Multiple Symbol Differential Detection for Noncoherent Communications with Large-Scale Antenna Arrays

Yue Wang, *Member, IEEE* and Zhi Tian, *Fellow, IEEE*

***Abstract*—For communications over large-scale antennas, noncoherent differential detection is an attractive option to avoid expensive channel estimation, but cannot maximally collect the performance benefits as the antenna number increases. For a desired performance-complexity tradeoff, this letter develops multiple symbol differential detection (MSDD) with simple implementation for single-input multiple-output (SIMO) systems, which jointly detects a block of symbols within the channel coherence time. The generalized likelihood ratio test criterion is adopted that maximizes the likelihood function over both the information symbols and the unknown channels. Simple formulas of the MSDD-SIMO detector are derived, whose performance is competitive to training-based coherent detection.**

***Index Terms*—Multiple symbol differential detection, noncoherent communications, generalized likelihood ratio test, large-scale antenna arrays.**

I. Introduction

IN contemporary wireless systems such as Internet of Things and millimeter-wave communications, large-scale antenna arrays have been considered as a promising technique to provide enhanced performance beyond traditional small-scale antenna array implementations in terms of increased link reliability and data rate [1], [2]. As the number of antennas increases, great opportunities arise to collect much enhanced antenna gain, in the form of either diversity gain for rich scattering channels or array gain for channels with sparse angular directivity [3], [4]. On the other hand, it is increasingly challenging to acquire the channel state information (CSI) required by coherent communications. This is because the conventional channel estimation techniques applicable to small-scale antenna systems cannot be effectively applied due to the large number of unknown channel coefficients, which may result in prohibitively long training time and huge power consumption [3]. For channels with sparse scattering such as in millimeter-wave systems, compressive sensing (CS) techniques have been applied in channel estimation to reduce training resources [5]–[7]. Still, they may incur high computational complexity in the process of CS-based signal recovery, which is related to the problem size decided by the number of antennas [8], [9]. Besides, CS methods are not applicable for radio frequency systems experiencing rich scattering. A simple alternative using beam sweeping has been proposed for low-complexity training in millimeter-wave systems [10]. However, when there are not enough training resources in practical systems, the problem of pilot contamination may occur and cause non-negligible channel estimation errors [11], [12], which then severely degrades the performance of training-based coherent detection [13].

The authors are with the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030, USA (e-mail: ywang56@gmu.edu; ztian1@gmu.edu).

Alternatively, noncoherent communication, as an efficient transmission paradigm without need for instantaneous channel estimation, utilizes the autocorrelation of the received signals to demodulate data information in the absence of CSI. Therefore, it has attracted recent attention in the context of large-scale antenna array systems [14]–[17]. In [14], a system with a single transmit antenna and a large number of receive antennas is analyzed to shed light on its asymptotic performance under the assumption of an infinite number of antennas. Noncoherent detection is coupled with energy-based communications to construct a multiuser system [14]. Along this line, the constellation design is further studied for noncoherent large-scale antenna systems in [15], [16]. Since most of these works aim for theoretical understanding of the achievable system performance, they all require some form of channel statistical information for analysis, and focus on the asymptotic results as the number of antennas grows infinite. However, they adopt the conventional energy detection or differential detection, which are not as effective in collecting the antenna gain. In fact, as the number of antenna increases, conventional noncoherent detectors exhibit an increasing performance gap from coherent detection, which is a major technical hindrance that motivates this work. To this end, this letter focuses on the design of efficient noncoherent detection schemes for large-scale antenna systems, in the absence of any channel statistics. Such effort is found in [17], which proposes a decision-feedback differential detection scheme. However, it is developed for a proof-of-concept geometrical channel model, and it needs the channel statistics for its weighting process.

