RESEARCH ARTICLE

Detection of Spatial-Modulated Signals in the Presence of CSI Uncertainty and Spatially- and Time-Selective Fading

HSUAN-CHENG CHANG[1], YEN-CHENG LIU[ID][1], AND YU T. SU[ID][2], (Life Senior Member, IEEE)
[1]MediaTek Inc., Hsinchu 30078, Taiwan
[2]Institute of Communications Engineering, National Yang-Ming Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding author: Yu T. Su (ytsu@nycu.edu.tw)

ABSTRACT Although the channel state information (CSI) is never perfect and varies in time, most studies on spatial modulation (SM) systems assume perfectly known CSI and time-invariant channel. The spatial correlations among multiple spatial subchannels, which have to be considered when CSI is imperfect, are also often neglected. In this paper, we release the above assumptions and take the CSI uncertainty along with the spatial-temporal selectivities into account. We derive the channel estimation error aware maximum likelihood (CEEA-ML) detectors as well as several low-complexity alternatives for PSK and QAM signals. As the CSI uncertainty depends on the channel estimator (CE) used, we consider two well-known CEs in our study. The error rate performance of the ML and some suboptimal detectors is analyzed. Numerical results obtained by simulations and analysis show that the CEEA-ML detectors offer clear performance gain against conventional mismatched SM detectors and, in many cases, the proposed suboptimal detectors incur only minor performance loss.

INDEX TERMS Imperfect channel state information, maximum likelihood, signal detection, space-time channel correlation, spatial modulation.

I. INTRODUCTION

Spatial modulation (SM), as it allows only a subset of the transmit antennas to be active in any transmission interval and exploits the transmit antenna index to carry extra information [1], is a low-complexity and spectral-efficient multi-antenna-based transmission scheme. Besides requiring no or smaller number of multiple radio frequency (RF) transmit chains, its low complexity requirement is also due to the fact that the inter-spatial channel interference (ICI) is either nonexistent or less severe.

Most receiver performance assessments on multi-antenna systems, coded or uncoded, assume that the channel state information (CSI) is perfectly known by the receiver (e.g., [1], [3], and [5]). In practice, the CSI at the receiver (CSIR) is obtained by a pilot-assisted or decision-directed (DD) estimator [2] and is never perfect. The impact of imperfect

CSIR on some MIMO detectors was considered in [4], [6], and [7] to evaluate the detectors' performance loss. Furthermore, the channel is usually not static, especially in a mobile environment, but the channel-aging effect, i.e., the impact of outdated CSI, is neglected in most of these studies except for [7]. Reference [8] discussed the effect of the channel coefficients' *phase estimation error* while [9] analyzed the performance of the conventional detector in the presence of white Gaussian channel estimation error that is independent of the channel estimator (CE) used. Other earlier works on SM detection all employ the perfect CSI and time-invariant channel assumptions thus give suboptimal performance in practical environments [10]. One can reduce the CSI uncertainty induced performance degradation by using a better CSI estimator or using a differential spatial modulation (DSM) scheme [11]. Although DSM can be noncoherently detected without the need of CSI, it is less power efficient and works for slowly time-varying environments only. The DSM also imposes limitations on antenna selection therefore is less

spectral efficient and is applicable for constant modulus signals (e.g., PSK) but not for QAM or PAM signals. Moreover, CSI acquisition is an essential part of an advanced communication receiver not only for data detection but also for link adaptation and resource allocation purposes.

We adopt another approach by taking into account the CSI uncertainty and the spatial/time selectivity, substantially extending the work reported in [12]. For some regular constellations, the complexity of the corresponding ML detectors can be reduced [13], [14]). However, these detectors assume perfect CSI and therefore the complexity reduction is not attainable when the detector takes CSI uncertainty into account. Moreover, their performance in the presence of CSI error may degrade significantly because of the mismatched channel statistics.

The difficulty in deriving the optimal SM detector in a spatially- and time-selective fading channel with imperfect CSI is due to the fact that the likelihood function depends not only on the transmitted symbol but also on the CSI estimator and its performance. In this paper, we remove the perfect CSI and static channel assumptions and consider two MIMO CEs/trackers for systems using a frame structure which consists of a pilot block and several data blocks. The first scheme is a decision-directed (DD) estimator (tracker) while the second one is referred to as the model-based (MB) estimator [15]. The frame structure and the CEs considered are described in Section II-C.

