{a}(\phi_r)]'$ is an $r \times M$ matrix of full row rank with $r \ll \min\{M, KL\}$ due to poor scattering, and $\mathbf{G}_i$ consists of $KL \times r$ independent and identically distributed zero-mean channel gains. Then, we have $\lim_{M \to \infty} \mathbf{A}\mathbf{A}^H = \mathbf{I}_r$ and

$$\lim_{M \to \infty} \mathbf{H}_i \mathbf{H}_i^H = \mathbf{G}_i \mathbf{G}_i^H \not\propto \mathbf{I}_{KL} \qquad (15)$$

which implies that the correlation among wireless links does not diminish with the increasing BS antenna size. In such a situation, it becomes crucial to obtain the full CSI of wireless links for effective interference management.

3) SBL Method: Sharing the same perspective as the study [27], the work in [29] also considers acquiring the full CSI of wireless links and proposes a sparse Bayesian learning (SBL) method to achieve this goal. SBL was first presented in [30] and has been proved to outperform some prevailing ℓ_1 minimization algorithms [31]. The SBL method proceeds by first transforming the channel matrix into the angular domain via DFT, as mentioned in the joint CSI recovery method, i.e., $\underline{\mathbf{H}}_i = \mathbf{H}_i \mathbf{U}$. Interestingly, instead of taking advantage of the sparsity in the angular domain, the sparsity in the UE domain, which has been empirically shown to exist, is utilized. In other words, the column vectors of the channel matrix $\underline{\mathbf{H}}_i$ are considered one by one. As each column vector consists of elements due to different UEs, the independence among elements can be reasonably assumed. This independence together with the sparsity

in the UE domain leads to an effective Gaussian-mixture (GM) model, which well describes the joint distributions of the channel elements. More surprising, empirical results show that there are only few parameters involved in the GM model that need to be determined. Therefore, the practical Bayes estimation can be implemented by evaluating marginal probability density functions via the approximate message passing (AMP) algorithm [32] and learning GM parameters by means of the expectation-maximization (EM) algorithm [33]. The numerical results show that this Bayesian method can achieve a significant reduction in estimation errors.

The assumption of channel vectors being sparse in the UE domain may not hold when the UE dimension KL is not large enough. A possible remedy for this situation is suggested in the following. First, it is desirable to understand if the GM model is also applicable for modeling distributions of spare channel vectors in the angular domain. Second, as angular-domain channels are very likely to consist of a small number of blockwise nonzero segments resulting from few clusters of scatterers, it is eminently reasonable to assume some dependence among angular-domain channel elements. Hence, the distribution of the channel vector could be a mixture of Gaussian random vectors, and the original AMP and EM algorithms should be modified accordingly to this new GM model.

C. Discussion and Comparison

In the previous subsections, several methods for efficient high-dimensional CSI acquisition have been discussed for massive MIMO communications. Table I provides a brief summary of the advantages and disadvantages of these methods. It is shown in the table that each method utilizes a distinct sparsity structure. However, all sparsity structures considered in massive MIMO are based on the observation that angular-domain

channels are sparse. As a result, the second-order statistics of massive MIMO channels inherit the sparsity structure, yielding low-rank channel covariance matrices. In addition, as sparse channels are collectively examined, it leads to either block-sparse or low-rank channel matrices. When the UE dimension is comparable with the channel dimension, sparsity in the angular domain also results in sparsity in the UE domain. On the basis of the aforementioned sparsity structures, different sparsity-inspired methods are developed either to reduce training overhead or to mitigate pilot contamination.

In FDD massive MIMO, without feeding back channel measurements to the BS side, less sparsity structures are available for developing efficient CSI acquisition methods. Despite this limitation, the weighted ℓ_1 minimization method shows that achieving further overhead reduction is feasible if partial support information can be obtained in advance and properly harnessed. Interestingly, by enabling the BS to gather perfect channel measurements from its served UEs, the joint CSI recovery method offers an effective way of utilizing sparsity structures across multiple UEs. If the performance superiority of this method still holds when taking rate-limited feedback channels into account, it will establish the fact that offloading CSI acquisition tasks to the BS is feasible and beneficial.