In this work, we develop multiple symbol differential detection (MSDD) to implement noncoherent communications for uplink (massive) single-input multiple-output (SIMO) transmissions, in the absence of channel statistical knowledge. The goal is to avoid any channel estimation, while effectively improve the detection performance of noncoherent detection for large-antenna systems. The idea of MSDD is to jointly detect a block of consecutive symbols from signal autocorrelations under the assumption of channel time-invariance at least within the data block size [18]. In this work, we collect measurements at the receiver by spatial autocorrelation, not temporal autocorrelation [19]. Then, we resort to the generalized likelihood ratio test (GLRT) criterion [20], in which the maximization of the likelihood function is performed over both the unknown symbols and the unknown channels. The derived formula of the general MSDD-SIMO detector is applicable for various modulation modes. And, such a general detector formulation can be further simplified in the case of constant-amplitude

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LWC.2017.2764018

constellations, e.g. M-ary phase shift keying (PSK). The proposed MSDD solution allows simple hardware implementation and enables flexible options for performance-complexity tradeoffs in large-antenna communication systems. It can effectively collect the antenna gain to considerably outperform the conventional (one-symbol) differential detection (DD) in both sparse and rich scattering channel environments. Its bit error rate (BER) performance approaches that of the idealized coherent detection (CD) with perfect channel information, without invoking complicated channel estimation or subject to pilot contamination.

Notations: a denotes a scalar, $\mathbf{a}$ is a vector, $\mathbf{A}$ represents a matrix, and $\mathbb{A}$ means a set. $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ are the transpose, conjugate, and conjugate transpose of a matrix or vector, respectively. The operation vec$(\cdot)$ stacks all the columns of a matrix into a vector, tr$(\mathbf{A})$ calculates the trace of $\mathbf{A}$, and $\mathcal{R}(\cdot)$ returns the real part of a complex argument. $|\mathbb{A}|$ denotes the cardinality of $\mathbb{A}$.

II. Signal Model

In the context of noncoherent detection, a sequence of independent information-bearing symbols $a_i \in \mathbb{M}$ are differentially encoded into the transmit symbols $b_i \in \mathbb{M}$ via the rule $b_i = a_i b_{i-1}$, where $\mathbb{M}$ denotes the M-ary constellation set, e.g., $\mathbb{M} = \{e^{j2\pi i/M} | \, i = 0, 1, \cdots, M-1\}$ for M-ary PSK.

Consider a SIMO system with one transmit antenna and N_r receive antennas. Assume a block-fading channel model, where the channel is time-invariant within the duration of multiple consecutive symbols in the same data block. The input-output signal model over the block-fading SIMO channel with frequency flat fading can be written as

$$\mathbf{Y} = \mathbf{h}\mathbf{b}^T + \mathbf{W} \tag{1}$$

where $\mathbf{Y}$ is the $N_r \times (N+1)$ received signal collected from the N_r receive antennas over the block length of $N+1$ symbols, the $N_r \times 1$ vector $\mathbf{h}$ denotes the block fading SIMO channels, the transmit signal vector $\mathbf{b}$ contains the initial symbol b_0 and the following N differentially encoded symbols in the form of $\mathbf{b} = [b_0 \; b_1 \; b_2 \; \cdots \; b_N]^T$, and $\mathbf{W}$ represents additive white Gaussian noise (AWGN) with identical variance σ_n^2.

III. Multiple Symbol Differential Detection

In developing an MSDD scheme for noncoherent communications, we seek to jointly detect the N consecutive information symbols $\mathbf{a} = [a_1 \; a_2 \; \cdots \; a_N]^T$ from the received $\mathbf{Y}$ in (1), without acquiring the channel $\mathbf{h}$ through expensive training. Because of differential encoding, we have $b_i = b_0 \prod_{k=1}^{i} a_k$, $i = 1, \cdots, N$, which turns (1) into

