Interference Cancellation Based Detection for V-BLAST With Diversity Maximizing Channel Partition

Djelili Radji, *Student Member, IEEE*, and Harry Leib, *Senior Member, IEEE*

***Abstract*—Multiple-input multiple-output (MIMO) systems achieve very high bandwidth efficiencies through spatial multiplexing. However, the complexity of optimal detection in such systems motivates the need for more practical alternatives. Recently, a suboptimal lower complexity detection scheme called "Generalized Parallel Interference Cancellation" (GPIC), with close to optimal performance, was introduced. The reported performance of GPIC, however, was assessed by computer simulations only. In this paper, we show that with its original design, GPIC does not always provide close to optimal performance. Based on a diversity analysis of GPIC like techniques, we propose two new improved algorithms, referred to as Sel-MMSE and Sel-MMSE-OSIC, and derive sufficient conditions for achieving optimal performance asymptotically. We also provide a complexity analysis of these two schemes, and show that for large constellation sizes it is lower than the original GPIC. While still more complex than the fixed complexity sphere decoder by a factor in the range of 2-3 (for most configurations), our algorithms are also applicable to undetermined MIMO systems. Simulations results confirm that the new schemes provide maximal diversity gains. Furthermore, Sel-MMSE-OSIC provides a significant gain over Sel-MMSE, making its performance nearly indistinguishable from optimal for all signal-to-noise ratio (SNR) levels.**

***Index Terms*—Interference cancellation detection, minimum mean-squared error (MMSE) detection, multiple-input multiple-output (MIMO) detection, selection diversity.**

I. Introduction

MULTIPLE Antennas are essential in modern wireless communications [1], and the Vertical Bell Laboratories Layered Space-Time (V-BLAST) architecture [2], [3] launched an important technology for broadband wireless access. Its chief merit is the very high bandwidth efficiency provided through spatial multiplexing (SM) [3], [4]. However the optimal receiver for V-BLAST, implementing maximum likelihood (ML) detection, has a complexity that increases exponentially with the number of transmit antennas [5]. To alleviate this problem, suboptimal approaches based on successive interference cancellation (SIC), using the zero-forcing (ZF) or the minimum mean-squared error (MMSE) criteria with optimized ordering, are commonly employed [2], [3], [6]. Analytical and numerical results have shown that because of the diversity loss incurred, such schemes perform rather poorly compared to ML detection [7]–[9]. Therefore, other alternatives capable of bridging this performance gap with reasonable complexity are of significant interest for practical applications.

Sphere decoding (SD) [10], [11] is a low-complexity technique that exploits the lattice structure of a V-BLAST system model. Although it can achieve ML performance, for large systems, its complexity is still exponential in the number of transmit antennas [12]. Furthermore, SD has a variable complexity which depends on the channel realization and the operating signal-to-noise-ratio (SNR), thus making its integration into practical communication systems problematic. The fixed-complexity sphere decoder (FSD) of [13]–[15] relieves the problems associated with variations in complexity, and achieves quasi-ML performance. In addition, the FSD, unlike the SD, has a parallelizable structure that can be exploited to provide higher throughputs [13].

Several other alternatives have been proposed to improve the performance of the conventional V-BLAST receivers. Among them, is the class of list-based detectors, which are analogous to the Chase decoding algorithm in the sense that they first generate a list of candidate decision vectors, and then the best candidate in the list is chosen as the final decision [16]–[19]. In [16], the approach used to create the candidates list is to first choose an appropriate data substream, then consider all possible transmitted symbol hypotheses for this substream, and for each hypothesis use a low-complexity detector to estimate the remaining substreams. A performance within one dB of ML detection for a 5×5 system with QPSK modulation was reported [16]. The scheme from [17] differs from the one of [16] in that all substreams are considered in the first step and the number of hypotheses considered for each of them can be less than the constellation size. The overall size of the candidates list is therefore generally larger than that of [16]. The scheme from [17] can perform closely to ML detection for a 4×4 system but is poorer for larger system size. The scheme of [18] is similar to [16] except that it uses a different criterion to select the appropriate substream. Its performance is close to ML detection for a 4×4 system with QPSK modulation. The technique of [19] differs from [16] in the sense that it allows the number of hypotheses considered to be less than the constellation size, thereby providing a tradeoff between complexity and performance. In fact, as pointed out in [19], the scheme of [16] is a special case of [19] where the number of hypotheses equals the constellation

Manuscript received September 01, 2008. Current version published January 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gerald Matz.

The authors are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada (e-mail: djelili.radji@mail.mcgill.ca; harry.leib@mcgill.ca).

Digital Object Identifier 10.1109/JSTSP.2009.2035860

size. New criteria are proposed for the selection of the first substream, and an efficient implementation is suggested, yielding further reductions in complexity. It is shown in [19] that for 16QAM, while performing close to ML detection, this system exhibits a degradation that increases with system size.

Recently, a new technique called "Generalized Parallel Interference Cancellation" (GPIC) was introduced in [20]. It is a generalization of [16] and [18] where several data substreams are considered in the first step. In addition, a postcancellation step is included to increase the size of the candidates list and thereby improve the overall performance. Compared to [16]–[19], it potentially achieves the best performance. Since [16] is a special case of [19], and as we shall see below, GPIC can outperform the latter, it can also outperform [16]. Furthermore, for a 6×6 system with QPSK, at a bit error rate (BER) of 10^{-6}, GPIC's performance is nearly indistinguishable from ML while the scheme from [17] suffers a loss of about 1.4 dB. The scheme of [18] can be viewed as a special case of GPIC without postcancellation, where only one substream is used in the first step and SIC detectors are used for precancellation. For a 6×6 system with 16QAM at a BER of 10^{-3} the scheme of [19] exhibits a loss of about 0.6 dB relative to ML detection, while GPIC performs very closely to the latter. The channel partition procedure that GPIC employs plays a significant role in the overall performance of this algorithm. However, no analytical results were provided in [20] to substantiate the good performance that was assessed by computer simulations only. Despite the improved performance of GPIC [20] with respect to the schemes from [16]–[19], we found that in certain cases it incurs a diversity loss at high SNR as compared to ML detection.

In this paper, we provide a theoretical performance analysis of GPIC like algorithms that leads to improved schemes, providing better performance with a lower complexity. We propose the use of a diversity maximizing selection procedure for channel partition, and derive sufficient conditions for achieving optimal performance asymptotically. Our results show that the improved algorithms can achieve a performance indistinguishable from ML detection even with more transmit than receive antennas. Furthermore, we show that for large constellation sizes, the complexity of the improved algorithms is lower than that of the original GPIC.

The rest of this paper is organized as follows. Section II presents the system model and reviews the main features of the GPIC algorithm [20]. A performance analysis of GPIC like algorithms is provided in Section III, that also introduces the new and improved algorithms. Section IV provides a complexity analysis of the new algorithms. Computer simulation results, including the effects of channel estimation error and channel spatial correlation, are presented in Section V. Section VI concludes the paper.

II. System Model and Review of the GPIC Algorithm

Consider a communication system employing N_T transmit and N_R receive antennas over a Rayleigh flat fading channel that is time invariant over one use. Assuming perfect synchronization, the corresponding discrete-time system model is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ has independent identically distributed (i.i.d) circularly symmetric complex Gaussian (CSCG) zero mean entries of unit variance, with $\mathbf{h}_m$, the mth column of $\mathbf{H}$, representing the channel from the mth transmit to all receive antennas. Furthermore, $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is the additive noise vector whose components are i.i.d CSCG with zero mean and variance σ_n^2. The components of the transmitted signal vector $\mathbf{x} = [x_1 \cdots x_{N_T}]^T$ where $(.)^T$ stands for transposition, are equally probable and mutually independent, drawn from a symbol constellation $\mathcal{A}$ with zero mean and average symbol energy σ_s^2. Hence, $\mathcal{E}[\mathbf{x}\mathbf{x}^H] = \sigma_s^2 \mathbf{I}_{N_T}$, where $\mathcal{E}[.]$ stands for the expectation operation, $(.)^H$ is the conjugate transpose, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. The SNR is defined here as SNR $= \sigma_s^2/\sigma_n^2$.

With $\mathbf{H}$ known at the receiver, the lowest average error probability is achieved by the maximum-likelihood (ML) detector using a minimum Euclidean distance (MED) decision rule [5]

$$\hat{\mathbf{x}}^{\mathrm{ML}} = \arg \min_{\mathbf{d} \in \mathcal{A}^{N_T}} \|\mathbf{y} - \mathbf{H}\mathbf{d}\|^2 \tag{2}$$

where $\|.\|$ denotes the Euclidean norm. Consider an alternative representation of the ML decision rule. Let the matrix $\mathbf{H}_1 \in \mathbb{C}^{N_R \times N}$ be composed of a subset of N $(1 \le N \le N_T)$ columns of $\mathbf{H}$, with the remaining columns forming $\mathbf{H}_2 \in \mathbb{C}^{N_R \times (N_T - N)}$. Clearly, there are a total of $N_U = \binom{N_T}{N}$ such possible subsets. Furthermore, let $\mathbf{x}_1$ denote the vector whose elements are symbols transmitted from the N antennas corresponding to $\mathbf{H}_1$ and $\mathbf{x}_2$ similarly defined for $\mathbf{H}_2$. Hence, $\mathbf{H}\mathbf{x} = \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2$. Let $M = |\mathcal{A}|$ be the cardinality of the set $\mathcal{A}$, that is the size of the symbol constellation, and denote the elements of the sets $\mathcal{A}^N$ and $\mathcal{A}^{N_T - N}$, respectively, as $\{\tilde{\mathbf{x}}_1^1, \tilde{\mathbf{x}}_1^2, \ldots, \tilde{\mathbf{x}}_1^K\}$ and $\{\tilde{\mathbf{x}}_2^1, \tilde{\mathbf{x}}_2^2, \ldots, \tilde{\mathbf{x}}_2^Q\}$. Hence, $K = |\mathcal{A}^N| = M^N$ and $Q = |\mathcal{A}^{N_T - N}| = M^{N_T - N}$. The ML decision rule of (2) can now be rewritten as

$$\hat{\mathbf{x}}_1^{\mathrm{ML}} = \tilde{\mathbf{x}}_1^{k^*}, \quad \hat{\mathbf{x}}_2^{\mathrm{ML}} = \tilde{\mathbf{x}}_2^{q^*} \tag{3}$$

where

$$(k^*, q^*) = \arg \min_{k \in \{1,2,\ldots,K\}, q \in \{1,2,\ldots,Q\}} \left\|\mathbf{y} - \mathbf{H}_1\tilde{\mathbf{x}}_1^k - \mathbf{H}_2\tilde{\mathbf{x}}_2^q\right\|^2. \tag{4}$$

Defining

$$\mathbf{y}^k = \mathbf{y} - \mathbf{H}_1\tilde{\mathbf{x}}_1^k = \mathbf{H}_2\mathbf{x}_2 + \mathbf{H}_1\left(\mathbf{x}_1 - \tilde{\mathbf{x}}_1^k\right) + \mathbf{n} \tag{5}$$

and

$$\mathbf{d}_2^k = \arg \min_{\mathbf{d}_2 \in \left\{\tilde{\mathbf{x}}_2^1, \tilde{\mathbf{x}}_2^2, \ldots, \tilde{\mathbf{x}}_2^Q\right\}} \|\mathbf{y}^k - \mathbf{H}_2\mathbf{d}_2\|^2 \qquad k = 1, 2, \ldots, K \tag{6}$$

we have the following equivalent representation of the ML decision rule

$$\hat{\mathbf{x}}_1^{\mathrm{ML}} = \tilde{\mathbf{x}}_1^{k^*}, \qquad \hat{\mathbf{x}}_2^{\mathrm{ML}} = \tilde{\mathbf{x}}_2^{q^*} = \mathbf{d}_2^{k^*} \tag{7}$$

where

$$k^* = \arg \min_{k \in \{1,2,\ldots,K\}} \left\| \mathbf{y}^k - \mathbf{H}_2 \mathbf{d}_2^k \right\|^2 \tag{8}$$

with $\mathbf{y}^k$ and $\mathbf{d}_2^k$ given by (5) and (6). The complexity of this receiver is exponential in N_T [5].

The lower complexity GPIC algorithm of [20] relies on three steps: channel partition, precancellation, and postcancellation.

1) *Channel partition*: The GPIC algorithm uses the following procedure for channel partition to obtain $\mathbf{H}_1$ [20]:
   - Compute $\mathbf{H}^\dagger$, the Moore-Penrose pseudoinverse of $\mathbf{H}$ (see. e.g. [21], [22], $\mathbf{H}^\dagger = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$ if $\mathbf{H}$ has full column rank, and $\mathbf{H}^\dagger = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$ if $\mathbf{H}$ has full row rank).
   - Find the indexes (denoted by $i_1, \ldots, i_N$) of the N rows of $\mathbf{H}^\dagger$ with largest Euclidean norms.
   - Select the columns $\mathbf{h}_{i_1}, \mathbf{h}_{i_2}, \ldots, \mathbf{h}_{i_N}$ of $\mathbf{H}$ to form $\mathbf{H}_1$ and the other columns for $\mathbf{H}_2$.