Our *major contributions* are summarized as follows. We derive a general *channel estimation error-aware* (CEEA) maximum likelihood (ML) receiver structure for detecting general MIMO signals with two known CEs in spatially- and time-selective fading channels. The CEEA-ML detectors for M-PSK and M-QAM based SM systems are obtained by specializing to the combined SM and PSK/QAM signal formats. As the ML detectors require high computational complexity, we develop two classes of low-complexity suboptimal detectors. The first one detects the transmit antenna index and symbol sequentially (resulting in two-stage detectors), and has a detection complexity that is independent of the signal modulation order. The second class simplifies the likelihood functions by using lower-dimension approximations that neglect the spatial correlation on either the receive or transmit side. The approximations lead to zero transmit correlation (ZTC) and zero receive correlation (ZRC) receivers. Both simplifications–separate detections and dimension reduction of the likelihood function–can be combined to yield even simpler detector structures. As will be verified later, these low-complexity detectors do not incur much performance loss. Except for the two-stage detectors, we analyze the error rate performance and confirm the accuracy of the analysis by simulations. A model-based two-stage spatial correlation estimator is developed as well. We summarize the equation numbers of the derived detectors in Table 1 for convenience of reference.

The rest of this paper is organized as follows. In Section II we present a general space-time (S-T) MIMO channel model,

TABLE 1. Formulae associated with various CE-aided detectors.

CE		MB	DD
CEEA-ML	General	(24)	(35)
	PSK-SM	(27)	(37)
	QAM-SM	(29)	(38)
Two-stage		(34)	(40)
ZRC		(45)	(56)
Two-stage ZRC		(53a), (53b)	(61a), (61b)
ZTC		(50)	(58)

review the corresponding perfect CSI based ML detector and introduce both DD and MB based CEs. We refer to the latter two CEs as DD-CE and MB-CE henceforth. In Section III, we focus on MB-CE-aided systems. A general CEEA-ML detector for general MIMO signals is first derived (cf. (24)), followed by those for M-PSK and M-QAM SM signals (cf. (27) and (29)). A low-complexity two-stage receiver for M-PSK SM systems is given at the end of the section (cf. (34)). Section IV presents similar derivations for DD-CE-aided SM systems (cf. (35), (37), (38), and (40)). In Section V, we develop simplified CEEA-ML and two-stage detectors for both MB- and DD-CE-aided SM systems by using lower-dimension approximations of the likelihood functions. The error rate performance of various detectors we derived is analyzed in Section VI. The analytic approach is valid for all but the two-stage detectors. Because of space limitation, the presentation is concise, skipping most detailed derivations (cf. (45), (50), (53a), (56), (58), and (61a)). The computational complexity and memory requirement of the mentioned detectors is analyzed in Section VII. As far as we know, materials presented in Sections III–VII are new. Numerical performance of our detectors and conventional mismatched detector is given in Section VIII. Finally, we recapitulate our main results and findings in Section IX.

Notations: Upper and lower case bold symbols denote matrices and vectors, respectively. $\mathbf{I}_N$ is the $N \times N$ identity matrix and $\mathbf{0}_N$ the $N \times 1$ all-zero vector. $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^\dagger$ represent the transpose, element-wise conjugate, conjugate transpose, inverse, and pseudo-inverse of the enclosed items, respectively. $\text{vec}(\cdot)$ is the operator that forms one tall vector by stacking the columns of a matrix. $\mathbb{E}\{\cdot\}$, and $\| \cdot \|_F$ denote the expectation and Frobenius norm of the enclosed items, respectively, $\otimes$ denotes the Kronecker product and $\odot$ the Hadamard product. $(\cdot)_i$ and $[\cdot]_{ij}$ denote the ith row and (i, j)th element of the enclosed matrix, respectively. $\text{Diag}(\cdot)$ translates the enclosed vector or elements into a diagonal matrix with the nonzero terms being the enclosed items, whereas $\text{DIAG}(\cdot)$ is defined by

$$\text{DIAG}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M) = \begin{bmatrix} \mathbf{x}_1 & \mathbf{0}_{N_1} & \cdots & \mathbf{0}_{N_1} \\ \mathbf{0}_{N_2} & \mathbf{x}_2 & \cdots & \mathbf{0}_{N_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_M} & \mathbf{0}_{N_M} & \cdots & \mathbf{x}_M \end{bmatrix}$$

TABLE 2. Statistics used in MB-CE-aided detectors, where the subscript SSK stands for space-shift keying [8].

Conditional		Mean	Covariance
CEEA-ML	General	$\mathbf{m}_{mb}(k)$ (23a)	$\mathbf{C}_{mb}(k)$ (23b)
	PSK-SM	$\bar{\mathbf{m}}_{ssk}(k)$	$\bar{\mathbf{C}}_{psk}(k)$
	QAM-SM	$\bar{\mathbf{m}}_{ssk}(k)$	$\bar{\mathbf{C}}_{qam}(k)$ (30)
Two-stage		$\bar{\mathbf{m}}_{ssk}(k)$	$\bar{\mathbf{C}}_{psk}(k)$
ZRC		$\bar{\mathbf{M}}_{zrc}(k)$	$\bar{\mathbf{C}}_{zrc}(k)$ (44b)
Two-stage ZRC		$\bar{\mathbf{M}}_{ssk}(k)$	$\bar{\mathbf{C}}_{ssk}(k)$ (52)
ZTC		$\bar{\mathbf{M}}_{ztc}(k)$ (49)	$\bar{\mathbf{C}}_{ztc}(k)$ (48)

TABLE 3. Statistics used in DD-CE-aided detectors.