With regard to TDD massive MIMO, uplink training has more sparsity structures to utilize as high-dimensional channels are jointly recovered at the BS side. It is worth noting that only low-rank channel covariance matrices have been used for pilot decontamination. Other sparsity structures such as low-rank channel matrices and sparse UE-domain channels have not been considered for mitigating the effects of pilot reuse. In this regard, there is still much room for innovation in sparsity-inspired pilot decontamination. It is also worth noting that using perfect covariance matrices of both desired channels and interference channels in the coordinated MMSE method has drawn criticism [34]. It would be intriguing to assess if there exist efficient algorithms for learning low-rank covariance matrices. If such algorithms are developed or identified, they should be integrated into the coordinated MMSE method.

D. Implementation Issues

Recently, investigators have examined the practical implementation of compressed-sensing-based algorithms for sparse channel recovery [35]–[37]. Although the design targets are channel models in the 3GPP LTE standard, several insights that have been provided are still valuable and applicable to realistic implementation of sparse massive MIMO channel recovery. It has been pointed out that greedy algorithms such as OMP or matching pursuit (MP) are more desirable from a hardware perspective. It is because these algorithms require lower computational complexity and lower numerical precision when compared with convex relaxation algorithms such as basis pursuit [36]. The tradeoff between hardware complexity and denoising performance of three greedy algorithms has been characterized in [37], and it is indicated that the chip area overhead required to implement the gradient pursuit algorithm can be three times larger than MP. The power consumption is normally proportional to this area overhead. When it comes

to the design of channel recovery algorithms in FDD massive MIMO, which are typically performed at the UE side, the issue of hardware complexity should be carefully taken into account. On the other hand, at the BS side, high-dimensional channels can be recovered by more advanced algorithms such as SBL or joint CSI recovery. To improve the computational efficiency, recent advances in large-scale convex optimization could be helpful [38].

E. Implications of New Propagation Models

Most existing studies have based their CSI acquisition approaches on the conventional MIMO channel models, which may fail to capture some unique characteristics of massive MIMO channels. For instance, the far-field and plane wavefront assumptions no longer hold when antenna arrays become physically larger than the Rayleigh distance [39]. On the other hand, the sheer size of antenna arrays, where different antenna elements observe varying subsets of scatterer clusters, makes the assumption of spatial channels being wide-sense stationary on the array axis no longer valid [40]. While new channel models have been proposed in [41] and [42] by making a more accurate spherical wavefront assumption and taking the nonstationarities into consideration, there is still very little understanding of how these characteristics affect the sparsity structures of the channels in massive MIMO systems. One previous result [43], however, suggests that the spherical wavefront model does adequately characterize the rank of the channel matrix. This implies that the new channel models can potentially affect the SDP method, which exploits the sparsity in the form of the channel matrix rank. In addition, the possibility that none of clusters are perceptible to some antenna elements cannot be categorically excluded; thus, it indicates the possible presence of the sparsity on the array axis. These inferences suggest that there is abundant room for further progress in identifying utilizable sparsity structures based on the latest models.

IV. CONCLUSION

In this paper, the challenges of acquiring high-dimensional CSI in FDD/TDD massive MIMO systems have been discussed. To address these challenges and break the curse of dimensionality, one can effectively utilize sparsity structures that uniquely appear in massive MIMO channels. Several state-of-the-art sparsity-inspired approaches for high-dimensional CSI acquisition have been examined and compared in terms of the sparsity structures being exploited, while their own advantages and disadvantages are identified. As a result of this study, the following conclusions can be drawn. The sparsity structures that can be harnessed are conditional on the radio propagation environments. In TDD massive MIMO, uplink training inherently has more sparsity structures to exploit as high-dimensional channels are jointly recovered at the BS. On the contrary, in the FDD mode, the desired channel is normally recovered at the UE where utilizable sparsity structures are limited. Finally, based upon existing approaches, we have identified the potential research problems in need of further investigation.

References

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

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

[3] G. Bartoli *et al.*, "Beamforming for small cell deployment in LTE-advanced and beyond," *IEEE Wireless Commun.*, vol. 21, no. 2, pp. 50–56, Apr. 2014.

[4] J. Choi, D. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

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

[6] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," in *Proc. IEEE*, vol. 98, no. 6, Jun. 2010, pp. 1058–1076.

[7] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

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

[9] J. W. Kang, Y. Whang, H. Y. Lee, and K. S. Kim, "Optimal pilot sequence design for multi-cell MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3354–3367, Oct. 2011.

[10] D. Love *et al.*, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[11] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE*. Upper Saddle River, NJ, USA: Prentice-Hall, 2010.

[12] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[13] J.-C. Shen, J. Zhang, and K. B. Letaief, "User capacity of pilot-contaminated TDD massive MIMO systems," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 3713–3718.

