Low-Complexity MIMO Detection for Approaching Near-ML Performance

Yinman Lee
Department of Electrical Engineering
National Chi Nan University
Puli, Taiwan, R.O.C.
Email: ymlee@ncnu.edu.tw

***Abstract*—The research of finding better ways for multiple-input multiple-output (MIMO) signal detection is important and still goes on. In this paper, we propose a low-complexity MIMO detection method mainly with basic and hardware-friendly operations, i.e., zero-forcing (ZF), successive interference cancelation (SIC), double-symbol maximum likelihood (ML) detection, and M-algorithm. We show that it can approach the optimal error rate employing quadrature phase-shift keying (QPSK) signals with affordable complexity in many cases. Simulation results show that in terms of giving a good trade-off between the error rate and implementation complexity, the proposed detection schemes can significantly outperform some contemporary detectors, including the ordered SIC-ZF method and the conventional ML detector employing QR-decomposition and M-algorithm (QRM-MLD), and own an error rate quite close to that of the optimal ML detector.**

***Index Terms*—Multiple-input multiple-output (MIMO), spatial multiplexing, zero-forcing (ZF) detection, maximum-likelihood (ML) detection, M-algorithm, successive interference cancelation (SIC).**

I. Introduction

Multiple-input multiple-output (MIMO) spatial multiplexing systems have recently shown great potential to dramatically increase the spectrum efficiency in wireless communications. In practice, the main challenge for such MIMO systems is the receiver design that can obtain good error-rate performance with acceptable computational complexity. It is well-understood that the maximum likelihood (ML) detector is able to provide the optimal error-rate performance, but has a disadvantage of extremely high computational complexity. Linear detectors, mainly based on the zero-forcing (ZF) or minimum mean-squared error (MMSE) criteria, are low in complexity but poor in error rate. Ordered successive interference cancelation (SIC) detectors, which extract the transmitted symbols one-by-one according to the post-detection signal-to-noise ratio (SNR) and perform successive interference elimination, can achieve better error-rate performance with increased complexity. Nevertheless, there is still a big performance gap between these suboptimal detectors and the ML one [1].

In the literature, various advanced near-ML detection techniques were studied to approach the optimal error rate with reduced computational complexity, such as the sphere decoder (SD) [2], [3], semidefinite relaxation (SDR) detector [4], [5], and ML detector employing QR-decomposition and M-algorithm (QRM-MLD) [6], etc. However, it is known that many near-ML methods require precise and complicated log-likelihood ratio (LLR) calculation, optimization algorithms, and/or matrix computations which are not desirable for hardware implementation in general. Partly based on our previous ideas [7], in this paper, we propose a low-complexity MIMO detection method mainly with basic and hardware-friendly operations, i.e., ZF, SIC, double-symbol ML detection, and M-algorithm. We show that it can approach the optimal error rate employing quadrature phase-shift keying (QPSK) signals with affordable complexity in many cases. Essentially, a 2×2 nulling mechanism is utilized, and only the “painless” 2×2 matrix inverse is involved for ZF. Besides, the dimension of ML detection is restricted to be two. All these limit the increase of complexity as the number of transmit/receive antennas goes large. Here, we need to stress that this mixture is completely different from those previous approaches which usually employ ZF to reduce the number of candidates in the ML search space. For our cases, ZF is used to lower the dimension of the signal model. Specifically, the included ZF stage comprises a series of 2×2 nulling steps which decompose the MIMO system step by step. Then, low-complexity double-symbol ML detection becomes possible for obtaining candidates for decisions. M-algorithm and SIC are used to enhance the error rate and make it come near the optimum. Simulation results show that in terms of giving a good trade-off between the error rate and implementation complexity, the proposed detection schemes can significantly outperform some contemporary detectors, including the ordered SIC-ZF method and the conventional QRM-MLD, and own an error rate quite close to that of the optimal ML detector.

II. MIMO Signal Model, Conventional ML, ZF and SIC Detection

First, we consider a wireless communication system with N antennas at the transmitter and N antennas at the receiver, assuming $N = 2L$ and $L \geq 2$. We define $h_{N_r N_t}$ to represent the flat fading complex channel response from transmit antenna N_t to receive antenna N_r, with $N_t, N_r = 1, 2, \cdots, N$.