Conditional		Mean	Covariance
CEEA-ML	General	$\mathbf{m}_{dd}$ (36a)	$\mathbf{C}_{dd}$ (36b)
	PSK-SM	$\tilde{\mathbf{m}}_{ssk}$	$\tilde{\mathbf{C}}_{psk}$
	QAM-SM	$\tilde{\mathbf{m}}_{ssk}$	$\tilde{\mathbf{C}}_{qam}$
Two-stage		$\tilde{\mathbf{m}}_{ssk}$	$\tilde{\mathbf{C}}_{psk}$
ZRC		$\tilde{\mathbf{M}}_{zrc}$	$\tilde{\mathbf{C}}_{zrc}$ (55)
Two-stage ZRC		$\tilde{\mathbf{M}}_{ssk}$ (59a)	$\tilde{\mathbf{C}}_{ssk}$ (59b)
ZTC		$\tilde{\mathbf{M}}_{ztc}$ (57b)	$\tilde{\mathbf{C}}_{ztc}$ (57c)

with vector length of $\mathbf{x}_i$ being N_i. Finally, the determinant of a matrix $\mathbf{A}$ is denoted by $\det \mathbf{A}$.

The detector structures we develop need to compute a common quadratic form and involve various conditional mean vectors and covariance matrices as a function of the data block index k. For notational brevity and the convenience of reference, we list the latter conditional parameters in Tables 2 and 3 and define

$$\mathcal{G}(\Gamma, \boldsymbol{\chi}) \overset{def}{=} \boldsymbol{\chi}^H \Gamma^{-1} \boldsymbol{\chi} \quad (1)$$

where $\Gamma \in \mathbb{C}^{N_1 \times N_1}$ is invertible and $\boldsymbol{\chi} \in \mathbb{C}^{N_1 \times N_2}$.

II. PRELIMINARIES
A. S-T CORRELATED CHANNELS
We consider a MIMO system with N_T transmit and N_R receive antennas and assume a block-faded scenario in which the MIMO channel remains static within a block of B channel uses but varies from block to block. The parameter B controls the channel's time-selectivity and should be chosen to be less than the channel's coherent time. For this system, the received sample matrix of block k can be expressed as

$$\mathbf{Y}(k) \overset{def}{=} [\mathbf{y}_1(k), \cdots, \mathbf{y}_B(k)] \overset{def}{=} \left[\underline{\mathbf{y}}_1^T(k), \cdots, \underline{\mathbf{y}}_{N_R}^T(k) \right]^T$$
$$= \mathbf{H}(k)\mathbf{X}(k) + \mathbf{Z}(k), \quad (2)$$

where $\mathbf{y}_i(k)$ are column vectors, $\underline{\mathbf{y}}_i(k)$ are row vectors, and

$$\mathbf{H}(k) \overset{def}{=} [h_{ij}(k)] = [\mathbf{h}_1(k), \cdots, \mathbf{h}_{N_T}(k)] \quad (3a)$$
$$= [\underline{\mathbf{h}}_1^T(k), \cdots, \underline{\mathbf{h}}_{N_R}^T(k)]^T \quad (3b)$$

is the narrowband $N_R \times N_T$ channel matrix and

$$\mathbf{X}(k) = [\mathbf{x}_1(k), \cdots, \mathbf{x}_B(k)] = [\underline{\mathbf{x}}_1^T(k), \cdots, \underline{\mathbf{x}}_{N_T}^T(k)]^T \quad (4)$$

is the $N_T \times B$ ($B \geq N_T$) matrix containing the modulated data and/or pilot symbols and the entries of the noise matrix $\mathbf{Z}(k)$ are i.i.d. random variables with distribution $\mathcal{CN}(0, \sigma_z^2)$.

Let $\Phi = \mathbb{E}\left\{ \text{vec}(\mathbf{H}(k)) \text{vec}^H(\mathbf{H}(k)) \right\}$ be the matrix whose $(N_R N_T)^2$ entries represent the correlation coefficients amongst spatial subchannels.

$$\text{vec}(\mathbf{H}(k)) = \Phi^{\frac{1}{2}} \text{vec}(\mathbf{H}_w(k)), \quad (5)$$

where $\mathbf{H}_w(k)$ is an $N_R \times N_T$ matrix with i.i.d. zero-mean, unit-variance complex Gaussian random variables as its elements. We assume that the spatial correlation matrix Φ is either completely or partially known to the receiver in deriving various detector structures. The numerical effect of using an estimated Φ is studied in Section VIII.