$$\begin{aligned} \mathbf{Y} &= \mathbf{h}[b_0 \; b_1 \; b_2 \; \cdots \; b_N] + \mathbf{W} \\ &= \mathbf{h} b_0 \left[1 \; a_1 \; a_1 a_2 \; \cdots \; \prod_{k=1}^{N} a_k \right] + \mathbf{W} \\ &= \mathbf{v}\boldsymbol{\alpha}^T + \mathbf{W} \end{aligned} \tag{2}$$

where $\mathbf{v} = \mathbf{h} b_0$ is the unknown channel scaled by the predefined initial symbol b_0, and $\boldsymbol{\alpha} = [1 \; a_1 \; a_1 a_2 \; \cdots \; \prod_{k=1}^{N} a_k]^T$ contains the unknown information symbols.

Since $\mathbf{v}$ is unknown in the absence of the channel knowledge $\mathbf{h}$, a reasonable alternative to the optimal maximum likelihood (ML) criterion for detecting $\mathbf{a}$ is the GLRT approach. When $\mathbf{W}$ is AWGN, vec$(\mathbf{Y} - \mathbf{v}\boldsymbol{\alpha}^T)$ follows the multivariate normal distribution with a probability density function

$$f(\mathbf{Y}; \mathbf{v}, \boldsymbol{\alpha}) = \frac{1}{\sqrt{2\pi\sigma_n^2}^{N_r N}} \exp\left(- \frac{\left(\text{vec}\left(\mathbf{Y} - \mathbf{v}\boldsymbol{\alpha}^T\right)\right)^H \left(\text{vec}\left(\mathbf{Y} - \mathbf{v}\boldsymbol{\alpha}^T\right)\right)}{2\sigma_n^2} \right). \tag{3}$$

Then, the task of GLRT-based noncoherent MSDD amounts to maximizing the following log-likelihood metric over both $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{v}}$:

$$\begin{aligned} \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}] &= (\text{vec}(\mathbf{Y}))^H \, \text{vec}(\tilde{\mathbf{v}}\tilde{\boldsymbol{\alpha}}^T) \\ &+ \left(\text{vec}(\tilde{\mathbf{v}}\tilde{\boldsymbol{\alpha}}^T)\right)^H \text{vec}(\mathbf{Y}) \\ &- \left(\text{vec}(\tilde{\mathbf{v}}\tilde{\boldsymbol{\alpha}}^T)\right)^H \text{vec}(\tilde{\mathbf{v}}\tilde{\boldsymbol{\alpha}}^T), \end{aligned} \tag{4}$$

where $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{v}}$ are candidate values of $\mathbf{a}$ and $\mathbf{v}$ respectively, and $\tilde{\boldsymbol{\alpha}} = [1 \; \tilde{a}_1 \; \tilde{a}_1\tilde{a}_2 \; \cdots \; \prod_{k=1}^{N} \tilde{a}_k]^T$ is the candidate value of $\boldsymbol{\alpha}$. Using the properties of inner products, (4) can be reformulated to yield the following equivalent metric:

$$\begin{aligned} \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}] &= \langle \tilde{\mathbf{v}}, \mathbf{Y}^* \tilde{\boldsymbol{\alpha}} \rangle + \langle \tilde{\mathbf{v}}^*, \mathbf{Y} \tilde{\boldsymbol{\alpha}}^* \rangle \\ &- \left(1 + \sum_{n=1}^{N} \prod_{k=1}^{n} |\tilde{a}_k|^2 \right) \langle \tilde{\mathbf{v}}^*, \tilde{\mathbf{v}} \rangle. \end{aligned} \tag{5}$$

Accordingly, the GLRT rule for noncoherent detection of the multiple information symbols $\mathbf{a}$ is given by

$$\hat{\mathbf{a}} = \arg\max_{\tilde{\mathbf{a}}} \left\{ \max_{\tilde{\mathbf{v}}} \{ \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}] \} \right\}. \tag{6}$$