This work was supported by the Ministry of Science and Technology, Taiwan, R.O.C., under Grant MOST 103-2221-E-260 -010-.

978-1-4799-5832-0/14/$31.00 ©2014 IEEE

Then the complete $N \times N$ channel matrix can be presented as

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NN} \end{bmatrix}. \quad (1)$$

The spatially-multiplexed data symbols from the N transmit antennas can be collected into an $N \times 1$ complex-valued vector, denoted as $\mathbf{s} = [s_1 \ \ s_2 \ \ \cdots \ s_N]^T \in \mathbb{S}^N$, where $\mathbb{S}$ denotes the set of constellation points for QPSK signals. They are sent over the $N \times N$ channel environment. The equivalent baseband signals obtained by the antennas at the receiver yields another $N \times 1$ complex-valued vector $\mathbf{y} = [y_1 \ \ y_2 \ \ \cdots \ y_N]^T$, given by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{z} \quad (2)$$

where $\mathbf{z}$ is an $N \times 1$ complex Gaussian noise vector with zero mean and equal variance in each dimension. In addition, all the noise components are assumed to be independent.

A. Optimal ML Detection

From (2), the ML detection of the transmitted symbol vector can be written as [1]

$$\hat{\mathbf{s}}_{\text{ML}} = \arg \min_{\mathbf{s} \in \mathbb{S}^N} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|^2 . \quad (3)$$

The major challenge of this ML criterion is the huge cardinality $\mathbb{S}^N$ in the search, which increases exponentially with the number of transmit antenna N.

B. Linear ZF Detection

For the linear ZF detection, the received signal vector $\mathbf{r}$ is multiplied by an equalization matrix which is equal to the inverse (or pseudo-inverse, in general) of the channel matrix. In the considered MIMO scenario, after this equalization, the estimate of its transmitted symbol vector is presented as [1]

$$\hat{\mathbf{s}}_{\text{ZF}} = \mathbf{H}^{-1}\mathbf{y} = \mathbf{s} + \tilde{\mathbf{z}} \quad (4)$$

where $\tilde{\mathbf{z}} = \mathbf{H}^{-1}\mathbf{z}$ is the resultant noise vector which disturbs the transmitted symbol vector. From (4), we see that the crosstalk generated by the channel in the received signal is removed. However, the noise may be amplified and become colored. This largely degrades the error-rate performance.

C. SIC Detection

In contrast to the linear detection where all symbols are detected jointly at the same time, the SIC detection uses a serial decision-feedback approach to extract the transmitted symbol in each layer of the MIMO channel separately. When a symbol has been detected, an estimate of the corresponding contribution to the received signal vector $\mathbf{y}$ is subtracted; this result is then used to detect the next symbol, etc. In the absence of detection errors, SIC progressively cleans $\mathbf{y}$ from the interference corresponding to the symbols already detected. To extract a specific symbol, the symbols that have not been detected yet are equalized according to the ZF or MMSE criteria. Note that error propagation may be a problem because incorrect decisions actually increase the interference when detecting subsequent symbols. To deal with this, the detection is usually ordered according to the post-detection SNR to lessen the error-propagation effect.

III. Proposed ZF/ML-M/SIC (ZMS) Detection

In this section, we address a new way to tackle the signal detection in MIMO spatial multiplexing systems with a good compromise between the error-rate performance and the required amount of computation. Our focus here is to develop a hardware-friendly method in which only traditional filtering and weighting operations are utilized. To be more specific, our proposed method, named ZF/ML-M/SIC (ZMS) detection, includes two kinds of stages for processing, which are clarified separately in the following.

A. ZF Stage with Multiple 2×2 Nulling Steps

Conventionally, to extract or separate the desired data symbol from interference, ZF can be used. To do so, matrix inverse on the order of $O(N^3)$ is generally required. It is obvious that the complexity is often not desirable as the dimension of the signal model N goes large. To avoid this, we propose a special nulling mechanism for the implementation of ZF, which only involves matrix inverse with size 2×2 no matter how large N is. This essentially lowers the complexity and fulfills the task for obtaining some particular data symbols at the same time.