We further assume that the spatial and temporal correlations are separable [16], i.e.,

$$\mathbb{E}\{h_{ij}(k)h_{mn}^*(\ell)\} = \rho_S(i-m, j-n) \times \rho_T(k-\ell)$$

with $\rho_T(k-\ell) \overset{def}{=} \mathbb{E}\{h_{ij}(k)h_{ij}^*(\ell)\}$ denoting the (i,j)th spatial subchannel's time correlation coefficient with lag $k-\ell$ while

$$\rho_S(i-m, j-n; k) \overset{def}{=} \mathbb{E}\{h_{ij}(k)h_{mn}^*(k)\}$$
$$= [\Phi]_{(j-1)N_R+i,\,(n-1)N_R+m} \quad (6)$$

is the correlation between the (i,j)th and (m,n)th spatial subchannels. As mentioned before, conditioned on the estimated channel matrix, the likelihood function is a function of these S-T correlations and so are the corresponding CEEA-ML detectors.

### B. DETECTION OF SM SIGNALS WITH PERFECT CSIR
The SM scheme can avoid ICI by imposing the single active antenna constraint and compensates for the corresponding data rate reduction by using the transmit antenna index to carry extra information bits [1].

An m-bit/transmission SM system partitions the data stream into groups of $m = \log_2(MN_T)$ bits of which the first $\log_2 N_T$ bits of a group are used to determine the transmit antenna and the remaining bits of the group are mapped into a symbol in the constellation $\mathcal{A}_M$ of size M for the selected antenna to transmit. Since only one transmit antenna is active in each transmission, the jth column of $\mathbf{X}(k)$ is of the form

$$\mathbf{x}_j(k) = [0, \cdots, 0, x_{\ell_j}(k), 0, \cdots, 0]^T, \quad (7)$$

where $\ell_j \overset{def}{=} \ell_j(k)$ is the active antenna index and $x_{\ell_j}(k) \overset{def}{=} s_j(k) \in \mathcal{A}_M$ is the modulated symbol transmitted at the jth symbol interval of the kth block. A transmitted symbol block $\mathbf{X}$ can thus be decomposed as $\mathbf{X} = \mathbf{LS}$, where $\mathbf{S} = \text{Diag}(\mathbf{s})$, $\mathbf{s} = [s_1, \cdots, s_B]^T$, $s_j \in \mathcal{A}_M$ and $\mathbf{L}$ is the $N_T \times B$ space-shift keying (SSK) matrix [9] defined as

$$[\mathbf{L}]_{ij} = \begin{cases} 1, & \text{if } i = \ell_j; \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$
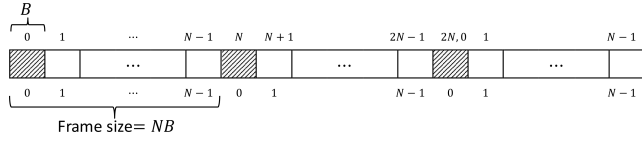
FIGURE 1. Proposed frame structure. A pilot block, shown as a shaded area, is inserted every N blocks. The block indices for MB-CE-aided systems ($0 \leq k \leq 2N$) and those for DD counterparts ($0 \leq k \leq N - 1$) are labeled on the top and bottom of the blocks, respectively.

We define $\mathcal{X}$ as the set of all $N_T \times B$ matrices whose columns are of the type (7). The average power

$$\varepsilon_s \stackrel{def}{=} \frac{1}{B} \mathbb{E} \left\{ \|\mathbf{X}(k)\|_F^2 \right\} = \frac{1}{B} \mathbb{E} \left\{ \text{tr} \left(\mathbf{X}(k) \mathbf{X}^H(k) \right) \right\} \quad (9)$$

is equivalent to the average power of $\mathcal{A}_M$. Therefore, for $j = 1, \cdots, B$, the jth column vector in (2) can be written as

$$\mathbf{y}_j(k) = \mathbf{h}_{\ell_j}(k) s_j(k) + \mathbf{z}_j(k). \quad (10)$$

Assuming perfect CSIR, i.e., $\hat{\mathbf{H}}(k) = \mathbf{H}(k)$, and i.i.d. source, we have the ML detector

$$\hat{\mathbf{X}}(k) = \arg \max_{\mathbf{X} \in \mathcal{X}} P\left(\mathbf{Y}(k) | \mathbf{H}(k), \mathbf{X} \right) \quad (11)$$

where

$$P\left(\mathbf{Y}(k) | \mathbf{H}(k), \mathbf{X}(k) \right)$$
$$= \left(\pi \sigma_z^2 \right)^{-N_R} \exp \left(-\frac{1}{\sigma_z^2} \|\mathbf{Y}(k) - \mathbf{H}(k)\mathbf{X}(k)\|_F^2 \right).$$

With $\mathbb{L} \stackrel{def}{=} \{1, \cdots, N_T\}$, (11) is simplified as

$$\left(\hat{s}_j(k), \hat{\ell}_j(k) \right) = \arg \min_{(s, \ell) \in \mathcal{A}_M \times \mathbb{L}} \|\mathbf{y}_j(k) - \mathbf{h}_\ell s\|_F^2 \quad (12)$$

for $j = 1, \cdots, B$.