Since $\mathbf{v}$ is a nuisance parameter, we solve (6) by first keeping $\tilde{\mathbf{a}}$ fixed and computing

$$\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}] = \max_{\tilde{\mathbf{v}}} \{ \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}] \}. \tag{7}$$

According to Taylor's theorem, we resort to the variational technique by imposing

$$\tilde{\mathbf{v}} = \mathbf{v}_0 + \lambda \boldsymbol{\epsilon}, \tag{8}$$

where $\mathbf{v}_0 \in \mathbb{C}^{N_r}$ denotes the optimal solution to $\mathbf{v} \in \mathbb{C}^{N_r}$, $\boldsymbol{\epsilon} \in \mathbb{C}^{N_r}$ measures the distortion of $\tilde{\mathbf{v}}$ from $\mathbf{v}_0$, and $\lambda \in \mathbb{R}$ is the coefficient of the distortion. After substituting (8) into (5) and then taking the first-order derivative of $\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}]$ with respect to λ at zero, we obtain

$$\begin{aligned} \forall \boldsymbol{\epsilon}, \quad \frac{\partial}{\partial \lambda} \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}, \tilde{\mathbf{v}}] \Big|_{\lambda=0} &= \langle \boldsymbol{\epsilon}, (\mathbf{Y}\tilde{\boldsymbol{\alpha}}^*)^* \rangle + \langle \boldsymbol{\epsilon}^*, \mathbf{Y}\tilde{\boldsymbol{\alpha}}^* \rangle \\ &- \left(1 + \sum_{n=1}^{N} \prod_{k=1}^{n} |\tilde{a}_k|^2 \right) \left(\langle \boldsymbol{\epsilon}, \mathbf{v}_0^* \rangle + \langle \boldsymbol{\epsilon}^*, \mathbf{v}_0 \rangle \right). \end{aligned} \tag{9}$$

Setting the derivative in (9) to zero yields the optimal $\mathbf{v}$ as

$$\mathbf{v}_0 = \frac{1}{1 + \sum_{n=1}^{N} \prod_{k=1}^{n} |\tilde{a}_k|^2} \mathbf{Y} \tilde{\boldsymbol{\alpha}}^*. \tag{10}$$

Note that (10) yields the channel estimate $\mathbf{h} = b_0^{-1} \mathbf{v}_0$ given b_0 and after obtaining the estimate of $\boldsymbol{\alpha}$, even though the channel estimate is not explicitly used during noncoherent detection.

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/LWC.2017.2764018

Next, substituting $\tilde{\mathbf{v}}$ in (5) by $\mathbf{v}_0$ in (10) yields

$$\begin{aligned}\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}] &= \frac{1}{1+\sum_{n=1}^{N}\prod_{k=1}^{n}|\tilde{a}_k|^2}\,\tilde{\boldsymbol{\alpha}}^T\mathbf{Y}^H\mathbf{Y}\tilde{\boldsymbol{\alpha}}^* \\ &= \frac{1}{1+\sum_{n=1}^{N}\prod_{k=1}^{n}|\tilde{a}_k|^2}\,\mathrm{tr}\left(\left(\tilde{\boldsymbol{\alpha}}^*\tilde{\boldsymbol{\alpha}}^T\right)\left(\mathbf{Y}^H\mathbf{Y}\right)\right) \\ &= \frac{1}{1+\sum_{n=1}^{N}\prod_{k=1}^{n}|\tilde{a}_k|^2}\left(\sum_{l=0}^{N}\prod_{p=0}^{l}|\tilde{a}_p|^2 z_{ll}\right. \\ &\quad \left.+\sum_{i=1}^{N}\sum_{l=0}^{i-1}\prod_{p=0}^{l}|\tilde{a}_p|^2\, 2\,\mathcal{R}\left(\prod_{k=1}^{i-l}\tilde{a}_{l+k}\, z_{il}\right)\right), \quad (11)\end{aligned}$$