[14] J.-C. Shen, J. Zhang, and K. B. Letaief, "Downlink user capacity of massive MIMO under pilot contamination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3183–3193, Jun. 2015.

[15] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[16] K. Zheng *et al.*, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1738–1760, 3rd Quart. 2015.

[17] X. Rao and V. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.

[18] J. Poutanen *et al.*, "Significance of common scatterers in multi-link indoor radio wave propagation," in *Proc. 4th Eur. Conf. Antennas Propag.*, Barcelona, Spain, Apr. 2010, pp. 1–5.

[19] P.-H. Kuo, H. Kung, and P.-A. Ting, "Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Shanghai, China, Apr. 2012, pp. 492–497.

[20] J.-C. Shen, J. Zhang, E. Alsusa, and K. B. Letaief, "Compressed CSI acquisition in FDD massive MIMO with partial support information," in *Proc. IEEE Int. Conf. Commun.*, London, U.K., Jun. 2015.

[21] W. Bajwa, J. Haupt, G. M. Raz, S. Wright, and R. Nowak, "Toeplitz-structured compressed sensing matrices," in *Proc. IEEE/SP 14th Workshop Stat. Signal Process.*, Madison, WI, USA, Aug. 2007, pp. 294–298.

[22] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.

[23] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, "On projection matrix optimization for compressive sensing systems," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2887–2898, Jun. 2013.

[24] 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.

[25] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with Interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.

[26] J. Foutz, A. Spanias, and M. K. Banavar, *Narrowband Direction of Arrival Estimation for Antenna Arrays*. San Rafael, CA, USA: Morgan & Claypool, 2008.

[27] S. L. H. Nguyen and A. Ghrayeb, "Compressive sensing-based channel estimation for massive multiuser MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Shanghai, China, Apr. 2013, pp. 2890–2895.

[28] T. Marzetta, "How much training is required for multiuser MIMO," in *Proc. 40th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2006, pp. 359–363.

[29] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, Mar. 2015.

[30] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.

[31] Z. Zhang and B. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 912–926, Sep. 2011.

[32] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

[33] J. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, Oct. 2013.

[34] J. Zhang *et al.*, "Pilot contamination elimination for large-scale multiple-antenna aided OFDM systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 759–772, Oct. 2014.

[35] J. Lofgren, L. Liu, O. Edfors, and P. Nilsson, "Improved matching-pursuit implementation for LTE channel estimation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 1, pp. 226–237, Jan. 2014.

[36] P. Maechler, P. Greisen, N. Felber, and A. Burg, "Matching pursuit: Evaluation and implementation for LTE channel estimation," in *Proc. IEEE Int. Symp. Circuits Syst.*, Paris, France, May 2010, pp. 589–592.

[37] P. Maechler *et al.*, "Implementation of greedy algorithms for LTE sparse channel estimation," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2010, pp. 400–405.

[38] Y. Shi, J. Zhang, B. O. Donoghue, and K. B. Letaief, "Large-scale convex optimization for dense wireless cooperative networks," *IEEE Trans. Signal Process.*, to be published.

[39] S. Payami and F. Tufvesson, "Channel measurements and analysis for very large array systems at 2.6 GHz," in *Proc. Eur. Conf. Antennas Propag.*, Prague, Czech Republic, Mar. 2012, pp. 433–437.

[40] X. Gao, F. Tufvesson, and O. Edfors, "Massive MIMO channels—Measurements and models," in *Proc. 47th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 280–284.

[41] S. Wu, C.-X. Wang, E.-H. Aggoune, M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, Jun. 2014.

[42] S. Wu *et al.*, "A non-stationary wideband channel model for massive MIMO communication systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1434–1446, Mar. 2015.

[43] J.-S. Jiang and M. Ingram, "Spherical-wave model for short-range MIMO," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1534–1541, Sep. 2005.

Juei-Chin Shen (S'10–M'14) received the Ph.D. degree in communication engineering from The University of Manchester, Manchester, U.K., in 2013.

In the same year, he joined the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong, as a Research Associate. His research interests include statistical signal processing, cognitive radio, spectrum sensing, compressed sensing, massive multiple-input–multiple-output systems, millimeter-wave communications, and random matrix theory.

Dr. Shen has been a corecipient of the IEEE PIMRC 2014 Best Paper Award.