For notational consistency in the derivation, we let $\mathbf{y}^{(1)} = \mathbf{y}$, $\mathbf{H}^{(1)} = \mathbf{H}$, $\mathbf{s}^{(1)} = \mathbf{s}$, and $\mathbf{z}^{(1)} = \mathbf{z}$. Neglecting the noise term, we first illustrate how the channel matrix is separated into blocks at the ZF stage. As shown in the left side of Fig. 1, the $N \times N$ channel matrix $\mathbf{H}^{(1)}$ is partitioned, from up to down and from left to right, with dimension $(N-2) \times (N-2)$, $(N-2) \times 2$, $2 \times (N-2)$, and 2×2. We name them $\mathbf{H}_a^{(1)}$, $\mathbf{H}_b^{(1)}$, $\mathbf{H}_c^{(1)}$, and $\mathbf{H}_d^{(1)}$, respectively. Next, we formulate a weighting matrix $\mathbf{W}^{(1)}$ with size $(N-2) \times 2$ to let $\mathbf{W}^{(1)}\mathbf{H}_d^{(1)}$ approach $\mathbf{H}_b^{(1)}$, i.e.,

$$\mathbf{W}^{(1)}\mathbf{H}_d^{(1)} = \mathbf{H}_b^{(1)} \quad (5)$$

with

$$\mathbf{W}^{(1)} = \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ \vdots & \vdots \\ w_{(N-2)1}^{(1)} & w_{(N-2)2}^{(1)} \end{bmatrix}. \quad (6)$$

To obtain $\mathbf{W}^{(1)}$, we perform

$$\mathbf{W}^{(1)} = \mathbf{H}_b^{(1)} \left(\mathbf{H}_d^{(1)}\right)^{-1} \quad (7)$$

in which only 2×2 matrix inverse is involved. Besides, we partition the received signal vector into $\mathbf{y}_a^{(1)} = [y_1 \ \ y_2 \ \ \cdots \ y_{N-2}]^T$ and $\mathbf{y}_b^{(1)} = [y_{N-1} \ \ y_N]^T$ with length $N-2$ and 2, respectively, and process these two signal vectors as

$$\mathbf{y}^{(2)} = \mathbf{y}_a^{(1)} - \mathbf{W}^{(1)}\mathbf{y}_b^{(1)}. \quad (8)$$

It is not difficult to see that the resultant received signal vector $\mathbf{y}^{(2)}$ in (8) contains the transmitted symbols s_1 to s_{N-2} only,

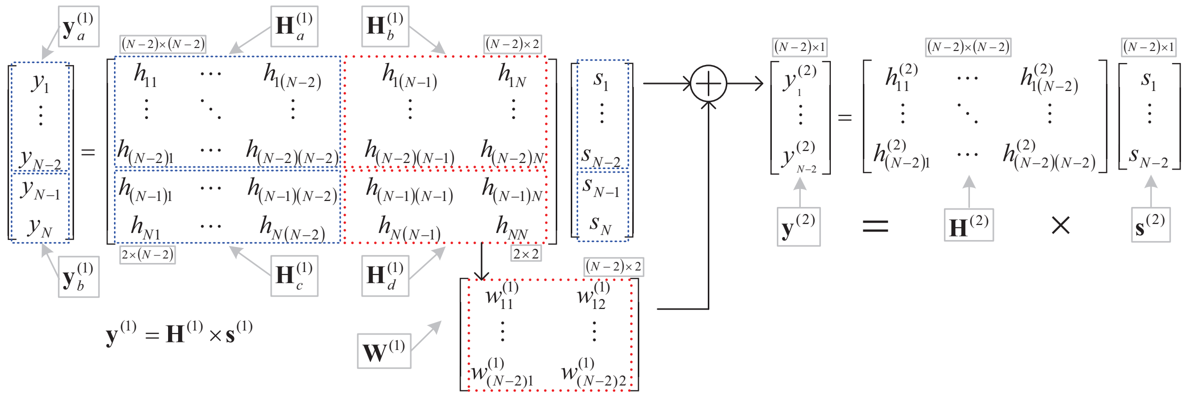

Fig. 1. First 2×2 nulling step for ZF stage.