   Clearly, any partitioning scheme is applicable, resulting in different performances. It is thus desirable to derive a selection procedure that improves performance. This is one of the main contributions of our work, which is addressed in the next section.
2) *Precancellation step*: Instead of using a MED rule to obtain the candidates $\mathbf{d}_2^k$ of (6), the precancellation step of the GPIC algorithm employs linear detection (LD). Assuming that $\mathbf{H}_2$ has full column rank, first an estimate of $\mathbf{d}_2^k$ $(k = 1, \ldots, K)$ is obtained

$$\bar{\mathbf{d}}_2^k = \mathbf{W}\mathbf{y}^k \tag{9}$$

where

$$\mathbf{W} = \mathbf{H}_2^\dagger = \left(\mathbf{H}_2^H\mathbf{H}_2\right)^{-1}\mathbf{H}_2^H \tag{10}$$

for the zero-forcing (ZF) criterion, and

$$\mathbf{W} = \left(\mathbf{H}_2^H\mathbf{H}_2 + \mathrm{SNR}^{-1}\mathbf{I}_{N_T-N}\right)^{-1}\mathbf{H}_2^H \tag{11}$$

for the minimum mean-squared error (MMSE) criterion. Then each component of $\bar{\mathbf{d}}_2^k$ is independently mapped to the closest point in the symbol constellation yielding $\mathbf{d}_2^k$. Clearly, a substantial reduction in complexity is achieved compared to using the MED rule. However, the candidates obtained in this manner may differ from those of (6) and consequently a certain amount of performance loss may be incurred as compared to the ML algorithm. Our work shows that such a loss can be significantly reduced if better detection techniques are used for precancellation.
3) *Postcancellation step*: This stage aims at improving the overall performance of the algorithm by allowing for detection of erroneous symbols after generating the candidates $\mathbf{d}_2^k$ in the precancellation step. This technique is described below [20]:
   - Choose an integer E such that with high probability $\mathbf{d}_2^k$ contains no more than E erroneous symbols.
   - For each of the $P = \binom{N_T-N}{E}$ possible combinations of E erroneous symbols and $N_T - N - E$ correct symbols, generate an updated value $\mathbf{d}_2^{k,p}$, $p \in \{1, 2, \ldots, P\}$ for $\mathbf{d}_2^k$, by first canceling the interference caused by the correct symbols and then detecting the erroneous symbols.

This process is repeated for $k = 1, 2, \ldots, K$. Hence, after postcancellation there are a total of KP candidates for the final decision about the transmitted vector instead of K with the precancellation only approach. Our work shows that with improved channel partition and precancellation, close to optimal performance can be achieved without postcancellation.

III. Performance Analysis

In this section, we consider the performance of GPIC like algorithms in terms of average probability of detection error $P_e(\mathrm{SNR})$, defined as the average probability that not all detected symbols are correct. Specifically, our aim is to quantify the diversity gain that it provides, where we recall that a given scheme is said to provide a diversity gain d if [4]

$$\lim_{\mathrm{SNR}\to\infty} \frac{\log P_e(\mathrm{SNR})}{\log \mathrm{SNR}} = -d. \tag{12}$$

The relevance of our diversity analysis approach comes from the fact that the performance loss incurred by suboptimal schemes such as linear ZF or MMSE algorithms and their OSIC counterparts, as compared to the ML algorithm, is due in part to a decrease in diversity gain. For a MIMO system that transmits independent data substreams on each antenna, thereby providing maximal multiplexing gain [4], the diversity gain provided by the ML receiver is N_R [5], [23], whereas the suboptimal schemes cited above can only provide an $N_R - N_T + 1$ diversity gain [5], [7]–[9]. Hence, a diversity analysis can indeed capture an essential part of the performance advantage provided by a specific scheme as compared to another. To evaluate (12), we use tight upper and lower bounds to $P_e(\mathrm{SNR})$. We ignore the postcancellation stage in this analysis, and will investigate its effects on performance through computer simulations in Section V. To simplify notations, we refer to the average probability of error as P_e, where its dependence on SNR is not explicitly written unless necessary.

Denote by $\mathcal{S}(\mathbf{y}, \mathbf{H})$ the set of all candidate pairs $\{(\tilde{\mathbf{x}}_1^k, \mathbf{d}_2^k), k = 1, 2, \ldots, K\}$ generated by the GPIC algorithm for a given received vector $\mathbf{y}$ and a given channel matrix $\mathbf{H}$. Then the output of the GPIC algorithm can be expressed as

$$\left(\hat{\mathbf{x}}_1^{\mathrm{GPIC}}, \hat{\mathbf{x}}_2^{\mathrm{GPIC}}\right) = \arg \min_{(\tilde{\mathbf{x}}_1, \mathbf{d}_2) \in \mathcal{S}(\mathbf{y}, \mathbf{H})} \|\mathbf{y} - \mathbf{H}_1\tilde{\mathbf{x}}_1 - \mathbf{H}_2\mathbf{d}_2\|^2. \tag{13}$$

Comparison of (13) with (3) shows that the only difference between the ML and the GPIC algorithms is that the latter searches

for the optimal pair $(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2)$ over a subset of the set used by the former. Hence,

$$\left(\hat{\mathbf{x}}_1^{\mathrm{GPIC}}, \hat{\mathbf{x}}_2^{\mathrm{GPIC}}\right) = \left(\hat{\mathbf{x}}_1^{\mathrm{ML}}, \hat{\mathbf{x}}_2^{\mathrm{ML}}\right) \Leftrightarrow \left(\hat{\mathbf{x}}_1^{\mathrm{ML}}, \hat{\mathbf{x}}_2^{\mathrm{ML}}\right) \in \mathcal{S}(\mathbf{y}, \mathbf{H}). \tag{14}$$

A. Upper Bound on P_e

We first expand P_e as

$$P_e = \mathcal{E}_{\mathbf{H},\mathbf{x}} \left[\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x})\right] \tag{15}$$

where $\Pr(A)$ is the probability that the event A occurs. We now further have

$$\begin{aligned}&\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) \\ &\quad = \Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) \\ &\qquad + \Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}|\mathbf{H}, \mathbf{x}).\end{aligned} \tag{16}$$

The first term in the right side of (16) can be upper bounded as

$$\begin{aligned}&\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) \\ &\qquad \leq \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}).\end{aligned} \tag{17}$$

To simplify the second term in the right side of (16), we first make the expansion

$$\begin{aligned}&\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}) \\ &\quad = \Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}, (\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{S}(\mathbf{y}, \mathbf{H})\right) \\ &\qquad \times \Pr\left((\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{S}(\mathbf{y}, \mathbf{H})|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) \\ &\qquad + \Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}, (\mathbf{x}_1, \mathbf{x}_2) \notin \mathcal{S}(\mathbf{y}, \mathbf{H})\right) \\ &\qquad \times \Pr\left((\mathbf{x}_1, \mathbf{x}_2) \notin \mathcal{S}(\mathbf{y}, \mathbf{H})|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right).\end{aligned} \tag{18}$$

Then we observe that

$$\Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}, (\mathbf{x}_1, \mathbf{x}_2) \notin \mathcal{S}(\mathbf{y}, \mathbf{H})\right) = 1 \tag{19}$$

since here the transmitted pair $(\mathbf{x}_1, \mathbf{x}_2)$ is not an element of $\mathcal{S}(\mathbf{y}, \mathbf{H})$. Furthermore, from (14)

$$\begin{aligned}&\Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}, (\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{S}(\mathbf{y}, \mathbf{H})\right) \\ &\quad = \Pr\left(\hat{\mathbf{x}} \neq \hat{\mathbf{x}}^{\mathrm{ML}}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}, \left(\hat{\mathbf{x}}_1^{\mathrm{ML}}, \hat{\mathbf{x}}_2^{\mathrm{ML}}\right) \in \mathcal{S}(\mathbf{y}, \mathbf{H})\right) \\ &\quad = 0.\end{aligned} \tag{20}$$

Substituting (19) and (20) into (18), we have

$$\begin{aligned}&\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}) \\ &\quad = \Pr\left((\mathbf{x}_1, \mathbf{x}_2) \notin \mathcal{S}(\mathbf{y}, \mathbf{H})|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) &(21)\\ &\quad = \Pr\left((\mathbf{x}_1, \mathbf{x}_2) \neq \left(\mathbf{x}_1, \mathbf{d}_2^{k_1}\right)|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) &(22)\\ &\quad = \Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) &(23)\end{aligned}$$

where $\mathbf{x}_1 = \tilde{\mathbf{x}}_1^{k_1}$ for some $k_1 \in \{1, 2, \ldots, K\}$. Furthermore,

$$\begin{aligned}&\Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}\right) \\ &\quad = \Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}|\mathbf{H}, \mathbf{x}) \\ &\qquad + \Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}\right) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) \\ &\quad \geq \Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}\right) \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}|\mathbf{H}, \mathbf{x}).\end{aligned} \tag{24}$$

Using (23) in (24), we have

$$\begin{aligned}\Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}\right) &\geq \Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}, \hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}) \\ &\quad \times \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} = \mathbf{x}|\mathbf{H}, \mathbf{x}).\end{aligned} \tag{25}$$

Recall that $\mathbf{d}_2^{k_1}$ is the output of the suboptimal algorithm used to generate the candidates for $\mathbf{x}_2$ when supplied with the received vector $\mathbf{y}^{k_1}$ given by

$$\mathbf{y}^{k_1} = \mathbf{y} - \mathbf{H}_1 \tilde{\mathbf{x}}_1^{k_1} = \mathbf{H}_2 \mathbf{x}_2 + \mathbf{H}_1 \left(\mathbf{x}_1 - \tilde{\mathbf{x}}_1^{k_1}\right) + \mathbf{n} = \mathbf{H}_2 \mathbf{x}_2 + \mathbf{n}. \tag{26}$$

Thus, we have

$$\Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H}, \mathbf{x}\right) = \Pr(\hat{\mathbf{x}}_2 \neq \mathbf{x}_2|\mathbf{H}_2, \mathbf{x}_2). \tag{27}$$

The use of (17), (25) and (27) in (16) yields

$$\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) \leq \Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x}) + \Pr(\hat{\mathbf{x}}_2 \neq \mathbf{x}_2|\mathbf{H}_2, \mathbf{x}_2). \tag{28}$$

Averaging both sides of (28) with respect to $\mathbf{H}$ and $\mathbf{x}$ yields the desired upper bound

$$P_e \leq P_e^{\mathrm{ML}} + P_{e2} \tag{29}$$

where

$$\begin{aligned}P_e^{\mathrm{ML}} &= \mathcal{E}_{\mathbf{H},\mathbf{x}} \left[\Pr(\hat{\mathbf{x}}^{\mathrm{ML}} \neq \mathbf{x}|\mathbf{H}, \mathbf{x})\right] \quad \text{and} \\ P_{e2} &= \mathcal{E}_{\mathbf{H}_2,\mathbf{x}_2} \left[\Pr(\hat{\mathbf{x}}_2 \neq \mathbf{x}_2|\mathbf{H}_2, \mathbf{x}_2)\right].\end{aligned} \tag{30}$$

Since the reduced channel $\mathbf{H}_2$ corresponds to a smaller number of transmit antennas than the full channel $\mathbf{H}$, we are not constrained by $P_e^{\mathrm{ML}} < P_{e2}$. We will illustrate this point by simulation results in Section V.

B. Lower Bound on P_e

Since the optimum receiver for our system implements the ML decision rule, the average probability of error associated with this receiver is a lower bound to the one achieved by any other detection scheme. Thus we have

$$P_e \geq P_e^{\mathrm{ML}}. \tag{31}$$

However, a less trivial lower bound can be obtained by rewriting (16) as

$$\begin{aligned}\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H},\mathbf{x}) = {} & \Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H},\mathbf{x},(\mathbf{x}_1,\mathbf{x}_2) \in \mathcal{S}(\mathbf{y},\mathbf{H})\right) \\ & \times \Pr\left((\mathbf{x}_1,\mathbf{x}_2) \in \mathcal{S}(\mathbf{y},\mathbf{H})|\mathbf{H},\mathbf{x}\right) \\ & + \Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H},\mathbf{x},(\mathbf{x}_1,\mathbf{x}_2) \notin \mathcal{S}(\mathbf{y},\mathbf{H})\right) \\ & \times \Pr\left((\mathbf{x}_1,\mathbf{x}_2) \notin \mathcal{S}(\mathbf{y},\mathbf{H})|\mathbf{H},\mathbf{x}\right). \quad (32)\end{aligned}$$

Clearly, we have

$$\Pr\left(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H},\mathbf{x},(\mathbf{x}_1,\mathbf{x}_2) \notin \mathcal{S}(\mathbf{y},\mathbf{H})\right) = 1 \quad (33)$$

since $(\mathbf{x}_1,\mathbf{x}_2)$ is not an element of $\mathcal{S}(\mathbf{y},\mathbf{H})$. Also, following similar steps as those used in the derivation of the upper bound (29), we have

$$\begin{aligned}&\Pr\left((\mathbf{x}_1,\mathbf{x}_2) \notin \mathcal{S}(\mathbf{y},\mathbf{H})|\mathbf{H},\mathbf{x}\right) \\ &\quad = \Pr\left((\mathbf{x}_1,\mathbf{x}_2) \neq \left(\mathbf{x}_1,\mathbf{d}_2^{k_1}\right)|\mathbf{H},\mathbf{x}\right) \quad (34) \\ &\quad = \Pr\left(\mathbf{d}_2^{k_1} \neq \mathbf{x}_2|\mathbf{H},\mathbf{x}\right) \quad (35) \\ &\quad = \Pr(\hat{\mathbf{x}}_2 \neq \mathbf{x}_2|\mathbf{H}_2,\mathbf{x}_2). \quad (36)\end{aligned}$$

Hence, from (32) with (33) and (36), we obtain

$$\Pr(\hat{\mathbf{x}} \neq \mathbf{x}|\mathbf{H},\mathbf{x}) \geq \Pr\left(\hat{\mathbf{x}}_2 \neq \mathbf{x}_2|\mathbf{H}_2,\mathbf{x}_2\right). \quad (37)$$

Then averaging both sides of (37) with respect to $\mathbf{H}$ and $\mathbf{x}$, we obtain

$$P_e \geq P_{e2}. \quad (38)$$

Combining (38) and (31), we get the desired lower bound

$$P_e \geq \max\left\{P_{e2}, P_e^{\mathrm{ML}}\right\}. \quad (39)$$

C. Channel Independent Selection

The combination of (29) and (30) gives

$$\max\left\{P_{e2}, P_e^{\mathrm{ML}}\right\} \leq P_e \leq P_e^{\mathrm{ML}} + P_{e2}. \quad (40)$$

Thus, the overall performance of the GPIC algorithm is determined by the relative values of P_{e2} and P_e^{ML}. However, for a given N, different channel partitioning schemes will result in different performances in terms of P_{e2}. In particular, we will see that the GPIC method using any channel independent partitioning scheme yields a suboptimal diversity gain in P_{e2} leading to poor overall performance for small values of N. However, the improved partitioning scheme that we introduce in the next subsection yields an optimal diversity gain in P_{e2}, making possible to achieve maximal overall diversity gain with relatively small N.