Jun Zhang (S'06–M'10–SM'15) received the B.Eng. degree in electronic engineering from the University of Science and Technology of China, Hefei, China, in 2004, the M.Phil. degree in information engineering from The Chinese University of Hong Kong, Hong Kong, in 2006, and the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, Austin, TX, USA, in 2009.

He is currently a Research Assistant Professor with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong. He has coauthored *Fundamentals of LTE* (Prentice Hall, 2010). His research interests include wireless communications and networking, green communications, and signal processing.

Dr. Zhang served as a MAC Track Cochair for IEEE WCNC 2011. He was a recipient of the 2014 Best Paper Award from the EURASIP Journal on Advances in Signal Processing and the PIMRC 2014 Best Paper Award.

Kwang-Cheng Chen (M'89–SM'93–F'07) received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1983 and the M.S. and Ph.D. degrees from the University of Maryland, College Park, MD, USA, in 1987 and 1989, respectively, all in electrical engineering.

From 1987 to 1998, he was with SSE; COMSAT; IBM Thomas J. Watson Research Center; and National Tsing Hua University, Hsinchu, Taiwan, working on mobile communications and networks. From 1998 to 2006, he was a Distinguished Professor with National Taiwan University, and also served as the Director, Graduate Institute of Communication Engineering, Director, Communication Research Center, and the Associate Dean for Academic Affairs, College of Electrical Engineering and Computer Science from 2009 to 2015. Since 2016, he has been with the Department of Electrical Engineering, University of South Florida, Tampa, FL, USA. His recent research interests include wireless networks, cybersecurity, cyberphysical systems, social networks, and data analytics.

Prof. Chen has been actively involving in the organization of various IEEE conferences as General/TPC chair/Cochair, and has served in editorial roles with a few IEEE publications. He also actively participates in and has contributed essential technology to various IEEE 802, Bluetooth, and LTE and LTEA wireless standards. He was the recipient of numerous awards including the 2011 IEEE COMSOC WTC Recognition Award, the 2014 IEEE Jack Neubauer Memorial Award, and the 2014 IEEE COMSOC AP Outstanding Paper Award.

Khaled B. Letaief (S'85–M'86–SM'97–F'03) received the B.S. (with distinction), M.S., and Ph.D. degrees from Purdue University, West Lafayette, IN, USA, in 1984, 1986, and 1990, respectively, all in electrical engineering.

From January 1985 and as a Graduate Instructor with Purdue University, he taught courses in communications and electronics. From 1990 to 1993, he was a Faculty Member with The University of Melbourne, Melbourne, Australia. Since 1993, he has been with Hong Kong University of Science and Technology (HKUST), Kowloon, Hong Kong, where he is currently a Chair Professor and the Dean of Engineering, with expertise in wireless communications and networks. In these areas, he has over 500 journal and conference papers and has given invited keynote talks and courses all over the world. He is the holder of 13 patents, including 11 U.S. patents.

Dr. Letaief is a Fellow of The Hong Kong Institution of Engineers. He has served as an Elected Member of the IEEE Communications Society (ComSoc) Board of Governors, an IEEE Distinguished Lecturer, an IEEE ComSoc Treasurer, and an IEEE ComSoc Vice President for Conferences. He is currently serving as the IEEE ComSoc Vice President for Technical Activities as a member of the IEEE Product Services and Publications Board and is a member of the IEEE Fellow Committee. He serves as a consultant for different organizations and is the founding Editor-in-Chief of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has served on the editorial board of other prestigious journals, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Wireless Series (as Editor-in-Chief). He has been involved in organizing a number of major international conferences, including WCNC'07 in Hong Kong, ICC'08 in Beijing, ICC'10 in Cape Town, TTM'11 in Hong Kong, and ICCC'12 in Beijing. He has been a dedicated teacher committed to excellence in teaching and scholarship. He was a recipient of the Mangoon Teaching Award from Purdue University in 1990, the HKUST Engineering Teaching Excellence Award, and the Michael Gale Medal for Distinguished Teaching (highest university-wide teaching award at HKUST). He was also the recipient of many other distinguished awards, including the 2007 IEEE Communications Society Publications Exemplary Award, the 2009 IEEE Marconi Prize Award in Wireless Communications, the 2010 Purdue University Outstanding Electrical and Computer Engineer Award, the 2011 IEEE Communications Society Harold Sobol Award, the 2011 IEEE Wireless Communications Technical Committee Recognition Award, and 11 IEEE Best Paper Awards. He is also recognized by Thomson Reuters as an ISI Highly Cited Researcher.