and the effect of s_{N-1} and s_N is eliminated due to the nulling mechanism of $\mathbf{W}^{(1)}$. A new subsystem is then formed as that shown in the right side of Fig. 1. For this case, the channel matrix becomes

$$\mathbf{H}^{(2)} = \mathbf{H}_a^{(1)} - \mathbf{W}^{(1)}\mathbf{H}_c^{(1)} \tag{9}$$

with size $(N-2) \times (N-2)$. As a result, the dimension of the signal model is reduced from N to $N-2$. In such a way, we continue this 2×2 nulling step to obtain another new subsystem with dimension $N-4$ in which only the transmitted symbols s_1 to s_{N-4} are included. Accordingly, we observe that after $L-1$ steps, a new subsystem with dimension 2 can be obtained as

$$\begin{aligned}\mathbf{y}^{(L)} &= \mathbf{H}^{(L)}\mathbf{s}^{(L)} + \mathbf{z}^{(L)} \\ &= \begin{bmatrix} h_{11}^{(L)} & h_{12}^{(L)} \\ h_{21}^{(L)} & h_{22}^{(L)} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} z_1^{(L)} \\ z_2^{(L)} \end{bmatrix}\end{aligned} \tag{10}$$

where $\mathbf{H}^{(L)}$ is the resultant 2×2 channel matrix and $\mathbf{z}^{(L)}$ is the noise vector after performing 2×2 nulling steps $L-1$ times. Essentially, this subsystem only contains the transmitted symbols s_1 and s_2. In one word, through these specially-designed 2×2 nulling steps, the dimension of the MIMO system is reduced step-by-step. After obtaining the subsystem in (10), the ZF stage is accomplished.

B. Detection Stage with Double-Symbol ML Detection with M-Algorithm (ML-M) and SIC

Once the subsystem with size 2×2 as that derived above is obtained, the first double-symbol ML Detection with M-Algorithm (ML-M) Detection is triggered. We apply ML Detection operating on $\mathbf{y}^{(L)}$ in (10) for the detection of s_1 and s_2. Specifically, the ML detection as that shown in (3) is employed. Instead of keeping the best decisions for s_1 and s_2 only in the normal way for the conventional ML detection, in this detection, M groups of s_1 and s_2 are saved for SIC in the next subsystem with dimension 4. Then, the first double-symbol ML-M detection finishes.

With M groups of s_1 and s_2, we perform SIC M times. We subtract $[s_1 \;\; s_2]^T$ from the 4×4 subsystem containing s_1 to s_4 as

$$\begin{aligned}&\begin{bmatrix} y_1^{(L-1)} \\ y_2^{(L-1)} \\ y_3^{(L-1)} \\ y_4^{(L-1)} \end{bmatrix} - \begin{bmatrix} h_{11}^{(L-1)} & h_{12}^{(L-1)} \\ h_{21}^{(L-1)} & h_{22}^{(L-1)} \\ h_{31}^{(L-1)} & h_{32}^{(L-1)} \\ h_{41}^{(L-1)} & h_{42}^{(L-1)} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \\ &= \begin{bmatrix} h_{13}^{(L-1)} & h_{14}^{(L-1)} \\ h_{23}^{(L-1)} & h_{24}^{(L-1)} \\ h_{33}^{(L-1)} & h_{34}^{(L-1)} \\ h_{43}^{(L-1)} & h_{44}^{(L-1)} \end{bmatrix} \begin{bmatrix} s_3 \\ s_4 \end{bmatrix} + \begin{bmatrix} z_1^{(L-1)} \\ z_2^{(L-1)} \\ z_3^{(L-1)} \\ z_4^{(L-1)} \end{bmatrix}.\end{aligned} \tag{11}$$

We apply double-symbol ML-M detection again operating on the right side of (11) to obtain s_3 and s_4 based on the particular decisions s_1 and s_2. Note that the new channel matrix is 4×2 in size. After that, we have M^2 groups of decisions s_1 to s_4. Here, we choose to keep M most probable groups of decisions for the next subsystem with size 6×6. Likewise, the SIC and double-symbol ML-M detection steps come alternately until the original $N \times N$ system is reached, and s_{N-1} and s_N are detected by double-symbol ML-M detection. The best decisions s_1 to s_N are kept at last. Typically, in the $N \times N$ scenario, SIC is performed $M(L-1)$ times and double-symbol ML-M detection is performed L times, respectively, in total. This way, the transmitted data symbols are all detected, The detection stage and so the proposed ZMS detection end.