C. TRACKING TIME-VARYING MIMO CHANNELS

The frame structure considered is depicted in Fig. 1, where each frame consists of a pilot (shadowed) block and $N-1$ data blocks with a total length of NB symbol intervals. A smaller N implies higher pilot density and better channel tracking capability but reduced spectral efficiency [17]. Let the pilot block be transmitted at the k_pth block and denote the $N_T \times B$ pilot matrix by $\mathbf{X}(k_p) = \mathbf{X}_p$. The corresponding received block is

$$\mathbf{Y}(k_p) = \mathbf{H}(k_p)\mathbf{X}_p + \mathbf{Z}(k_p) \quad (13)$$

with the average pilot symbol power given by $\varepsilon_p \stackrel{def}{=} \frac{1}{B}\|\mathbf{X}_p\|_F^2$ and $\mathbf{X}_p$ a unitary matrix. In particular, for SM systems, we assume that $B = N_T$ and $\mathbf{X}_p = \sqrt{\varepsilon_p}\mathbf{I}_{N_T}$.

We now describe the pilot-assisted MB and DD channel trackers (estimators) to be considered in subsequent discourse.

1) MB CHANNEL ESTIMATOR (MB-CE)

For a single link with moderate mobility and frame size, it is reasonable to assume that the (i, j)th component of the channel matrix $\mathbf{H}(k)$ is a quadratic function of the sampling epoch, $\{k\}$ [15]

$$h_{ij}(k) = \alpha_{ij}(k)k^2 + \beta_{ij}(k)k + \gamma_{ij}(k). \quad (14)$$

Define the coefficient vector $\boldsymbol{\xi}_{ij}(k) \stackrel{def}{=} [\alpha_{ij}(k), \beta_{ij}(k), \gamma_{ij}(k)]^T$. By collecting the received samples at three consecutive pilot locations $\mathbf{Y}(k_p), \mathbf{Y}(k_p + N), \mathbf{Y}(k_p + 2N)$, we update the estimate for $\boldsymbol{\xi}_{ij}(k)$ every two frames via

$$\hat{\boldsymbol{\xi}}_{ij}(k_p) \stackrel{def}{=} \begin{bmatrix} \hat{\alpha}_{ij}(k_p) \\ \hat{\beta}_{ij}(k_p) \\ \hat{\gamma}_{ij}(k_p) \end{bmatrix} = \mathbf{T}^{-1}(k_p) \, \tilde{\mathbf{y}}_{ij}(k_p) \quad (15)$$

where

$$\mathbf{T}(k) \stackrel{def}{=} \begin{bmatrix} k^2 & k & 1 \\ (k + N)^2 & k + N & 1 \\ (k + 2N)^2 & k + 2N & 1 \end{bmatrix}, \quad (16a)$$

$$\tilde{\mathbf{y}}_{ij}(k_p) \stackrel{def}{=} \frac{1}{\sqrt{\varepsilon_p}} \begin{bmatrix} y_{ij}(k_p) \\ y_{ij}(k_p + N) \\ y_{ij}(k_p + 2N) \end{bmatrix}$$
$$= \begin{bmatrix} h_{ij}(k_p) \\ h_{ij}(k_p + N) \\ h_{ij}(k_p + 2N) \end{bmatrix} + \frac{1}{\sqrt{\varepsilon_p}}\tilde{\mathbf{z}}_{ij}(k_p), \quad (16b)$$

and $\tilde{\mathbf{z}}_{ij}(k_p) \sim \mathcal{CN}(\mathbf{0}_3, \frac{\sigma_z^2}{\varepsilon_p}\mathbf{I}_3)$. From (14), (15) and the definition $\mathbf{t}(k) = [k^2, k, 1]^T$, we obtain the MB-CEs

$$\hat{h}_{ij}(k) = \mathbf{t}^H(k)\hat{\boldsymbol{\xi}}_{ij}(k_p) = \mathbf{t}^H(k)\,\mathbf{T}^{-1}(k_p)\,\tilde{\mathbf{y}}_{ij}(k_p) \quad (17)$$

for the channel responses at the $k_p, \cdots, (k_p + 2N - 1)$th blocks. Performance of this MB-CE as a function of the pilot density (N) and signal-to-noise ratio (SNR) can be found in [15].