where $z_{il} = \mathbf{y}_i^H\mathbf{y}_l$ is the autocorrelation of the received signal with $\mathbf{y}_i$ being the i-th column of $\mathbf{Y}$, and $|\tilde{a}_0|^2 = 1$ is introduced just for the convenience of mathematical notation. In reaching the third equation of (11), we use the property that $\mathrm{tr}\left(\left(\tilde{\boldsymbol{\alpha}}^*\tilde{\boldsymbol{\alpha}}^T\right)\left(\mathbf{Y}^H\mathbf{Y}\right)\right)$ equals the sum of all entries in the Hadamard product $\left(\tilde{\boldsymbol{\alpha}}^*\tilde{\boldsymbol{\alpha}}^T\right)\circ\left(\mathbf{Y}^H\mathbf{Y}\right)^T$, while the sum of the diagonal entries and that of off-diagonal entries of the Hadamard product are organized in the first and second terms, respectively. According to (6) and with $\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}]$ given by (11), the proposed MSDD rule can be formalized as

$$\hat{\mathbf{a}} = \arg\max_{\tilde{\mathbf{a}}}\{\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}]\}, \quad \text{where } \Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}] \text{ is in (11).} \quad (12)$$

It is worth noting that (11) can be significantly simplified for the constant-amplitude modulation schemes. Without loss of generality, we consider the case of normalized amplitude $|\tilde{a}_i| = 1,\ i \in \{1,\cdots,N\}$, which simplifies (11) to

$$\Lambda[\mathbf{Y} \mid \tilde{\mathbf{a}}] = \frac{1}{1+N}\sum_{l=0}^{N} z_{ll} + \frac{2}{1+N}\sum_{i=1}^{N}\sum_{l=0}^{i-1}\mathcal{R}\left(\prod_{k=1}^{i-l}\tilde{a}_{l+k}\, z_{il}\right). \quad (13)$$

Further, since the first summation term and the constant coefficient of the second summation term in (13) are irrelevant to $\tilde{\mathbf{a}}$, (12) can be reformulated as a simple MSDD-SIMO detector for all the constant-amplitude constellation cases:

$$\hat{\mathbf{a}} = \arg\max_{\tilde{\mathbf{a}}}\left\{\sum_{i=1}^{N}\sum_{l=0}^{i-1}\mathcal{R}\left(\prod_{k=1}^{i-l}\tilde{a}_{l+k}\, z_{il}\right)\right\}, \quad \text{for constant } |\tilde{a}_i|. \quad (14)$$

IV. Implementation and Complexity

Noticeably, the MSDD solution leads to an important benefit in terms of simplified hardware implementation for large-scale antenna systems with large N_r. As shown in Fig. 1, the receiver first collects autocorrelation values z's as inner products (denoted by $\odot$ in Fig. 1) of the received signals $\mathbf{y}$'s across antennas, where each inner product can be implemented using simple analog components such as shift registers, multipliers and adders. Then the MSDD rule is applied on the scalar z's only, rather than on $\mathbf{y}$'s of a large size N_r. The number of z's involved is a function of the fixed block size N instead of N_r, that is, $N(N+1)/2$ for (14) or $(N+2)(N+1)/2$ for (12).

The computational complexity of MSDD via exhaustive search is $\mathcal{O}(|\mathbb{M}|^N)$, which is exponential in N but independent of N_r. Reduced-complexity alternatives such as sphere decoding and Viterbi algorithms can be adopted to offer polynomial complexity [19], with a small drop in performance.

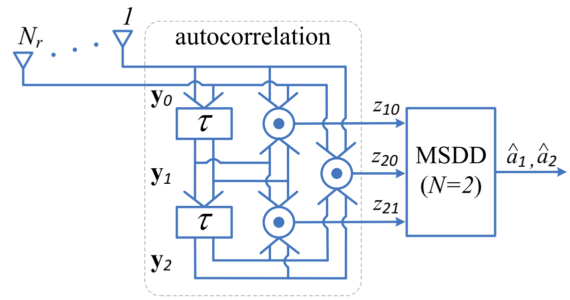

Fig. 1. Diagram of the MSDD-SIMO receiver with $N = 2$.