C. On Enhancing the Detection Performance

Through examining some characteristics of our detection approach, we introduce two ways to further improve the detection performance as follows.

- *E1:* First, for the proposed ZMS detection, we see that the detection order is that s_1 and s_2 come to be detected first, s_3 and s_4 come second, and s_5 and s_6 come third, and so on. Intuitively, for such SIC, signals with better quality should be placed and detected first to avoid error propagation. Therefore, before we start the actual detection process, we may order the channel with norm values from large to small for the original $N \times N$ system. Signals with the two largest channel-norm values will be placed in the first place as s_1 and s_2 in our signal model, and signals with the two smallest channel-norm values will then be placed as s_{N-1} and s_N. This way, less error-propagation effect and thus better error-rate performance can be expected.

- *E2:* Second, with the 2×2 nulling mechanism, 2×2 matrix inverse is performed each time for the cancelation of two transmitted symbols in the (sub)system. Actually, this is equivalent to the ZF operation, and noise amplification can be a problem if the involved 2×2 matrix is near singular. Fortunately, to perform matrix inverse with size 2×2, the simple Cramer's rule can be used, and thus it is quite simple to check the matrix condition before performing the actual nulling step. Here, we shuffle the channel rows to check the condition of different 2×2 matrices possible for the nulling, in which we only need to calculate the determinants of these matrices. The one with the best condition (with the largest determinant value) is used for the actual nulling. This simple checking step can largely reduce the noise amplification effect and improve the overall error rate at the same time.

D. Summary: Flow Diagram and Computational Complexity

For better understanding, the flow diagram of the complete ZMS detection is illustrated in Fig. 2, and the induced computational complexity is shown in Table I. In actual hardware implementation, it is understood that multiplication operation dominates the computational complexity. To not complicate the matter, in Table I, we describe the complexity through showing the required number of real multiplication (RM) in each step of our detection approach. Also note that our model is for complex-valued signals. With such a model, one complex multiplication requires four RM, and calculating the norm of a complex number requires two RM. In addition, in [8], it was derived that the total number of RM for the optimal ML detection of $N\times N$ MIMO systems is $4\{N^2M_{\mathbb{S}}\}+2\{NM_{\mathbb{S}}^N\}$, where $M_{\mathbb{S}}$ denotes the number of the constellation points, and the total number of RM for the conventional QRM-MLD of $N\times N$ MIMO systems is $4\left\{\left(\sum_{i=1}^{N-1} i + M_{\mathbb{S}}(N-1)\right)M + M_{\mathbb{S}}\right\} + 2\{MM_{\mathbb{S}}(N-1)+M_{\mathbb{S}}\}$ plus the RM for calculating QR decomposition, Q matrix, and implementing $\mathbf{Q}^H\mathbf{y}$.

IV. Simulations

In this section, comparison of the bit error rate (BER) performance and computational complexity among various common detection schemes and the proposed ZMS detection method is carried out. In the error-rate simulation, all coefficients in the channel matrix in (1) is independently generated according to Rayleigh fading. In addition, gray encoding is used for the mapping between the binary bits and symbols.

Figs. 3 and 4 are the BER simulation for 4×4 and 8×8 systems with QPSK signals, respectively. The BER curves for ZF, ordered SIC-ZF, conventional QRM-MLD, and full ML detection approaches are also included for completeness. From these two figures, we see that the BER of the proposed