2) DD CHANNEL ESTIMATOR (DD-CE)

The DD-CE uses the pilot block at the beginning of a frame to obtain an estimate of the S-T channel say $\hat{\mathbf{H}}(k_p)$ for detecting the data of the next data block. This newly detected data block $\hat{\mathbf{X}}(k_p + 1)$ is then treated as pilots (called pseudo-pilots) to update the CSI, e.g., for least squares (LS) CE, $\hat{\mathbf{H}}(k_p + 1) = \mathbf{Y}(k_p + 1)\hat{\mathbf{X}}^\dagger(k_p + 1)$, which is in turn used for detecting the following data block $\hat{\mathbf{X}}(k_p + 2)$. Such a pseudo-pilot-based channel estimation-data detection procedure repeats until the last (Nth) data block of the frame is detected and a new pilot block arrives to keep tracking the time-varying channel. Error propagation, if exists, is terminated when the new frame starts and a new pilot block $\mathbf{X}(k_p + N)$ becomes available.

We note that only one element in each column of the kth detected block $\hat{\mathbf{X}}(k)$ is nonzero, hence it is likely that not all vectors of CEs $\hat{\mathbf{h}}_\ell(k)$'s are updated at each data block. The outdated CSI will cause a DD system's performance loss in a time-varying channel. We denote by $\hat{\mathbf{H}}[\mathfrak{L}](k)$, the submatrix containing only the columns associated with the

set of antenna indices activated in block k, $\mathfrak{L} \subseteq \{1, \cdots, N_T\}$, i.e.,

$$\hat{\mathbf{H}}[\hat{\mathfrak{L}}(k)](k) \leftarrow \mathbf{Y}(k)\bar{\mathbf{X}}^\dagger(k),$$

where $\hat{\mathfrak{L}}(k) \stackrel{def}{=} \{\hat{\ell}_1(k), \cdots, \hat{\ell}_B(k)\}$ and $\bar{\mathbf{X}}(k)$ is the truncated $\hat{\mathbf{X}}(k)$ with its all-zero rows removed. By combining the submatrix which consists of those channel vectors estimated in the previous (the $(k-1)$th) block

$$\hat{\mathbf{H}}[\mathbb{L} \setminus \hat{\mathfrak{L}}(k)](k) = \hat{\mathbf{H}}[\mathbb{L} \setminus \hat{\mathfrak{L}}(k)](k-1),$$

we obtain a full-rank estimate $\hat{\mathbf{H}}(k)$ and, with a slight abuse of notation, denote $\mathbf{Y}(k)\hat{\mathbf{X}}^\dagger(k)$ as $\hat{\mathbf{H}}(k)$.

Remark 1: As the DD method uses the CE obtained in the previous block for demodulating the current block's data, error propagation within a frame is inevitable. The MB approach avoids error propagation at the cost of increased latency and storage requirement. The DD method updates the CSI estimate in each block while the channel's time variation is taken into account by the MB approach through the model (14) which leads to the estimate (17) in an N-block frame. $\square$

In the remainder of this paper, we use the normalization $\varepsilon_p = 1$ and in Section VIII, $\varepsilon_p = \varepsilon_s = 1$.

III. CHANNEL ESTIMATION ERROR-AWARE ML DETECTION WITH MB CHANNEL ESTIMATES

As defined in [4], a *mismatched detector* is the one which replaces $\mathbf{H}(k)$ in (11) by the estimated CSI $\hat{\mathbf{H}}$

$$\hat{\mathbf{X}}^{\mathrm{MM}}(k) \stackrel{def}{=} \arg\min_{\mathbf{X} \in \mathcal{X}} \|\mathbf{Y}(k) - \hat{\mathbf{H}}\mathbf{X}\|_F^2. \quad (18)$$

In our case, $\hat{\mathbf{H}}$ is either $\hat{\mathbf{H}}(k)$ or $\hat{\mathbf{H}}(k-1)$ depending on whether an MB or DD estimator is used.

We extend the basic approach of [4], which is similar to that used in deriving the classic discrete-time Kalman filter that was based on the following lemma:

Lemma 1: [18, Thm. 10.2] Let $\mathbf{z}_1$ and $\mathbf{z}_2$ be circularly symmetric complex joint Gaussian random vectors with zero means and full-rank covariance matrices $\Sigma_{ij} \stackrel{def}{=} \mathbb{E}\{\mathbf{z}_i \mathbf{z}_j^H\}$. Then, conditioned on $\mathbf{z}_2$, the random vector $\mathbf{z}_1$ is circularly symmetric Gaussian with mean $\Sigma_{12}\Sigma_{22}^{-1}\mathbf{z}_2$ and covariance matrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

As mentioned before, we further take the CE used, the channel aging effect and the spatial correlation into account. We assume a spatially/temporally correlated block fading channel and refer to the resulting detectors as *channel estimation error-aware (CEEA)-ML detectors*.