Recall that the ML receiver achieves a diversity of order N_R, and hence at high SNR

$$P_e^{\mathrm{ML}}(\mathrm{SNR}) \propto \mathrm{SNR}^{-N_R}. \quad (41)$$

As for P_{e2}, if the $L = N_T - N$ columns of the channel matrix $\mathbf{H}$ that form $\mathbf{H}_2$ are chosen at random without using any channel information, then from (26) and (30) P_{e2} is the average error probability of a given detection scheme for a MIMO flat Rayleigh fading channel with L transmit and N_R receive antennas. With the linear ZF/MMSE algorithms and their OSIC counterparts, the diversity gain is [5], [7]–[9]

$$d^{\mathrm{indep}} = N_R - L + 1 = N_R - N_T + N + 1 \quad (42)$$

and at high SNR, we have

$$P_{e2} = P_{e2}^{\mathrm{LD/OSIC}} \propto \mathrm{SNR}^{-N_R+N_T-N-1}. \quad (43)$$

Clearly, if $N < N_T - 1$, then $N + 1 - N_T < 0$ and

$$\lim_{\mathrm{SNR}\to\infty} \frac{P_e^{\mathrm{ML}}(\mathrm{SNR})}{P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR})} = 0. \quad (44)$$

Hence, at high SNR, both sides of (40) are dominated by the term $P_{e2}^{\mathrm{LD/OSIC}}$, and

$$P_e(\mathrm{SNR}) \approx P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR}) \propto \mathrm{SNR}^{-(N_R-N_T+N+1)}. \quad (45)$$

Thus, channel independent partitioning increases the diversity gain by N as compared to the conventional V-BLAST receivers with no partitioning. However, from (45), achieving the same diversity order as the ML receiver in this case requires $N = N_T - 1$. Clearly with such a value of N, no reduction in complexity as compared to the ML algorithm can be achieved. We now show that using the diversity maximizing selection procedure of [24], the GPIC algorithm can provide optimal diversity with $N < N_T - 1$.

D. Diversity Maximizing Selection

Let the set of all possible matrices $\mathbf{H}_2$ resulting from the partition of the channel matrix $\mathbf{H}$ be denoted as $\{\mathbf{H}_2^{(1)}, \mathbf{H}_2^{(2)}, \ldots, \mathbf{H}_2^{(N_U)}\}$. Then, from [24], we have that for the linear ZF and MMSE algorithms, the selection of columns from $\mathbf{H}$ to form $\mathbf{H}_2^{(p)}$ such that diversity is maximized should be

$$\mathbf{H}_2^{\mathrm{opt}} = \mathbf{H}_2^{(p)} \quad (46)$$

$$p = \arg \min_{1\leq j\leq N_U} (Z_j) \quad (47)$$

where

$$Z_j = \begin{cases} \max_{k=1,\ldots,L}\left\{\left(\mathbf{H}_2^{(j)H}\mathbf{H}_2^{(j)}\right)^{-1}_{kk}\right\}, & \text{for ZF} \\ \max_{k=1,\ldots,L}\left\{\left(\mathbf{H}_2^{(j)H}\mathbf{H}_2^{(j)} + \mathrm{SNR}^{-1}\mathbf{I}_L\right)^{-1}_{kk}\right\}, & \text{for MMSE} \end{cases} \quad (48)$$

with $(.)_{ij}$ standing for the (i,j)th entry of a matrix and $L = N_T - N$. The maximum achievable diversity gain with the LD or the OSIC algorithms can be bounded as [24]

$$d_L \leq \tilde{d} \leq d_U \quad (49)$$

where

$$d_L = (N_T - L + 1)(N_R - L + 1) = (N+1)(N_R - N_T + N + 1) \quad (50)$$

$$d_U = (N_T - L + 1)(N_R - 1) = (N+1)(N_R - 1) \quad (51)$$

for $N_T \geq 2$, $N_R \geq 2$, and $1 \leq L = N_T - N \leq \min\{N_T, N_R\}$, i.e, $\max\{0, N_T - N_R\} \leq N \leq N_T - 1$. Since for the GPIC algorithm we require N to be a strictly positive integer, then the range of allowable values for N with optimal selection is

$$\max\{1, N_T - N_R\} \leq N \leq N_T - 1. \quad (52)$$

We note here that contrary to the conventional approaches, (52) does not require $N_R \geq N_T$. Hence, this algorithm can handle the case $N_T > N_R$ provided that $N \geq N_T - N_R$. Reverting back to the bounds in (49), we point out that the authors of [24] make the conjecture

$$\tilde{d} = d_L = (N+1)(N_R - N_T + N + 1). \quad (53)$$

This conjecture is a generalization of a previous one made in [25], that was proven in [26] for the linear ZF/MMSE, the ZF-SIC with fixed/optimal ordering, and the MMSE-SIC with fixed ordering algorithms. However, to the best of our knowledge, the proof for the MMSE-SIC with optimal ordering is still lacking.

In [9] and [27], it was shown that LD and SIC algorithms based on the MMSE criterion can achieve significant performance gains over their ZF based counterparts. Therefore, we focus here on the MMSE criterion, and we now consider a scheme that employs the diversity maximizing selection procedure of Section III-D, with MMSE-OSIC rather than linear MMSE for detection in the precancellation step. Given a selected channel $\mathbf{H}_2 \in \mathbb{C}^{N_R \times L}$, the diversity gain achieved using linear MMSE detection is given by [24]

$$d^{\mathrm{LD}} = -\lim_{\mathrm{SNR}\to\infty} \frac{\log \Pr\left(R_{\min}^{\mathrm{LD}} \leq 1\right)}{\log \mathrm{SNR}} \quad (54)$$

where

$$R_{\min}^{\mathrm{LD}} = \min_{k=1\ldots L} \frac{\mathrm{SNR}}{\left(\mathbf{H}_2^H \mathbf{H}_2 + \mathrm{SNR}^{-1}\mathbf{I}_L\right)^{-1}_{kk}} - 1 \quad (55)$$

and the one achieved using MMSE-SIC with any ordering rule is given by [24]

$$d^{\mathrm{SIC}} = -\lim_{\mathrm{SNR}\to\infty} \frac{\log \Pr\left(R_{\min}^{\mathrm{SIC}} \leq 1\right)}{\log \mathrm{SNR}} \quad (56)$$

where

$$R_{\min}^{\mathrm{SIC}} = \min_{k=1\ldots L} \frac{\mathrm{SNR}}{\left(\mathbf{H}_{2,k}^H \mathbf{H}_{2,k} + \mathrm{SNR}^{-1}\mathbf{I}_k\right)^{-1}_{\pi(k)\pi(k)}} - 1 \quad (57)$$

with $\mathbf{H}_{2,k}$ being a submatrix obtained from $\mathbf{H}_2$ by removing the $k-1$ columns corresponding to the $k-1$ substreams previously detected, and $\pi(k)$ being the index of that column of $\mathbf{H}_{2,k}$ that corresponds to the substream chosen for detection. It can be shown that [27]

$$R_{\min}^{\mathrm{SIC}} \geq R_{\min}^{\mathrm{LD}} \quad \text{with probability one (w.p.1).} \quad (58)$$

Hence, from Appendix A, we get

$$\log \Pr\left(R_{\min}^{\mathrm{LD}} \leq 1\right) \geq \log \Pr\left(R_{\min}^{\mathrm{SIC}} \leq 1\right) \quad (59)$$

and thus we have

$$d^{\mathrm{LD}} = -\lim_{\mathrm{SNR}\to\infty} \frac{\log \Pr\left(R_{\min}^{\mathrm{LD}} \leq 1\right)}{\log \mathrm{SNR}} \leq -\lim_{\mathrm{SNR}\to\infty} \frac{\log \Pr\left(R_{\min}^{\mathrm{SIC}} \leq 1\right)}{\log \mathrm{SNR}} = d^{\mathrm{SIC}}. \quad (60)$$

Furthermore, from (53) and [26], for LD algorithms, this selection procedure provides optimum diversity, i.e., $d^{\mathrm{LD}} = \tilde{d} = d_L$. Hence, for SIC with any ordering rule, this selection procedure provides a diversity gain no less than d_L, the maximal diversity gain for LD. We thus propose the following improved GPIC algorithm.

1) For channel partitioning, use the MMSE-based diversity maximizing selection procedure described above to obtain the optimal matrix $\mathbf{H}_2$.
2) For each candidate $\tilde{\mathbf{x}}_1^k$ ($k \in \{1, \ldots, K\}$), use either linear MMSE detection or its OSIC counterpart to obtain the corresponding candidates $\mathbf{d}_2^k$
3) Choose the pair $(\tilde{\mathbf{x}}_1, \mathbf{d}_2)$ associated with the MED [cf. (13)] as the final decision.

We will refer to the algorithms using linear MMSE and MMSE-OSIC detection in step 2) as Sel-MMSE and Sel-MMSE-OSIC, respectively.

We now proceed with the analysis of the GPIC algorithm employing diversity optimal selection, and show how to choose N to guarantee asymptotically optimal performance. From (49), at high SNR we have

$$P_{e2} = P_{e2}^{\mathrm{LD/OSIC}} \propto \mathrm{SNR}^{-\tilde{d}}. \quad (61)$$

Using (40) and (41), we draw the following conclusions.

1) If $\tilde{d} > N_R$, and since from (41) and (61)

$$\frac{P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR})}{P_e^{\mathrm{ML}}(\mathrm{SNR})} \propto \mathrm{SNR}^{-(\tilde{d}-N_R)} \quad (62)$$

we have

$$\lim_{\mathrm{SNR}\to\infty} \frac{P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR})}{P_e^{\mathrm{ML}}(\mathrm{SNR})} = 0. \quad (63)$$

Using (63) with (40), we obtain

$$\lim_{\mathrm{SNR}\to\infty} \frac{P_e(\mathrm{SNR})}{P_e^{\mathrm{ML}}(\mathrm{SNR})} = 1. \quad (64)$$

Hence, GPIC achieves optimal performance asymptotically.

2) If $\tilde{d} = N_R$, then, $P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR}) \propto \mathrm{SNR}^{-N_R}$. Thus, from (40) and (41), at high SNR

$$P_e(\mathrm{SNR}) \propto \mathrm{SNR}^{-N_R} \tag{65}$$

showing that GPIC only achieves the same diversity gain as the ML algorithm.

3) If $\tilde{d} < N_R$, using (40) and (41) yields

$$\lim_{\mathrm{SNR}\to\infty} \frac{P_e(\mathrm{SNR})}{P_e^{\mathrm{LD/OSIC}}(\mathrm{SNR})} = 1 \tag{66}$$

and thus at high SNR, we have

$$P_e(\mathrm{SNR}) \propto \mathrm{SNR}^{-\tilde{d}} \tag{67}$$

showing that GPIC provides a diversity gain strictly lower than the ML's.

Define $f(N) = d_L - N_R$, and with (53) consider the following function:

$$f(s) = (s+1)(N_R - N_T + s + 1) - N_R, \quad s \in \mathbb{R}. \tag{68}$$

We seek the minimal integer $N_{\min}$, such that when $N \geq N_{\min}$ we have $f(N) > 0$. Since from (49) $\tilde{d} \geq d_L$, we have that $f(N) > 0 \Rightarrow \tilde{d} > N_R$. Therefore, if $N \geq N_{\min}$ then the GPIC algorithm achieves optimal performance asymptotically. Thus, by analyzing the function (68), we can derive sufficient conditions for the GPIC algorithm with diversity optimal channel partition to provide ML performance at high SNR. The two zeros of $f(s)$ are

$$s_1 = \sqrt{N_R + \frac{1}{4}(N_R - N_T)^2} - \frac{1}{2}(N_R - N_T) - 1 \tag{69}$$

$$s_2 = -\sqrt{N_R + \frac{1}{4}(N_R - N_T)^2} - \frac{1}{2}(N_R - N_T) - 1 \tag{70}$$

and $f(s)$ is negative if and only if $s_2 < s < s_1$. Now consider the following.