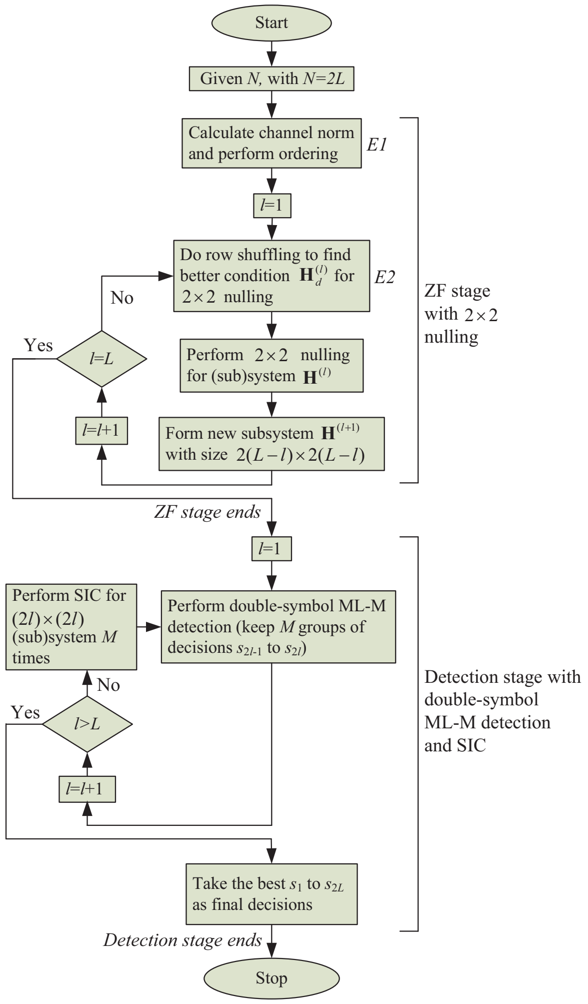

Fig. 2. Complete flow diagram of our proposed ZMS detection scheme for general $N\times N$ MIMO systems.

ZMS detection scheme is much smaller than that of the QRM-MLD with the same M, especially in the high SNR region. Interestingly, the resultant diversity gains of ZMS are much better than that of the QRM-MLD, and quite similar to that of the ML (or near ML) detection.

Next, we compare the required computational complexity. According to the derivation for the numbers of RM given in the last section, we present their computation load with different numbers of transmit/receive antennas. We observe that the required number of RM of the proposed ZMS detection scheme can be slightly smaller than that of the QRM-MLD.

V. Conclusions

In this paper, we have proposed a special formulation to combine the ZF, ML, SIC, and M algorithm for detecting $N\times N$ spatial multiplexing systems. We showed that the error-rate performance of this ZMQ detection can be much better than that of the conventional QRM-MLD with the same M, and the total number of RM required can be even lower than that of QRM-MLD in many cases.

TABLE I
COMPUTATIONAL COMPLEXITY REQUIRED FOR THE PROPOSED ZMS DETECTION IN TERMS OF REAL MULTIPLICATION, WHERE $M_{\mathbb{S}}$ DENOTES THE NUMBER OF THE CONSTELLATION POINTS ($M_{\mathbb{S}} = 4$) AND M IS THE SIZE OF M-ALGORITHM.

Stage	Operation	Complexity of each execution	Times of execution
ZF	*E1:* channel norm ordering	$2\{N^2\}$	Once
	E2: channel row shuffling	$4\left\{2\binom{2(L-l+1)}{2}\right\}+2\left\{\binom{2(L-l+1)}{2}\right\}$	$l = 1, 2, \cdots, L-1$
	2×2 nulling: form $\mathbf{W}^l$	$4\{8(L-l)+4\}$	$l = 1, 2, \cdots, L-1$
	2×2 nulling: obtain $\mathbf{y}^{l+1}$	$4\{4(L-l)\}$	$l = 1, 2, \cdots, L-1$
	2×2 nulling: form $\mathbf{H}^{l+1}$	$4\{8(L-l)^2\}$	$l = 1, 2, \cdots, L-1$
Detection	ML-M for 2×2 subsystem	$4\{4M_{\mathbb{S}}\}+2\{2M_{\mathbb{S}}^2\}$	Once
	ML-M for $2l \times 2$ subsystem	$4\{4lM_{\mathbb{S}}\}+2\{2lM_{\mathbb{S}}^2\} \times M$	$l = 2, 3, \cdots, L$
	SIC for $2l \times 2l$ (sub)system	$4\{(2l)(2(l-1))\} \times M$	$l = 2, 3, \cdots, L$