V. Numerical Results

This section presents simulation results to verify the performance of the proposed MSDD detector in terms of BER. In all simulation trials, the transmitter with a single antenna sends N consecutive binary PSK (BPSK) information-symbols[1], while the receiver with N_r antennas collects the signals passed through Rayleigh fading channels[2] and inflicted with AWGN noise as in (1). Two benchmark methods are simulated as well: DD which is actually a special case of MSDD with $N = 1$, and idealized CD under perfect channel knowledge.

Fig. 2 illustrates the BER results of MSDD, DD, and CD for various values of the signal-to-noise ratio (SNR) per bit, where the slope of the BER curve indicates the collected diversity order of the multi-antenna system. As expected, the proposed MSDD outperforms DD, but exhibits performance gaps from the idealized CD. In this sense, CD and DD provide the lower and upper bounds of the BER performance of MSDD respectively. To reveal the impact of N on the performance of the MSDD detector, Fig. 3 depicts how much the BER performance of MSDD improves as N increases. It shows that the BER of the MSDD detector quickly decreases as N begins to increase, and such improvement flattens out when N reaches a moderate value, say $N = 6$ in Fig. 3. The saturation value of N may vary slightly with N_r. This is because an increase in N_r would cause an enlarged performance gap between DD and CD. It then leaves more room for performance improvement by MSDD, which echoes the motivation to develop MSDD in this work. On the other hand, a small value of N leads to low complexity and simple implementation, as discussed in Section IV. Therefore, a moderate value of N is favored in the deployment of MSDD, which reflects an efficient performance-complexity tradeoff. In addition, choosing a small but effective value for the parameter N rationalizes the assumption of channel time-invariance within a data block size N, under which our MSDD solution is derived. Finally, it should be noted that

[1]Without loss of generality, we test our noncoherent transmission scheme in the simple BPSK case. Meanwhile, the proposed MSDD detectors in (12) and (14) are applicable for more general and complex cases, and we have tested higher-order modulations and observed similar performance results.

[2]The MSDD detectors are derived independently of any assumption on the channel propagation model or fading distribution, and hence are applicable to all channel conditions. Here simulation tests are carried out for rich scattering channels with independent and identically distributed Rayleigh fading, while similar test results are observed in other scenarios as well, e.g., the spatially sparse channel model in millimeter-wave systems and other fading distributions. Those results are not included due to space limit.

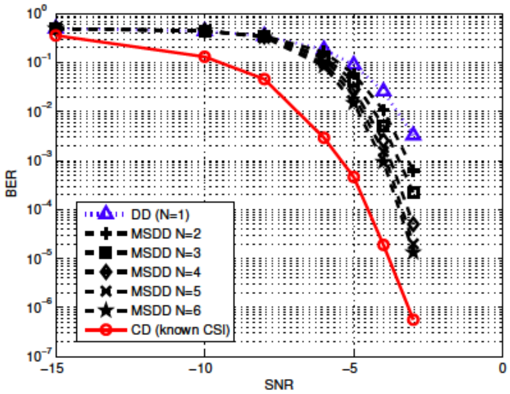

Fig. 2. BER of the proposed MSDD-SIMO noncoherent method versus SNR compared with that of DD and CD benchmarks in Rayleigh fading channels for BPSK when $N_r = 32$.