1) Assume $N_R \geq 1$. If $N_R \geq N_T$, then $s_2 = -(\sqrt{N_R + (1/4)(N_R - N_T)^2} + (1/2)(N_R - N_T) + 1) < 0$. Now if $N_R < N_T$, then $s_2 = (1/2)(N_T - N_R)(1 - \sqrt{1 + 4N_R(N_T - N_R)^{-2}}) - 1 < -1 < 0$.
2) Assume $N_R \geq 1$ and $N_T \geq 2$, then $(1/4)(N_R + N_T)^2 - (1/4)(N_R - N_T)^2 = N_R N_T > N_R$. Thus, $\sqrt{N_R + (1/4)(N_R - N_T)^2} < (1/2)(N_R + N_T)$. Therefore, $s_1 < (1/2)(N_R + N_T) - (1/2)(N_R - N_T) - 1 = N_T - 1$.
3) Assume $N_R \geq 1$ and $N_T \geq 2$.
   - If $N_R = N_T$, then $s_1 = \sqrt{N_R} - 1 > \sqrt{2} - 1 > 0$.
   - If $N_R < N_T$, then $s_1 = (1/2)(N_T - N_R)(1 + \sqrt{1 + 4N_R(N_T - N_R)^{-2}}) - 1 > (2/2)(N_T - N_R) - 1 = N_T - N_R - 1$.
   - If $N_R > N_T$, then since $N_T \geq 2 > 1$, we have $N_R > N_R - N_T + 1$. Thus, $N_R + (1/4)(N_R - N_T)^2 > (1/4)(N_R - N_T)^2 + N_R - N_T + 1 = ((1/2)(N_R - N_T) + 1)^2$. Hence, $s_1 = \sqrt{N_R + (1/4)(N_R - N_T)^2} - (1/2)(N_R - N_T) - 1 > 0$.

Combining these results, we conclude that if $N_T \geq 2$ and $N_R \geq 1$, then

$$s_2 < 0 \quad \text{and} \quad \max\{0, N_T - N_R - 1\} < s_1 < N_T - 1. \tag{71}$$

Since $s_2 < 0$ for all system configurations of interest and $N > 0$

$$f(N) > 0 \Leftrightarrow N > s_1 \Leftrightarrow N \geq \lfloor s_1 + 1 \rfloor = N_{\min} \tag{72}$$

where $\lfloor A \rfloor$ stands for rounding to the nearest integer less than or equal to A, and the last inequality follows from the fact that N must be an integer. Hence,

$$N_{\min} = \left\lfloor \sqrt{N_R + \frac{1}{4}(N_R - N_T)^2} - \frac{1}{2}(N_R - N_T) \right\rfloor. \tag{73}$$

Using (71), we obtain

$$\max\{1, N_T - N_R\} \leq N_{\min} \leq N_T - 1 \tag{74}$$

providing identical bounds as (52). Hence, the sufficient condition to guarantee that with optimum selection GPIC performs asymptotically optimum, is

$$N \geq N_{\min}. \tag{75}$$

From (73), when $N_R = N_T$, we have $N_{\min} = \lfloor \sqrt{N_R} \rfloor$. Thus, for systems with $N_R = N_T < 4$ and those with $4 \leq N_R = N_T < 9$, we have $N_{\min} = 1$ and $N_{\min} = 2$, respectively. Hence, the required $N_{\min}$ is relatively small for most practical systems and therefore the overall complexity of our algorithm can indeed be relatively low. Furthermore, from (72), we have $N_{\min} = 1$ if and only if $s_1 < 1$. It is straightforward to verify: $s_1 < 1 \Leftrightarrow N_R > 2N_T - 4 \Leftrightarrow N_R \geq 2N_T - 3$, where the last inequality follows because N_R is an integer. Thus, we have shown

$$N_{\min} = 1 \Leftrightarrow N_R \geq 2N_T - 3. \tag{76}$$

For any given N_T, we can always ensure $N_{\min} = 1$ by selecting N_R sufficiently large. Table I provides the values of $N_{\min}$ for various system configurations. From this table, we observe that for $N_T = N_R = 12$, $N_{\min} = 3$; thus, confirming the fact that $N_{\min}$ is a relatively small integer even for large systems. Furthermore, we also observe that for $N_T = 4$, $N_{\min} = 2$ for $N_R = 4$ and $N_{\min} = 1$ for $N_R = 6$. Hence, in this case, increasing N_R by 2 not only increases the optimal diversity gain by 2 but also reduces $N_{\min}$.

IV. Complexity Analysis and Comparison

In this section, we take a closer look at the computational complexity of the new algorithms and provide a comparison with the FSD algorithm of [13]–[15]. Similar to [20], as a measure of complexity, we count the number of complex additions/substractions and complex multiplications/divisions required to implement the algorithms. Here, we consider

TABLE I
VALUES OF $N_{\min}$ FOR VARIOUS SYSTEM CONFIGURATIONS

N_R \ N_T	4	6	8	10	12
2	2	4	6	8	10
4	2	3	4	6	8
6	1	2	3	5	6
8	1	2	2	4	5
10	1	1	2	3	4
12	1	1	2	2	3
14	1	1	1	2	2
16	1	1	1	2	2
18	1	1	1	1	2
20	1	1	1	1	2
22	1	1	1	1	1

GPIC with optimum selection and linear MMSE precancellation, referred to as Sel-MMSE, and the same algorithm with MMSE-SIC precancellation employing V-BLAST ordering (MMSE-OSIC), referred to as Sel-MMSE-OSIC. Postcancellation is not employed for neither Sel-MMSE nor Sel-MMSE-OSIC. We count the number of complex operations required for channel partition, generation of the MMSE weight vectors and the decision statistics used by the constellation demapper in precancellation, and the MED decision rule.

Recall that similar to the real case, the product of an $R_A \times C_A$ complex matrix by a $C_A \times C_B$ complex matrix requires $C_A R_A C_B$ complex multiplications and $(C_A - 1)R_A C_B$ complex additions [6], [28]. Also, regarding the Hermitian product $\mathbf{A}^H\mathbf{A}$ where $\mathbf{A}$ is $R_A \times C_A$, taking into account both the fact that only $(1/2)C_A(C_A + 1)$ entries need to be computed and also the fact that the diagonal entries of the resulting matrix are real, this operation requires $(1/2)C_A^2 R_A$ complex multiplications and $(1/2)C_A^2(R_A - 1)$ complex additions [6], [28].

Channel partition is common to both Sel-MMSE and Sel-MMSE-OSIC and it requires the computation of $N_U = \binom{N_T}{N}$ inverses of matrices of the form $(\mathbf{A}^H\mathbf{A}+\text{SNR}^{-1}\mathbf{I}_L)$ where $\mathbf{A}$ is an $N_R \times (N_T - N)$ complex matrix. From the results presented in [29], each such inversion requires $(1/2)((N_T-N)^3+(N_T-N)^2(N_R+1)+(N_T-N)(N_R-2)+2)$ complex multiplications and $(1/2)((N_T-N)^3+(N_T-N)^2(N_R-2)+N_R(N_T-N))$ complex additions.

A. Sel-MMSE

For Sel-MMSE, the next step is to generate the MMSE weight matrix according to (11), with $\mathbf{H}_2$ replaced by $\mathbf{H}_2^{\text{opt}}$. Now since the matrix $(\mathbf{H}_2^{\text{opt}H}\mathbf{H}_2^{\text{opt}} + \text{SNR}^{-1}\mathbf{I}_L)^{-1}$ has already been computed in the channel partition step, computing the MMSE weight matrix requires additional $N_R(N_T-N)^2$ complex multiplications and $N_R(N_T-N)(N_T-N-1)$ complex additions. Then obtaining each of the $K = M^N$ candidates for $\mathbf{x}_2$ requires first the computation of $\mathbf{y}^k$ as in (5). Clearly, this computation requires NN_R complex multiplications and NN_R complex additions. The decision statistics used for demapping are then obtained according to (9), which in turn needs $N_R(N_T - N)$ complex multiplications and $(N_R - 1)(N_T - N)$ complex additions. Hence, generating the K candidates requires KN_RN_T complex multiplications and $K(N_T(N_R - 1) + N)$ complex additions. Finally finding the MED solution (8) requires $K(N_T-N)N_R+(N_R/2)) = KN_R((N_T-N+(1/2)$ complex multiplications, and $K(N_T-N-1)N_R+N_R+((N_R-1)/2)= K(N_R(N_T - N) + (N_R - 1)/2)$ complex additions. Hence Sel-MMSE requires overall

$$\frac{N_U}{2}\big((N_T-N)^3+(N_T-N)^2(N_R+1)+(N_T-N)(N_R-2)+2\big) + N_R(N_T-N)^2+KN_R(2N_T-N+1/2) \quad (77)$$

complex multiplications, and

$$\frac{N_U}{2}\left((N_T-N)^3+(N_T-N)^2(N_R-2)+N_R(N_T-N)\right) + N_R(N_T-N)(N_T-N-1) + K\left((N_R-1)(N_T-N+1/2)+N_RN_T\right) \quad (78)$$

complex additions.

B. Sel-MMSE-OSIC

For Sel-MMSE-OSIC, at each detection stage ($i = 0, 1, \ldots N_T - N - 1$), the MMSE-OSIC algorithm first computes the inverse of a matrix of the form $(\mathbf{A}^H\mathbf{A} + \text{SNR}^{-1}\mathbf{I}_L)$ where $\mathbf{A}$ is an $N_R \times (N_T - N - i)$ matrix. The substream selected for detection corresponds then to the index of the smallest diagonal element of this matrix [6]. After that, the corresponding weight vector is obtained as $(\mathbf{Aq})^H$, where $\mathbf{q}$ is the corresponding column of $(\mathbf{A}^H\mathbf{A} + \text{SNR}^{-1}\mathbf{I}_L)$.

Now since $(\mathbf{H}_2^{\text{opt}H}\mathbf{H}_2^{\text{opt}} + \text{SNR}^{-1}\mathbf{I}_L)^{-1}$ has already been computed in the channel partition step, the generation of the MMSE weight vector for the first detection step requires $N_R(N_T - N)$ complex multiplications and $N_R(N_T - N - 1)$ complex additions. For each of the subsequent detection steps ($i = 1, \ldots N_T - N - 1$), using the results from [29], obtaining the weight vector requires $(1/2)(N_T - N - i)(N_T - N - i + 3) + N_R(N_T - N - i)$ complex multiplications and $(1/2)(N_T - N - i)(N_T - N - i + 1) + N_R(N_T - N - i - 1)$ complex additions. Thus, the weight vectors generation process requires overall $N_R(N_T - N) + \sum_{i=1}^{N_T-N-1}(1/2)(N_T - N - i)(N_T - N - i + 3) + N_R(N_T - N - i) = (1/6)(N_T - N)^3 + (1/2)(N_T-N)^2(N_R+1)+(1/6)(N_T-N)(3N_R-4)$ complex multiplications and $N_R(N_T-N-1)+\sum_{i=1}^{N_T-N-1}(1/2)(N_T-N-i)(N_T-N-i+1)+N_R(N_T-N-i-1) = (1/6)(N_T-N)^3 + (1/2)N_R(N_T - N)^2-(1/6)(N_T - N)(3N_R + 1)$ complex additions.

After that, for each of the K candidates for $\mathbf{x}_2$, $\mathbf{y}^k$ must be computed as in (5), which requires NN_R complex multiplications and NN_R complex additions. Then, at the ith ($i = 0, 1, \ldots N_T-N-1$) detection step, the decision statistic is obtained by an inner product of the corresponding weight vector and the modified received vector. This process requires N_R complex multiplications and $N_R - 1$ complex additions. The interference cancellation step used to obtain the modified received vector for each detection step requires N_R complex multiplications and N_R complex additions. Hence, the overall cost for all K candidates is $K(NN_R+(N_T-N)N_R+(N_T-N-1)N_R)= KN_R(2N_T - N - 1)$ complex multiplications and $K(NN_R + (N_T - N)(N_R - 1) + (N_T - N - 1)N_R)= K((N_R - 1)(N_T - N) + N_R(N_T - 1))$ complex additions.

Now due to the SIC approach used here, the metric $\mathbf{y}^k - \mathbf{H}_2\mathbf{d}_2^k$ required in (8) can be computed with N_R complex additions and multiplications. Therefore, obtaining the MED solution here adds overall $(3/2)KN_R$ complex multiplications and $K(3N_R-1)/2$ complex additions. Hence, Sel-MMSE-OSIC requires overall

$$\begin{aligned}&\frac{N_U}{2}\left((N_T-N)^3+(N_T-N)^2(N_R+1)\right.\\&\quad\left.+(N_T-N)(N_R-2)+2\right)+\frac{1}{6}(N_T-N)^3\\&\quad+\frac{1}{2}(N_T-N)^2(N_R+1)+\frac{1}{6}(N_T-N)(3N_R-4)\\&\quad+KN_R(2N_T-N+1/2)\end{aligned}\tag{79}$$

complex multiplications and

$$\begin{aligned}&\frac{N_U}{2}\left((N_T-N)^3+(N_T-N)^2(N_R-2)+N_R(N_T-N)\right)\\&\quad+\frac{1}{6}(N_T-N)^3+\frac{1}{2}N_R(N_T-N)^2\\&\quad-\frac{1}{6}(N_T-N)(3N_R+1)\\&\quad+K\left((N_R-1)(N_T-N+1/2)+N_RN_T\right)\end{aligned}\tag{80}$$

complex additions.