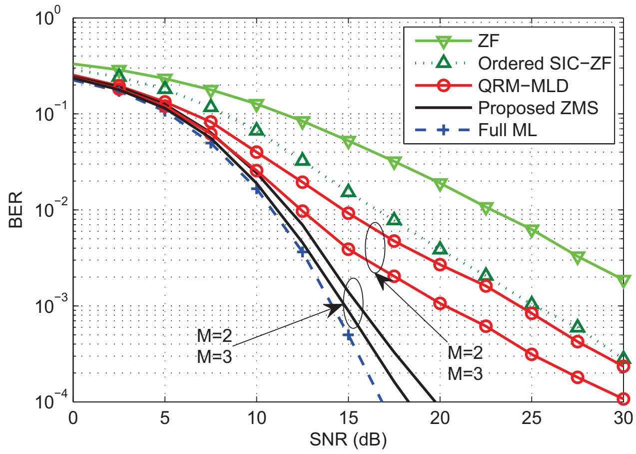

Fig. 3. BER of different detection schemes employing QPSK modulation with $N = 4$.

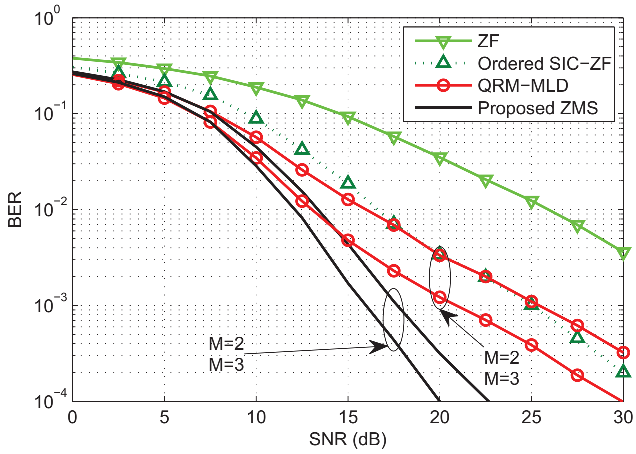

Fig. 4. BER of different detection schemes employing QPSK modulation with $N = 8$.

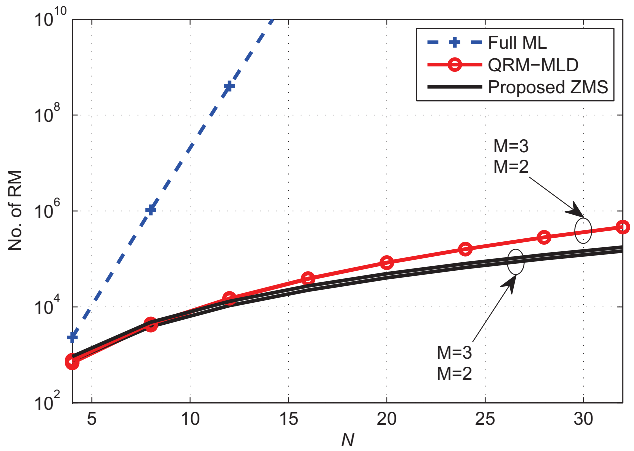

Fig. 5. Required number of RM of different detection schemes for one complete detection cycle.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.

[2] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[3] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[4] W.-K. Ma, P.-C. Ching, and Zhi Ding, "Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2862–2872, Oct. 2004.

[5] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near-maximum-likelihood decoding algorithm for MIMO systems based on semi-definite programming," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3869–3886, Nov. 2007.

[6] H. Kawai, K. Higuchi, N. Maeda, and M. Sawahashi, "Adaptive control of surviving symbol replica candidates in QRM-MLD for OFDM MIMO Multiplexing," *IEEE J. Select. Areas Commun.*, vol. 24, no. 6, pp. 1130–1140, Jun. 2006.

[7] Y. Lee and H.-W. Shieh, "Low-compleixty groupwise OSIC-ZF detection for $N \times N$ spatial multiplexing systems," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1930–1937, May 2011.

[8] J.-S. Kim, S.-H. Moon, and I. Lee, "A new reduced complexity ML detection scheme for MIMO systems," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1302–1310, Apr. 2010.