A. GENERAL MIMO SIGNAL DETECTION WITH IMPERFECT CSI

When the MB-CE $\hat{\mathbf{H}}(k)$ is used, the *CEEA-ML* MIMO detector has to compute

$$\arg\max_{\mathbf{X} \in \mathcal{A}_{M+}^{N_T \times B}} P\left[\mathrm{vec}(\mathbf{Y}(k)) \big| \mathrm{vec}(\mathbf{X}), \mathrm{vec}(\hat{\mathbf{H}}(k))\right], \quad (19)$$

where $\mathcal{A}_{M+} \stackrel{def}{=} \mathcal{A}_M \cup \{0\}$. Since all the entries of $\mathbf{Y}(k)$ and $\hat{\mathbf{H}}(k)$ are zero-mean random variables, we invoke *Lemma 1* with

$$\mathbf{z}_1 = \mathrm{vec}(\mathbf{Y}(k))$$
$$= \left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}\right)\mathrm{vec}(\mathbf{H}(k)) + \mathrm{vec}(\mathbf{Z}(k)), \quad (20a)$$

$$\mathbf{z}_2 = \mathrm{vec}(\hat{\mathbf{H}}(k))$$
$$= \left[\mathrm{vec}(\mathbf{Y}(k_p))\ \mathrm{vec}(\mathbf{Y}(k_p + N))\ \mathrm{vec}(\mathbf{Y}(k_p + 2N))\right]$$
$$\times \left(\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right)^T, \quad (20b)$$

for $k_p < k < k_p + 2N$, to obtain

$$\Sigma_{11} = \mathbb{E}\{\mathbf{z}_1 \mathbf{z}_1^H\}$$
$$= \left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}\right)\Phi(\mathbf{X}^*(k) \otimes \mathbf{I}_{N_R}) + \sigma_z^2 \mathbf{I}_{N_R N_T}, \quad (21a)$$

$$\Sigma_{12} = \mathbb{E}\{\mathbf{z}_1 \mathbf{z}_2^H\} = \mathbb{E}\{(\mathbf{z}_2 \mathbf{z}_1^H)^H\} = \Sigma_{21}^H$$
$$= \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}\right)\Phi, \quad (21b)$$

$$\Sigma_{22} = \mathbb{E}\{\mathbf{z}_2 \mathbf{z}_2^H\}$$
$$= \nu(k)\Phi + \sigma_z^2 \left\|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right\|_F^2 \mathbf{I}_{N_R N_T}, \quad (21c)$$

where

$$\mathbf{q}(k) = \left[\rho_T(k - k_p), \rho_T(k - k_p - N), \rho_T(k - k_p - 2N)\right]^T, \quad (22a)$$

$$\nu(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\begin{bmatrix} 1 & \rho_T(N) & \rho_T(2N) \\ \rho_T(N) & 1 & \rho_T(N) \\ \rho_T(2N) & \rho_T(N) & 1 \end{bmatrix}$$
$$\times \left(\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right)^H. \quad (22b)$$

Direct substitutions of the above covariance matrices give the mean vector $\mathbf{m}_{mb}(k)$ and covariance matrix $\mathbf{C}_{mb}(k)$

$$\mathbf{m}_{mb}(k) = \left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}\right)\mathbf{A}(k)\mathrm{vec}(\hat{\mathbf{H}}(k)), \quad (23a)$$

$$\mathbf{C}_{mb}(k) = \sigma_z^2 \mathbf{I}_{N_R N_T} + \left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}\right)\left[\mathbf{I}_{N_R N_T} - \mathbf{A}(k)\right.$$
$$\left.\left(\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\right)^*\right]\Phi\left(\mathbf{X}^*(k) \otimes \mathbf{I}_{N_R}\right) \quad (23b)$$

for the likelihood function of $\mathbf{Y}(k)$, where the spatial correlation matrix Φ also appears in

$$\mathbf{A}(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\ \Phi\ \Sigma_{22}^{-1}. \quad (23c)$$

Using (1) and (19)–(23c), we obtain a more compact form of the *CEEA-ML detector*

$$\hat{\mathbf{X}}_{mb}^{\mathrm{ML}}(k) = \arg\min_{\mathbf{X} \in \mathcal{A}_{M+}^{N_T \times B}} \mathcal{G}(\mathbf{C}_{mb}(k), \mathrm{vec}(\mathbf{Y}(k)) - \mathbf{m}_{mb}(k))$$
$$+ \log \det \mathbf{C}_{mb}(k). \quad (24)$$