C. FSD of [13]

The FSD algorithm of [13] implements

$$\hat{\mathbf{x}}^{\mathrm{FSD}}=\arg\min_{\mathbf{s}\in\mathcal{D}}\|\mathbf{U}(\mathbf{s}-\hat{\mathbf{s}})\|^2\tag{81}$$

where $\mathcal{D}$ is a subset of $\mathcal{A}^{N_T}$, $\mathbf{U}$ is an $N_T\times N_T$ upper triangular matrix, obtained through the Cholesky decomposition of the matrix $(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})$, $\hat{\mathbf{s}}=(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^{-1}\tilde{\mathbf{H}}^H\mathbf{y}$ is the ZF estimate of the transmitted vector $\mathbf{x}$ and $\tilde{\mathbf{H}}$ is a column permuted version of $\mathbf{H}$ obtained from the FSD channel matrix ordering algorithm [13], [15]. The Euclidean distance in (81) is computed recursively starting from level $i=N_T$ down to level $i=1$ using

$$D_i=u_{ii}^2|s_i-z_i|^2+\sum_{j=i+1}^{N_T}u_{jj}^2|s_j-z_j|^2=d_i+D_{i+1}\tag{82}$$

where $D_{N_T+1}=0$, $D_1=\|\mathbf{U}(\mathbf{s}-\hat{\mathbf{s}})\|^2$ and

$$z_i=\hat{s}_i-\sum_{j=i+1}^{N_T}\frac{u_{ij}}{u_{ii}}(s_j-\hat{s}_j).\tag{83}$$

In the first p levels, i.e., levels $N_T, N_T-1,\ldots,N_T-p+1$, a full search is performed by expanding all M branches per node [15]. For the remaining levels, only one branch per node, i.e., the one closest to z_i, is considered.

The FSD channel matrix ordering has the same complexity as the V-BLAST ordering. Thus, using the results from [29], this process needs $(2/3)N_T^3+(1/2)N_T^2(N_R+2)+(1/6)N_T(3N_R-10)+1$ complex multiplications and $(2/3)N_T^3+(1/2)N_T^2(N_R-2)+(1/6)N_T(3N_R-4)$ complex additions. Obtaining $\mathbf{U}$ from the Cholesky decomposition [28] requires $(1/6)N_T^3-(1/12)N_T^2-(1/12)N_T$ complex multiplications and $(1/6)N_T^3-(1/4)N_T^2+(1/12)N_T$ complex additions. Since, $(\tilde{\mathbf{H}}^H\tilde{\mathbf{H}})^{-1}$ has already been computed during the ordering stage, obtaining $\hat{\mathbf{s}}$ requires $N_T(N_T+N_R)$ complex multiplications and $N_T(N_T+N_R-2)$ complex additions.

Since u_{ii} $(i=1,\ldots,N_T)$ is real and recalling that a real multiplication costs 1/6 complex multiplication and a real addition costs 1/2 complex addition, the cost of computing D_i $(i\neq N_T)$ of (82) is 5/6 complex multiplications and 3/2 complex additions. As for D_{N_T}, since $D_{N_T+1}=0$, its computation needs only 5/6 complex multiplications and 1 complex addition. Similarly, recalling that the product ab where a is complex and b is real costs 1/3 complex multiplication, computing z_i $(i\neq N_T)$ from (83) requires $(4/3)(N_T-i)$ complex multiplications and $2(N_T-i)$ complex additions, while $z_{N_T}=\hat{s}_{N_T}$ is obtained with no additional cost.

Now since for most practical systems, we have $p\leq 2$ [14], [15], we will only consider the cases $p=1$ and $p=2$ in the following. For the case $p=1$, a total of M branches are considered by the FSD algorithm at each level. Thus the total cost of computing all required z_i is $\sum_{i=1}^{N_T-1}(4/3)M(N_T-i)=(2/3)MN_T(N_T-1)$ complex multiplications and $\sum_{i=1}^{N_T-1}2M(N_T-i)=MN_T(N_T-1)$ complex additions. Similarly, the total cost of computing all required D_i is $(5/6)MN_T$ complex multiplications and $(1/2)M(3N_T-1)$ complex additions. Therefore, for the case $p=1$, the FSD algorithm requires overall

$$\begin{aligned}&\frac{5}{6}N_T^3+\frac{1}{12}N_T^2(6N_R+23)+\frac{1}{4}N_T(6N_R-7)\\&\qquad+1+\frac{1}{6}MN_T(4N_T+1)\end{aligned}\tag{84}$$

complex multiplications and

$$\begin{aligned}&\frac{5}{6}N_T^3+\frac{1}{4}N_T^2(2N_R-1)+\frac{1}{12}N_T(18N_R-31)\\&\qquad+M(N_T^2+(1/2)N_T-1/2)\end{aligned}\tag{85}$$

complex additions.

Proceeding similarly, we can show that for the case $p=2$, the FSD algorithm requires overall

$$\begin{aligned}&\frac{5}{6}N_T^3+\frac{1}{12}N_T^2(6N_R+23)+\frac{1}{4}N_T(6N_R-7)+1\\&\qquad+\frac{1}{6}M\left(5+M(N_T-1)(4N_T+5)\right)\end{aligned}\tag{86}$$

complex multiplications and

$$\begin{aligned}&\frac{5}{6}N_T^3+\frac{1}{4}N_T^2(2N_R-1)+\frac{1}{12}N_T(18N_R-31)\\&\qquad+\frac{1}{2}M\left(2+M(N_T-1)(2N_T+3)\right)\end{aligned}\tag{87}$$

complex additions.

D. Complexity Comparison

We now provide a complexity analysis of the ML algorithm so that it can be compared to that of the algorithms discussed above. Finding the ML solution according to (2) requires for each of the M^{N_T} candidates two computations: one of the form $\mathbf{r}-\mathbf{A}\mathbf{s}$ and another one of the form $\|\mathbf{r}\|^2=\mathbf{r}^H\mathbf{r}$ where $\mathbf{r}$ and $\mathbf{s}$ are, respectively, N_R and N_T dimensional complex vectors, and $\mathbf{A}$ is an $N_R\times N_T$ complex matrix. Clearly, the first computation requires N_RN_T complex multiplications and additions,

TABLE II
FLOPS COUNT OF VARIOUS ALGORITHMS RELATIVE TO ML

System	M	4	16	64
$N_T = 4$, $N_R = 4$	GPIC (1,0)	$6.24\text{x}10^{-2}$	$5.29\text{x}10^{-4}$	$6.53\text{x}10^{-6}$
	GPIC (1,1)	$9.73\text{x}10^{-2}$	$1.10\text{x}10^{-3}$	$1.46\text{x}10^{-5}$
	Sel-MMSE (1,0)	$6.51\text{x}10^{-2}$	$5.53\text{x}10^{-4}$	$6.82\text{x}10^{-6}$
	Sel-MMSE-OSIC (1,0)	$6.39\text{x}10^{-2}$	$5.48\text{x}10^{-4}$	$6.80\text{x}10^{-6}$
	FSD $p = 1$	$3.84\text{x}10^{-2}$	$2.82\text{x}10^{-4}$	$3.16\text{x}10^{-6}$
$N_T = 6$, $N_R = 6$	GPIC (2,0)	$9.42\text{x}10^{-3}$	$2.43\text{x}10^{-5}$	$9.20\text{x}10^{-8}$
	GPIC (2,1)	$1.72\text{x}10^{-2}$	$5.42\text{x}10^{-5}$	$2.09\text{x}10^{-7}$
	Sel-MMSE (2,0)	$1.54\text{x}10^{-2}$	$2.65\text{x}10^{-5}$	$9.54\text{x}10^{-8}$
	Sel-MMSE-OSIC (1,0)	$8.10\text{x}10^{-3}$	$3.23\text{x}10^{-6}$	$2.01\text{x}10^{-9}$
	Sel-MMSE-OSIC (2,0)	$1.53\text{x}10^{-2}$	$2.65\text{x}10^{-5}$	$9.54\text{x}10^{-8}$
	FSD $p = 2$	$5.20\text{x}10^{-3}$	$1.14\text{x}10^{-5}$	$4.23\text{x}10^{-8}$
$N_T = 8$, $N_R = 8$	GPIC (2,0)	$7.26\text{x}10^{-4}$	$1.03\text{x}10^{-7}$	$2.41\text{x}10^{-11}$
	GPIC (2,2)	$3.50\text{x}10^{-3}$	$7.60\text{x}10^{-7}$	$1.84\text{x}10^{-10}$
	Sel-MMSE (2,0)	$2.20\text{x}10^{-3}$	$1.28\text{x}10^{-7}$	$2.49\text{x}10^{-11}$
	Sel-MMSE-OSIC (1,0)	$8.87\text{x}10^{-4}$	$1.86\text{x}10^{-8}$	$5.90\text{x}10^{-13}$
	Sel-MMSE-OSIC (2,0)	$2.20\text{x}10^{-3}$	$1.28\text{x}10^{-7}$	$2.49\text{x}10^{-11}$
	FSD $p = 2$	$3.68\text{x}10^{-4}$	$4.60\text{x}10^{-8}$	$1.05\text{x}10^{-11}$
$N_T = 4$, $N_R = 2$	GPIC (2,1)	$1.95\text{x}10^{-1}$	$1.06\text{x}10^{-2}$	$6.54\text{x}10^{-4}$
	Sel-MMSE (2,0)	$1.16\text{x}10^{-1}$	$5.60\text{x}10^{-3}$	$3.41\text{x}10^{-4}$
	Sel-MMSE-OSIC (2,0)	$1.16\text{x}10^{-1}$	$5.60\text{x}10^{-3}$	$3.41\text{x}10^{-4}$
	FSD	Not suitable	Not suitable	Not suitable

while the second one requires $N_R/2$ complex multiplications and $(N_R - 1)/2$ complex additions. Hence, the ML algorithm requires overall $M^{N_T} N_R(N_T + 1/2)$ complex multiplications and $M^{N_T}(N_R N_T + (N_R - 1)/2)$ complex additions.

In Table II, we present the ratio of the number of real floating point operations (flops) required for GPIC, Sel-MMSE, Sel-MMSE-OSIC, and FSD to that required by the ML algorithm for various system configurations. The flops counts are obtained from the complex multiplications and additions counts using the following conversion rule: one complex addition requires two flops and one complex multiplication requires six flops [6]. In this table, SchemeI (i,j) refers to SchemeI with $N = i$ and $E = j$. Thus, the new algorithms are denoted as Sel-MMSE (i,0) and Sel-MMSE-OSIC (i,0) as they do not use postcancellation. The results show that Sel-MMSE and Sel-MMSE-OSIC have similar complexities and that they have lower complexity than both the ML algorithm and the original GPIC algorithm with postcancellation. Furthermore, the complexity gain increases as the constellation size M increases. Indeed, for $N_T = N_R = 8$ and 64QAM modulation, compared to ML, the suboptimal schemes GPIC (2,2), Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0) lower the complexity by a factor greater than 10^{10} and Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0) have a lowered complexity with respect to GPIC (2,2) by a factor of 7.

In most of the considered scenarios, the FSD algorithm lowers the complexity with respect to Sel-MMSE-OSIC (2,0) by a factor of 2 to 3. For the 8×8 system with QPSK modulation, its complexity reduction relative to Sel-MMSE-OSIC (2,0) is a factor of 6. This is due to the fact that for this system configuration, the complexity of Sel-MMSE-OSIC (2,0) is dominated by the channel partition stage which requires the computation of $N_U = 28$ inverses of matrices of size 6. However, the FSD algorithm does not possess the flexibility of Sel-MMSE-OSIC in the sense that it cannot be applied for undetermined systems, i.e., $N_T > N_R$. Indeed, as shown in Table II, while our new algorithms lower the complexity with respect to the ML algorithm by several orders of magnitude, the FSD algorithm on the other hand is not suitable for this configuration. Furthermore, for the 6×6 system with 64QAM modulation, Sel-MMSE-OSIC (1,0) lowers the complexity with respect to the FSD algorithm by a factor of 21 while, as we shall see in the next section, performing within 0.6 dB from ML at SER $= 10^{-4}$.

V. COMPUTER SIMULATIONS RESULTS

This section provides computer simulations results for the performance of the various detection techniques discussed in this paper. In all cases except Fig. 2(b), that considers the average probability of error P_e (15) and the average bit error rate (BER), we measure the average symbol error rate (SER) defined as $\text{SER} = (1/N_T) \sum_{i=1}^{N_T} \Pr(x_i \neq \hat{x}_i)$.

We consider uncoded systems with QPSK, 16QAM and 64QAM modulation. For each channel use, a new vector of N_T randomly generated symbols is transmitted, and the channel varies randomly and independently from one use to another. The elements of the channel matrix are generated as i.i.d complex Gaussian random variables with zero mean and unit variance, modeling Rayleigh fading. In addition, for each SNR point the performance is averaged over a minimum of a 100 000 channel realizations and a minimum of 500 vector errors are accumulated. The schemes considered are the ML algorithm, the GPIC algorithm of [20] including postcancellation, and the new Sel-MMSE and Sel-MMSE-OSIC algorithms without postcancellation. When $N_R \geq N_T$, the ML performance is simulated using the exact complex sphere decoder [11], [30], [31], otherwise it is simulated using (2). We do not reproduce the results for the conventional V-BLAST receivers without channel partitioning, since these are presented in [20] showing the significance of channel partition in such techniques.