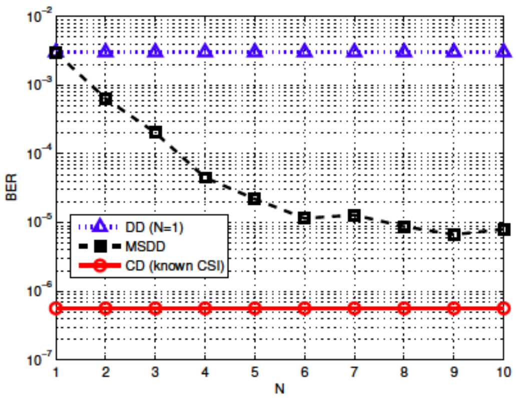

Fig. 3. BER of the proposed MSDD-SIMO noncoherent method versus N compared with that of DD and CD benchmarks in Rayleigh fading channels for BPSK when $N_r = 32$ and SNR= 3dB.

the performance gap between our noncoherent MSDD and the training-based CD will shrink, disappear, or even flip over in practice. This is because the BER curves of CD in both Fig. 2 and Fig. 3 are obtained under perfectly known CSI, yet the realistic performance of CD would suffer from the aggravated channel estimation errors in practical large-antenna systems.

VI. Conclusions

This work develops a new noncoherent solution for uplink SIMO transmissions based on the MSDD principle. According to the GLRT rule, we derive the MSDD formulas for general modulation constellations. By virtual of noncoherent detection, it bypasses the complicated channel estimation step, avoids the associated huge training overhead and high computational complexity, and allows a simple analog-digital hardware structure for large-scale antenna array systems. In terms of BER performance, it considerably outperforms the traditional differential detection, and is competitive to the coherent counterpart assuming perfect channel information. As a result, the proposed MSDD-based noncoherent solution offers a desired tradeoff between performance and complexity for future large-antenna communications. For future work, it is of interest to investigate the practical implementation of this MSDD principle in broad operating scenarios such as frequency selective channels and multiuser cases.

References

[1] 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.

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

[3] 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.

[4] J. G. Andrews, *et al*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

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

[6] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.

[7] Y. Wang, Z. Tian, S. Feng, and P. Zhang, "Efficient channel statistics estimation for millimeter-wave MIMO systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Shanghai, China, Mar. 2016, pp. 3411–3415.

[8] Y. Wang, Z. Tian, S. Feng, and P. Zhang, "A fast channel estimation approach for millimeter-wave massive MIMO systems," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Arlington, VA, USA, Dec. 2016, pp. 1413–1417.

[9] Y. Wang, P. Xu, and Z. Tian, "Efficient channel estimation for massive MIMO systems via truncated two-dimensional atomic norm minimization," in *Proc. IEEE Conf. Commun.*, Pairs, France, May 2017, pp. 1–6.

[10] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. H. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[11] 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.

[12] P. Liu, S. Jin, T. Jiang, Q. Zhang, and M. Matthaiou, "Pilot power allocation through user grouping in multi-cell massive MIMO systems," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1561–1574, Apr. 2017.

[13] 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.

[14] M. Chowdhury, A. Manolakos, and A. Goldsmith, "Scaling laws for noncoherent energy-based communications in the SIMO MAC," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1980–1992, Apr. 2016.

[15] A. Manolakos, M. Chowdhury, and A. Goldsmith, "Energy-based modulation for noncoherent massive SIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7831–7846, Nov. 2016.

[16] D. Kong, X. Xia, and T. Jiang, "A differential QAM detection in uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6371–6383, Sep. 2016.

[17] A. Schenk and R. F. H. Fisher, "Noncoherent detection in massive MIMO systems," in *Proc. Int. ITG/IEEE Workshop on Smart Antennas*, Stuttgart, Germany, Mar. 2013, pp. 1–8.

[18] M. K. Simon and M. S. Alouini, "Multiple symbol differential detection with diversity reception," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1312–1319, Aug. 2001.

[19] V. Lottici, and Z. Tian, "Multiple Symbol Differential Detection for UWB Communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1656–1666, May 2008.

[20] Z. Tian and G. B. Giannakis, "A GLRT approach to data-aided timing acquisition in UWB radios–part I: algorithms," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2956–2967, Nov. 2005.