A. P_{e2} and P_e^{ML} Diversity Comparison

We first illustrate the fact that $P_{e2}^{\text{LD/OSIC}}$ (30) can exhibit a diversity order greater or equal to that of P_e^{ML} (30). Regarding $P_{e2}^{\text{LD/OSIC}}$, for each channel realization, after optimum partitioning, we use $\mathbf{y}^{k_1}$ from (26) as an equivalent received vector and apply linear MMSE detection for Sel-MMSE (or its OSIC counterpart for Sel-MMSE-OSIC) to obtain $\hat{\mathbf{x}}_2$. If $\hat{\mathbf{x}}_2 \neq \mathbf{x}_2$, then an error is declared. For each SNR point, the performance is then averaged over a minimum of a 100 000 channel realizations and a minimum of 500 errors are accumulated. In Fig. 1, P_e^{ML}, P_{e2} Sel-MMSE and P_{e2} Sel-MMSE-OSIC are compared for two different system configurations: a 4×4 system with QPSK and a 6×6 system with 64QAM.

For the 4×4 system, from (53) we have $\tilde{d} = d_L = 4 = N_R$ for Sel-MMSE (1,0). Thus, in Fig. 1, both P_{e2} Sel-MMSE (1,0) and P_e^{ML} show the same diversity gain. Fig. 1 also shows that in this case P_{e2} Sel-MMSE-OSIC (1,0) provides the same diversity gain as P_e^{ML}. Now considering the 6×6 system, from (53) we have $\tilde{d} = d_L = 4 < N_R$ for Sel-MMSE (1,0), and $\tilde{d} = d_L = 9 > N_R$ for Sel-MMSE (2,0). Therefore, P_{e2} Sel-MMSE (1,0) has lower diversity gain than P_e^{ML}, while P_{e2} Sel-MMSE (2,0) has higher diversity gain than P_e^{ML} as confirmed by the results presented in Fig. 1. Similarly, P_{e2} Sel-MMSE-OSIC (1,0) has lower diversity gain than P_e^{ML} while P_{e2} Sel-MMSE-OSIC (2,0) has higher diversity gain than P_e^{ML}.

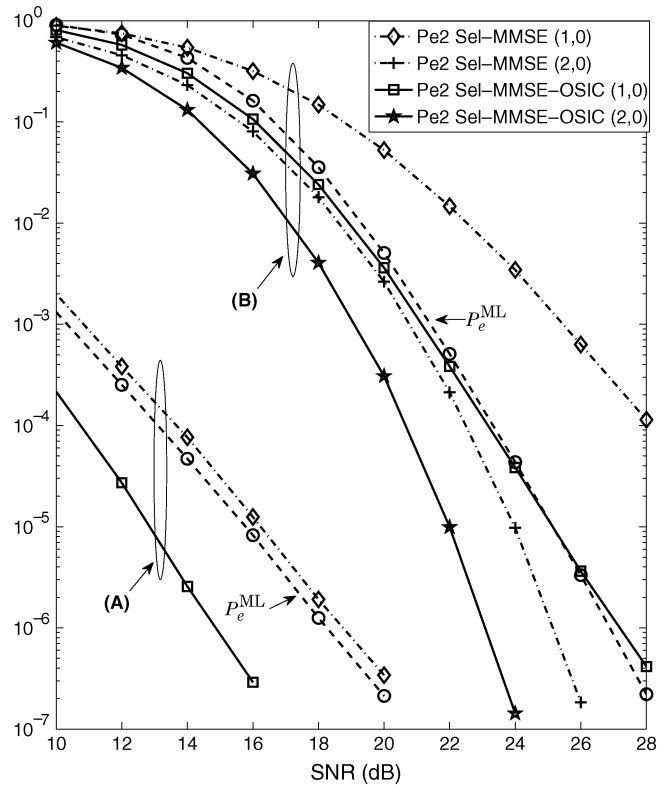
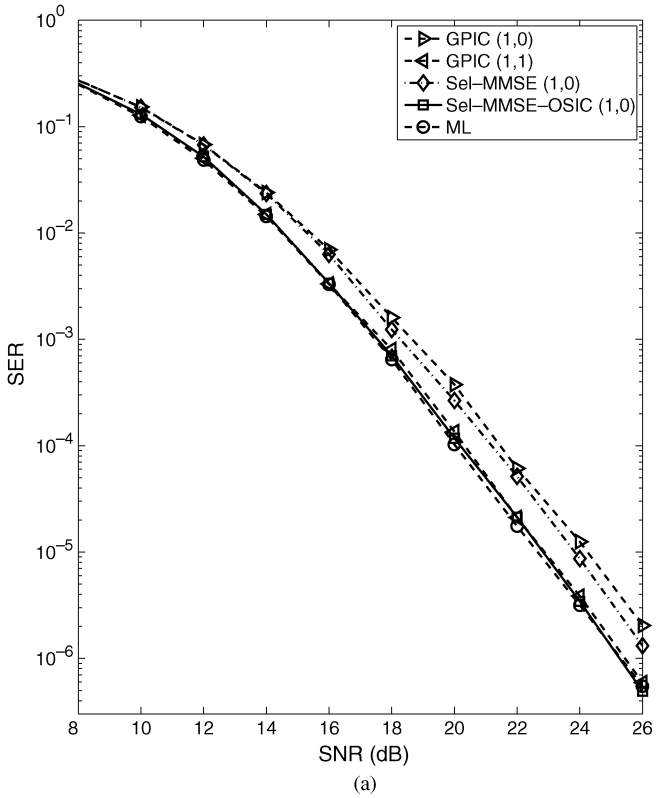


Fig. 1. P_e^{ML}, P_{e2} Sel-MMSE and P_{e2} Sel-MMSE-OSIC comparison. (A) $N_T = N_R = 4$ QPSK. (B) $N_T = N_R = 6$ 64QAM.

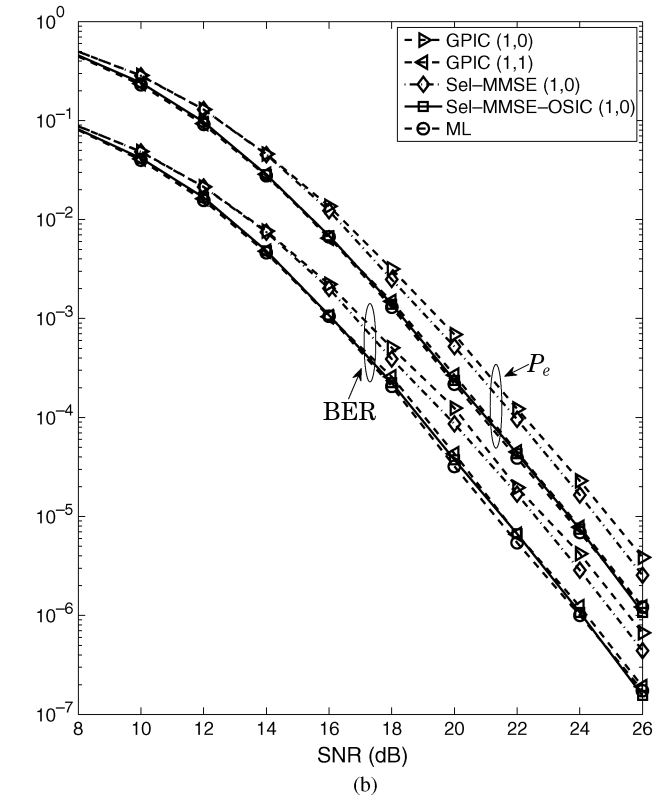
B. Error Rates Performance

In Fig. 2, a 16QAM system with $N_R = N_T = 4$ and Gray mapping [32] is considered. Since in this case, with $N = 1$, $d_L = N_R = 4$, both Sel-MMSE (1,0) and Sel-MMSE-OSIC (1,0) provide optimal diversity gain. While Sel-MMSE-OSIC (1,0) is nearly indistinguishable from ML, Sel-MMSE (1,0) is about 1 dB worse. Thus, the use of SIC can provide significant SNR gains over the linear detection approach. GPIC (1,1) also performs very closely to ML and has similar performance as Sel-MMSE-OSIC (1,0) but requires about twice more computations than the latter. Fig. 2(b) illustrates the fact that the P_e, BER, and SER plots all bear the same trend, and henceforth in the followings we choose to present only SER results due to space limits.

The results for $N_R = N_T = 6$ with QPSK and 64QAM modulation are shown in Figs. 3. Here, $N_{\min} = 2$ and therefore both Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0) achieve optimal performance, asymptotically. For both constellations, the performance of Sel-MMSE-OSIC (2,0) is nearly indistinguishable from ML for the entire range of SNR considered. In addition, from Fig. 3(b), we observe that despite the fact that they both provide the same diversity gain, Sel-MMSE-OSIC (1,0) achieves an SNR gain of more than 4 dB over Sel-MMSE (1,0) at SER $= 10^{-4}$. Furthermore, from the same figure, GPIC (2,0) performs nearly 2 dB worse than Sel-MMSE-OSIC (1,0) at SER $= 10^{-5}$. With QPSK, the performance of GPIC (2,1) is indistinguishable from ML for SER $\geq 10^{-3}$, and it is 0.4 dB worse than ML at SER $= 10^{-7}$. For 64QAM, GPIC (2,1) is



(a)



(b)

Fig. 2. Performance of the algorithms, $N_T = N_R = 4$ 16QAM. (a) SER. (b) P_e and BER.

0.7 dB worse than ML at SER $= 10^{-6}$. These observations illustrate the approximate optimal performance of the GPIC (2,1) claimed in [20], and also that Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0) can indeed provide indistinguishable performance

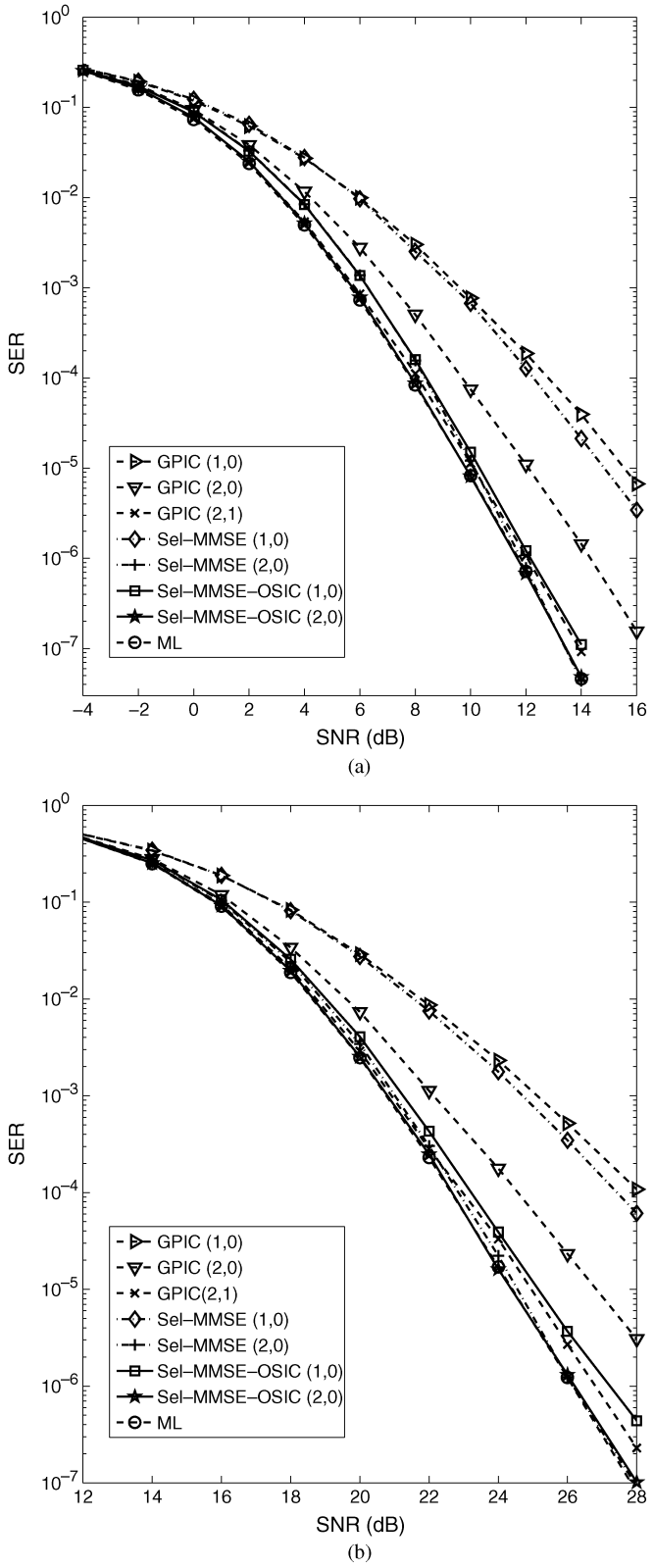


Fig. 3. Symbol error rate performance of the algorithms, $N_T = N_R = 6$. (a) QPSK. (b) 64QAM.

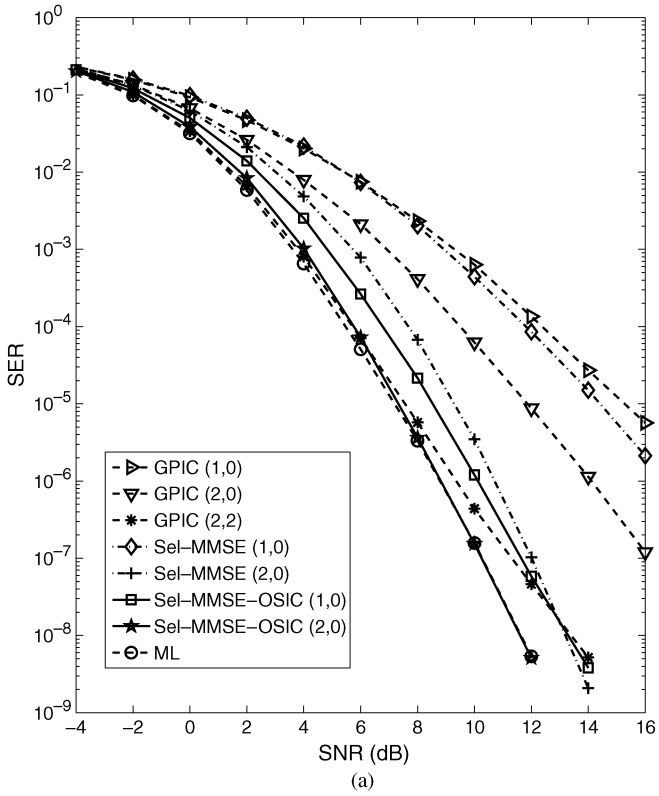
when compared to optimal. Furthermore, from Table II, with 64QAM, these two improved algorithms require only about half the complexity associated with "GPIC (2,1)" and compared to ML, they lower the complexity by a factor of 10^7. We further observe that for this case, Sel-MMSE-OSIC (1,0) lowers the complexity with respect to the FSD algorithm by a factor of 21 while performing within 0.6 dB from ML at SER $= 10^{-4}$.

Fig. 4 present the results for $N_R = N_T = 8$ with QPSK and 64QAM modulation. Again, here $N_{\min} = 2$ and thus theoretically asymptotic optimal performance is guaranteed for both Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0). Now, while the latter performs very close to ML in all cases, the former fails to match ML's performance for the range of SNR considered. Indeed, despite a convergence trend, from Fig. 4(a), the performance of Sel-MMSE (2,0) is still about 1.5 dB away from MLs for SER $= 10^{-8}$. Such a behavior can be explained by first noting that in this case, we have $d_L = 9$. Then, although the condition $d_L > N_R$ required for asymptotic convergence still holds here, we do have $d_L - N_R = 1$. Therefore, the required SNR level for $P_{e2}^{\mathrm{LD/OSIC}}(\mathrm{SNR}) \ll P_e^{\mathrm{ML}}(\mathrm{SNR})$, and thereby yielding convergence, can become very large. In addition, since the achievable diversity gain here is 8, the error rate level for which the two schemes exhibit similar performances can thus be extremely low. Hence, the strong SNR gain advantage provided by Sel-MMSE-OSIC is essential here to achieve optimal performance at error rates of practical interest. Indeed, with 64QAM, Sel-MMSE-OSIC (2,0) performs within 0.1 dB from ML for the entire SNR range. Both GPIC (2,0) and GPIC (2,2) fail to provide optimal performance asymptotically. In fact, contrary to the claim made in [20], the performance of GPIC (2,2) actually diverges from MLs. Indeed, from Fig. 4(b), while GPIC (2,2) is about 1 dB worse than ML at SER $= 10^{-4}$, at SER $= 10^{-6}$, it is nearly 3 dB worse than ML. Furthermore, from Table II, for 64QAM modulation GPIC (2,2) is about 7 times more complex than both Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0).

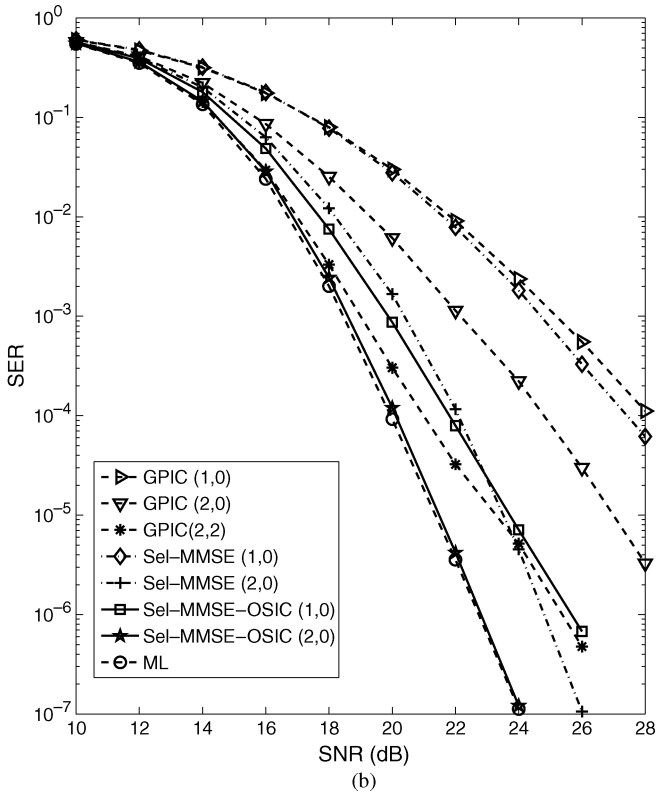
Fig. 5 considers a MIMO system with four transmit and two receive antennas. In this case, $N_{\min} = 2$, and thus both Sel-MMSE (2,0) and Sel-MMSE-OSIC (2,0) achieve optimal performance asymptotically. In fact, their respective performances are nearly indistinguishable from ML. On the other hand, both GPIC (2,0) and GPIC (2,1) fail to provide optimal diversity gain. For 16QAM, GPIC (2,1) is more than 3 dB worse than ML at SER $= 10^{-3}$, and provides a diversity order of only 1 as compared to the optimal diversity order of 2. From Table II, Sel-MMSE (2,0) requires only about 60% of the complexity of GPIC (2,1) and about 12% of the complexity of ML for the QPSK case, while for the 16QAM case, it has a lower complexity by a factor of nearly 2 with respect to GPIC (2,1), and by a factor of 178 with respect to ML. The very poor performance of GPIC in this case should be expected since when $N_R < N_T$, the rows of the pseudoinverse of the channel matrix can no longer be used to null the interstream interference, and therefore the rationale that lead to the channel partition procedure of [20] fails. The FSD algorithm on the other hand is not applicable here.

C. Impact of Spatial Correlation and Channel Estimation Errors

Next, we consider the impact of channel spatial correlation on the relative performance of these algorithms. We consider a 4×4 system with 16QAM modulation, and employ the channel

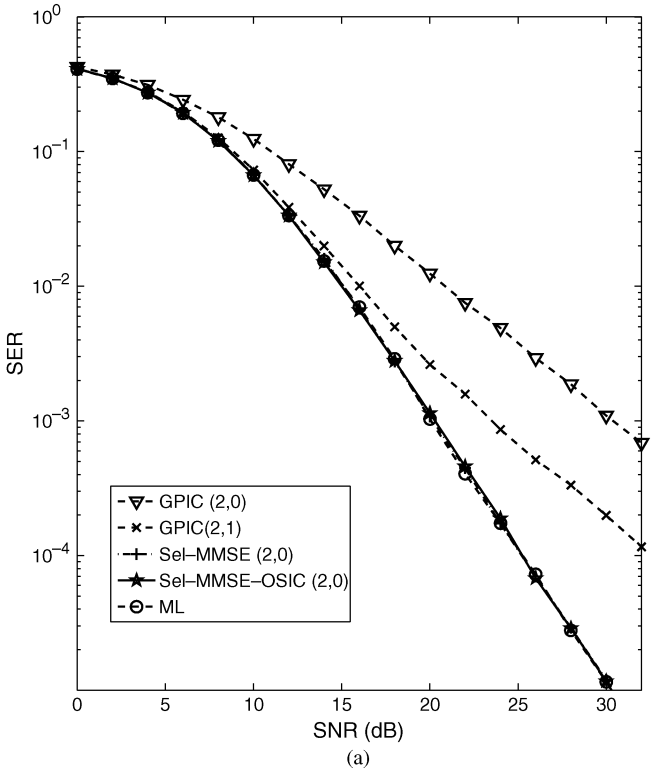


(a)

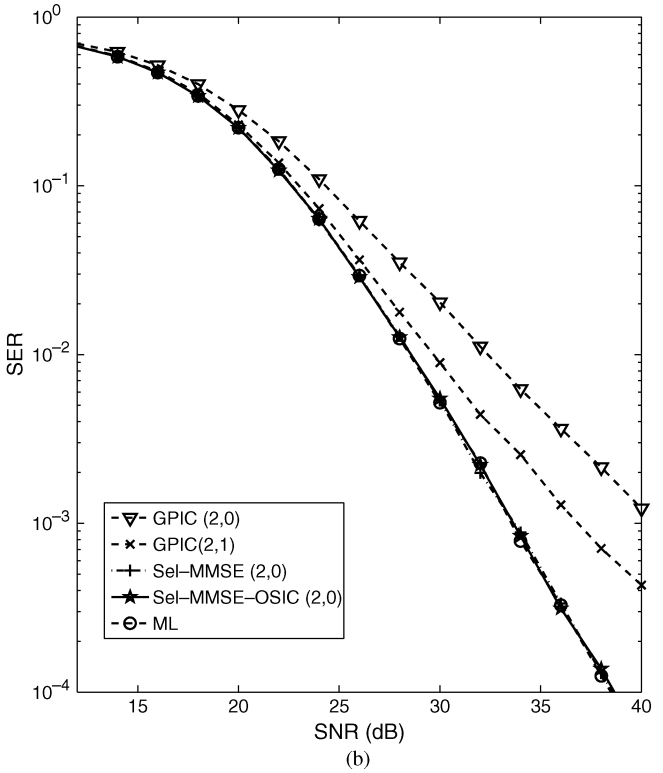


(b)

Fig. 4. Symbol error rate performance of the algorithms, $N_T = N_R = 8$. (a) QPSK. (b) 64QAM.



(a)



(b)

Fig. 5. Symbol error rate performance of the algorithms, $N_T = 4$, $N_R = 2$. (a) QPSK. (b) 16QAM.

correlation model of [33]. Assuming same space correlation at the transmitter and receiver, the channel matrix is written as $\tilde{\mathbf{H}} = \mathbf{R}^{1/2}\mathbf{H}\mathbf{R}^{1/2}$. Here, $\mathbf{H}$ is the channel matrix with i.i.d components as described in Section II, and $\mathbf{R}$ is the 4×4 symmetric Toeplitz spatial correlation matrix with first row given by $[1\ r\ r^4\ r^9]$, where r is a parameter that indicates the correlation level [33]. We consider both moderate and high correlation cases corresponding, respectively, to $r = 0.5$ and $r = 0.8$. The

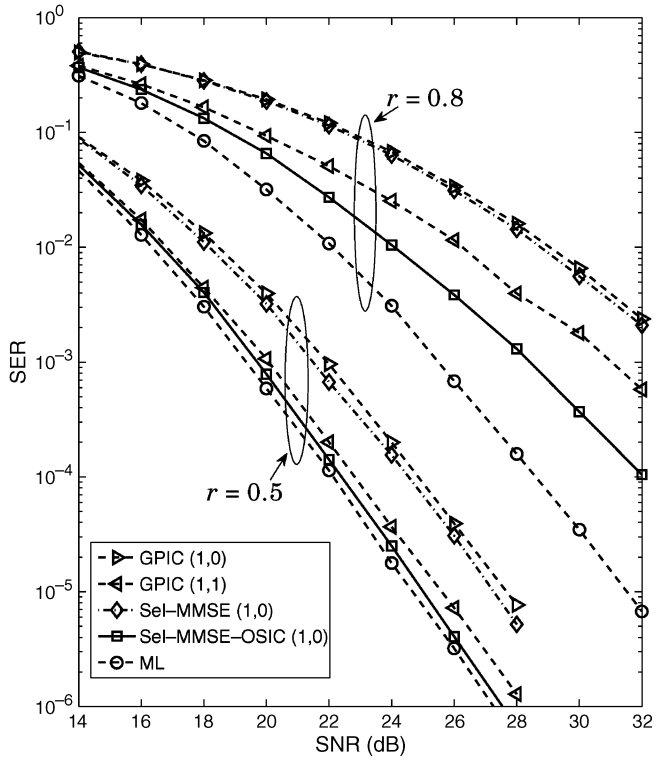


Fig. 6. Effect of spatial correlation, $N_T = N_R = 4$ 16QAM.

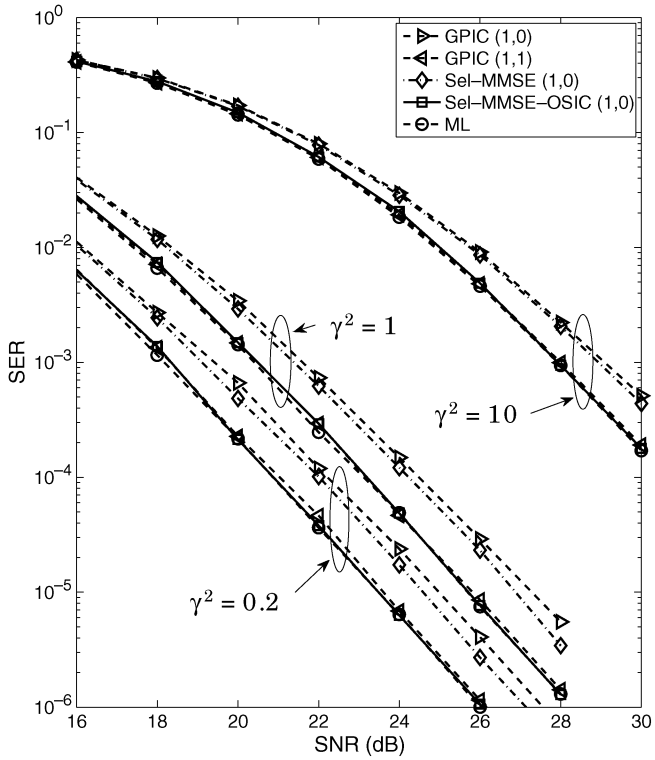


Fig. 7. Effect of channel estimation error, $N_T = N_R = 4$ 16QAM.

results are presented in Fig. 6. It is seen that the performances of the various algorithms degrade as the spatial correlation increases. In addition, the performance gap between the ML algorithm and the suboptimal approaches increases with correlation. However, Sel-MMSE-OSIC (1,0) is clearly the most robust among the suboptimal schemes; for $r = 0.5$, its performance is about 0.3 dB worse than ML and 0.6 dB better than GPIC (1,1) at a SER $= 10^{-5}$, while for $r = 0.8$, it performs about 3 dB worse than ML and about 2.5 dB better than GPIC (1,1) at a SER $= 10^{-3}$. Note that if the additive noise vector in our original system model (1) is space correlated with positive definite spatial correlation matrix $\mathbf{C}$, then by whitening we can obtain an equivalent spatially white noise system, with receive spatial correlation matrix given by $\mathbf{C}^{-1/2}$. Thus, the space correlated additive noise case can be regarded as a special case of channel spatial correlation.

Finally in Fig. 7, we investigate the effect of channel estimation error on the performance of these algorithms. We adopt the same model as [34] with the estimated channel modeled as $\hat{\mathbf{H}} = \mathbf{H} + \mathbf{H}_\epsilon$, where $\mathbf{H}$ is the true channel matrix and the $N_R \times N_T$ error matrix $\mathbf{H}_\epsilon$ is independent of $\mathbf{H}$ and its elements are i.i.d zero-mean complex Gaussian variables with variance $\gamma^2\sigma_n^2$, with γ^2 being a constant indicative of the error level. We consider a 4×4 system with 16QAM modulation and different values of γ^2. From the results presented in Fig. 7, we observe that similarly to the spatial correlation case, channel estimation error results in a certain amount of degradation in performance for all the detection schemes. However, contrary to the spatial correlation case, the estimation error has an almost equal impact on all schemes.

VI. Conclusion

In this paper, we have shown by analysis that with an improved version of the GPIC algorithm, ML error performance is achievable at high SNR with $N < N_T - 1$ and without postcancellation. Furthermore, we have derived the conditions to obtain such a performance: the diversity maximizing channel partition algorithm of [24] should be used instead of the suboptimal approach of [20], and the parameter N (i.e., the number of columns in the matrix $\mathbf{H}_1$), should be no less than $N_{\min}$, given by (73). For most practical cases, $N_{\min} \leq 2$. The use of SIC for precancellation yields large SNR gains over LD, and thereby mitigates the problem of slow convergence of SER to ML SER when increasing SNR.

We also showed that the improved algorithms have lower complexity than the original GPIC of [20] and that the complexity gain increases with the system constellation size M. The FSD algorithm on the other hand achieves a complexity reduction of 2 to 3 relative to the new algorithms for most of the considered scenarios. However, contrary to Sel-MMSE and Sel-MMSE-OSIC which are guaranteed to provide near optimal performance even for undetermined systems, i.e., $N_T > N_R$, the FSD algorithm is not applicable for such cases.

Simulation results reveal that the original GPIC fails to provide optimal performance asymptotically. The new improved algorithms, however, without postcancellation provide optimal diversity, with Sel-MMSE-OSIC performing indistinguishable from ML over a large SNR range. The improved algorithms perform well even when the number of transmit antennas exceeds the number of receive antennas. In addition, our results show that while channel estimation error has an almost equal

impact on all schemes, spatial correlation causes a more severe degradation in performance for the suboptimal schemes than for ML. The Sel-MMSE-OSIC scheme, however, can tolerate higher levels of spatial correlation than the other suboptimal schemes. Therefore, these new algorithms provide detection techniques for V-BLAST, that from performance and complexity point of views could be very attractive for practical applications.

APPENDIX A

Let X and Y be two continuous random variables with joint probability density function $p_{XY}(x,y)$. Then $Y \geq X$ with probability one means that $\int\int_{y \geq x} p_{XY}(x,y)dxdy = 1$. Hence, $p_{XY}(x,y) = 0$ in the region where $y < x$. Consider the following generalization of the latter: $p_{XY}(x,y) \leq p_{XY}(y,x)$ in the region where $y < x$. Define the three non-overlapping regions

$$\begin{aligned}\mathcal{R}_1 &= \{(x,y) : x \leq \alpha, y \leq \alpha\} \\ \mathcal{R}_2 &= \{(x,y) : x \leq \alpha, y > \alpha\} \\ \mathcal{R}_3 &= \{(x,y) : x > \alpha, y \leq \alpha\}\end{aligned}$$

with $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}\Pr(Y \leq \alpha) &= \iint_{\mathcal{R}_1 \cup \mathcal{R}_3} p_{XY}(x,y)dxdy \\ &= \iint_{\mathcal{R}_1} p_{XY}(x,y)dxdy \\ &\quad + \iint_{\mathcal{R}_3} p_{XY}(x,y)dxdy \qquad (88)\end{aligned}$$

and

$$\begin{aligned}\Pr(X \leq \alpha) &= \iint_{\mathcal{R}_1 \cup \mathcal{R}_2} p_{XY}(x,y)dxdy \\ &= \iint_{\mathcal{R}_1} p_{XY}(x,y)dxdy \\ &\quad + \iint_{\mathcal{R}_2} p_{XY}(x,y)dxdy \qquad (89)\end{aligned}$$

but

$$\begin{aligned}\iint_{\mathcal{R}_3} p_{XY}(x,y)dxdy &= \iint_{x>\alpha, y\leq\alpha} p_{XY}(x,y)dxdy \qquad (90) \\ &\leq \iint_{y\leq\alpha, x>\alpha} p_{XY}(y,x)dydx \\ &= \iint_{\mathcal{R}_2} p_{XY}(x,y)dxdy. \qquad (91)\end{aligned}$$

Hence,

$$\begin{aligned}\Pr(Y \leq \alpha) &\leq \iint_{\mathcal{R}_1} p_{XY}(x,y)dxdy + \iint_{\mathcal{R}_2} p_{XY}(x,y)dxdy \\ &= \Pr(X \leq \alpha). \qquad (92)\end{aligned}$$

REFERENCES

[1] S. D. Blostein and H. Leib, "Multiple antenna systems: Role and impact in future wireless access," *IEEE Commun. Mag.*, vol. 41, no. 7, pp. 94–101, Jul. 2003.

[2] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *IEEE Int. Symp. Signals, Syst., Electron.*, Pisa, Italy, Oct. 1998, pp. 295–300.

[3] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, Nov. 1999.

[4] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[5] A. Paulraj, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[6] J. Benesty, Y. A. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1722–1730, Jul. 2003.

[7] S. Loyka and F. Gagnon, "Performance analysis of the V-BLAST algorithm: An analytical approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1326–1337, Jul. 2004.

[8] N. Prasad and M. K. Varanasi, "Analysis of decision feedback detection for MIMO Rayleigh-fading channels and the optimization of power and rate allocations," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1009–1025, Jun. 2004.

[9] Y. Jiang, X. Zheng, and J. Li, "Asymptotic analysis of V-BLAST," in *Proc. IEEE GLOBECOM*, St. Louis, MO, Nov. 2005, pp. 3882–3886.

[10] M. O. Damen, H. E. 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.

[11] 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.

[12] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[13] L. G. Barbero and J. S. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2131–2142, Jun. 2008.

[14] J. Jálden, L. G. Barbero, B. Ottersten, and J. S. Thompson, "Full diversity detection in MIMO systems with a fixed-complexity sphere decoder," in *Proc. IEEE ICASSP*, Apr. 2007, pp. 49–52.

[15] J. Jálden, L. G. Barbero, B. Ottersten, and J. S. Thompson, "The error probability of the fixed-complexity sphere decoder," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2711–2720, Jul. 2009.

[16] Y. Li and Z. Q. Luo, "Parallel detection for V-BLAST system," in *Proc. IEEE ICC*, New York, May 2002, pp. 340–344.

[17] H. Sung, K. B. Lee, and J. W. Kang, "A simplified maximum likelihood detection scheme for MIMO systems," in *Proc. IEEE VTC*, Oct. 2003, pp. 419–423.

[18] Z. Lei, Y. Dai, and S. Sun, "A low complexity near ML V-BLAST algorithm," in *Proc. IEEE VTC-Fall*, Dallas, TX, Sep. 2005, pp. 942–946.

[19] D. W. Waters and J. R. Barry, "The Chase family of detection algorithms for multiple-input multiple-output channels," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 739–747, Feb. 2008.

[20] Z. Luo, M. Zhao, S. Liu, and Y. Liu, "Generalized parallel interference cancellation with near-optimal detection performance," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 304–312, Jan. 2008.

[21] G. H. Golub and C. F. V. Loan, *Matrix Compuations*, 3rd ed. Baltimore, MD: John Hopkins Univ. Press, 1996.

[22] G. A. F. Seber, *A Matrix Handbook for Statisticians*. New York: Wiley, 2007.

[23] X. Zhu and R. D. Murch, "Performance analysis of maximum likelihood detection in a MIMO antenna system," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 187–191, Feb. 2002.

[24] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes, "On the diversity order of spatial multiplexing systems with transmit antenna selection: A geometrical approach," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5297–5311, Dec. 2006.

[25] R. W. Heath, Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[26] J. Jaldén and B. Ottersten, "On the maximal diversity order of spatial multiplexing with transmit antenna selection," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4273–4276, Nov. 2007.

[27] Y. Jiang, M. K. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An in-depth study of the high SNR regime," *IEEE. Trans. Inf. Theory*, 2009, accepted for publication.

[28] D. S. Watkins, *Fundamentals of Matrix Computations*, 2nd ed. New York: Wiley, 2002.

[29] Y. Shang and X.-G. Xia, "An improved fast recursive algorithm for V-BLAST with optimal ordered detections," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 756–760.
[30] L. G. Barbero and J. S. Thompson, "Performance of the complex sphere decoder in spatially correlated MIMO channels," *IET, Commun.*, vol. 1, no. 1, pp. 122–130, Feb. 2007.
[31] D. Pham, K. R. Pattipati, P. K. Willett, and J. Luo, "An improved complex sphere decoder for V-BLAST systems," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 748–751, Sep. 2004.
[32] J. Lu, K. B. Letaief, J. C.-I. Chuang, and M. L. Liou, "M-PSK and M-QAM BER computation using signal-space concepts," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 181–184, Feb. 1999.
[33] A. van Zelst and J. Hammerschmidt, "A single coefficient spatial correlation model for multiple-input multiple-output (MIMO) radio channels," in *Proc. 27th General Assembly Int. Union of Radio Science (URSI)*, Aug. 2002, pp. 17–24.
[34] E. G. Larsson, "Diversity and channel estimation errors," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 205–208, Feb. 2004.

Djelili Radji (S'08) was born in Cotonou, Benin, in 1983. He received the B.Eng. degree in electrical engineering from McGill University, Montreal, QC, Canada, in 2005. He is currently pursuing the Ph.D. degree in electrical engineering at the same university.

Since 2005, he has been a Teaching Assistant with the Electrical Engineering Department, McGill University. His research interests are in the area of multiple-input multiple-output (MIMO) wireless communications, including MIMO spatial multiplexing schemes and channel coding for MIMO systems.

Harry Leib was born in 1953. He received the B.Sc. (*cum laude*) and M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, Israel, in 1977 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1987.

From 1977 to 1984, he was with the Israel Ministry of Defense, working in the area of Communication Systems. During his Ph.D. studies at the University of Toronto (1984–1987), he was a Teaching and Research Assistant in the Department of Electrical Engineering, working in the areas of Digital Communications and Signal Processing. After completing the Ph.D. degree, he was with the University of Toronto as a Postdoctoral Research Associate in Telecommunications (September 1987 to December 1988) and as an Assistant Professor (January 1989 to August 1989). Since September 1989, he has been with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, initially as an Assistant Professor, then as an Associate Professor, and now as a Full Professor. He spent part of his Sabbatical leave of absence at Bell Northern Research in Ottawa, ON, Canada (September 1995 to February 1996) working in a CDMA-related project. During the other part of his sabbatical leave of absence (March 1996 to August 1996), he was a Visiting Professor in the Communications Lab/Institute of Radio Communications, Helsinki University of Technology, Espoo Finland, where he taught a condensed course on channel coding and modulation and worked with graduate students in the communications area. At McGill University, he teaches undergraduate and graduate courses in communications, and directs the research of graduate students. His current research activities are in the areas of digital communications, wireless communication systems, and information theory.

Dr. Leib has been an Editor for communication and information theory for the IEEE TRANSACTIONS ON COMMUNICATIONS since 2000, and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY from 2001 to 2007. From 2003 to 2005, he has been a Guest Coeditor for a special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATION on differential and noncoherent wireless communication.