An alternate derivation of Σ_{11} and Σ_{12} begins with $\hat{\mathbf{H}}(k) = \mathbf{H}(k) + \mathbf{E}(k)$. As $\mathbb{E}\{\mathrm{vec}(\mathbf{E}(k))\} = \mathbf{0}_{N_R N_T}$ and

$$\Psi_E(k) \stackrel{def}{=} \mathbb{E}\left\{\mathrm{vec}(\mathbf{E}(k))\mathrm{vec}^H(\mathbf{E}(k))\right\}$$

$$= \left(v(k) - 2\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k) + 1 \right)\Phi$$
$$+ \sigma_z^2 \left\| \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p) \right\|_F^2 \mathbf{I}_{N_R N_T}. \quad (25)$$

the received sample vector is related to $\mathbf{E}(k)$ via

$$\mathbf{Y}(k) = \hat{\mathbf{H}}(k)\mathbf{X}(k) + \left(\mathbf{Z}(k) - \mathbf{E}(k)\mathbf{X}(k) \right)$$
$$= \hat{\mathbf{H}}(k)\mathbf{X}(k) + \tilde{\mathbf{Z}}(k), \quad (26)$$

where $\text{vec}(\tilde{\mathbf{Z}}(k)) \sim \mathcal{CN}(\mathbf{0}_{N_R N_T}, \sigma_z^2 \mathbf{I}_{N_R N_T} + (\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R}) \Psi_E(k)(\mathbf{X}^*(k) \otimes \mathbf{I}_{N_R}))$. With

$$\mathbf{z}_1 = \left(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R} \right) \text{vec}(\hat{\mathbf{H}}(k)) + \text{vec}(\tilde{\mathbf{Z}}(k)),$$

we obtain Σ_{11} and Σ_{12} as given by (21a) and (21b).

Remark 2: A closer look at the components, (20a)–(23c), of the *CEEA-ML detector* (24) reveals that the spatial correlation Φ affects the auto-correlations of the received signal and estimated channel, Σ_{11} and Σ_{22}, and their cross-correlations, Σ_{12} and Σ_{21}. The influences of the time selectivity $\rho_T(\cdot)$, frame structure (a pilot block in every N blocks) and estimator structure (17) on the latter three correlations can be found in (21b) and (21c) and through $\mathbf{q}(k)$ and $v(k)$. These channel and system factors also appear in the CE error covariance $\Psi_E(k)$.

In contrast, in analyzing the performance of *mismatched detectors* and the effects of imperfect CSI, [9], [10] assume a simplified CSI error model that $\mathbf{E}(k)$ consists only of white Gaussian components which are independent of the S-T correlations and CE used, ignoring the fact that the latter two are critical and inseparable part of the detector structure. $\square$

Note that (24) is general enough to describe the detector structures for arbitrary data format and/or modulation schemes. For spatial multiplexing (SMX) signals, the search range is modified to $\mathcal{A}_M^{N_T \times B}$, while a precoded MIMO system, we replace $\mathbf{X}$ by $\mathbf{WS}$ with $\mathbf{W}$ being the precoding matrix. We derive specific detector structures for various SM signals in the remaining part of this section and in Sections IV and V.

B. CEEA-ML DETECTORS FOR SM SIGNALS
1) M-PSK CONSTELLATION
With the decomposition $\mathbf{X} = \mathbf{LS}$ defined in Section II-B and $\mathcal{A}_M$ an M-PSK constellation, the *CEEA-ML detector* for the PSK-based SM system is derivable from (24) and is given by

$$\hat{\mathbf{X}}_{mb}^{\text{ML}}(k) = \arg \widetilde{\min_{(\mathbf{s},\mathbf{L}) \in \mathcal{A}_M^B \times \mathcal{L}}} \log \det \bar{\mathbf{C}}_{psk}(k)$$
$$+ \mathcal{G}(\bar{\mathbf{C}}_{psk}(k), \mathbf{y}_s(k) - \bar{\mathbf{m}}_{ssk}(k)) \quad (27)$$

where $\mathcal{L}$ denotes the set of all SSK matrices of the form (8) $\bar{\mathbf{m}}_{ssk}(k)$ and $\bar{\mathbf{C}}_{psk}(k)$, which are not exactly the conditional mean and covariance, are obtained from (23a) and (23b) by substituting $\mathbf{L}$ for $\mathbf{X}$ and $\frac{\sigma_z^2}{\varepsilon_s}$, $\varepsilon_s = |s_j|^2$, for σ_z^2. In (27), we have defined, for an implicit function f of $\mathbf{s}$ and $\mathbf{L}$,

$$\arg \widetilde{\min_{(\mathbf{s},\mathbf{L})}} f(\mathbf{s},\mathbf{L}) = \hat{\mathbf{L}} \,\text{Diag}(\hat{\mathbf{s}}), \quad (\hat{\mathbf{s}}, \hat{\mathbf{L}}) = \arg \min_{(\mathbf{s},\mathbf{L})} f(\mathbf{s},\mathbf{L}),$$
$$(28)$$

where $\mathbf{y}_s(k) \overset{def}{=} \text{vec}\left(\mathbf{Y}(k)\mathbf{S}^H \right) / \varepsilon_s$.

2) M-QAM CONSTELLATION
If $\mathcal{A}_M$ is an M-QAM constellation, the corresponding *CEEA-ML detector* is

$$\hat{\mathbf{X}}_{mb}^{\text{ML}}(k) = \arg \widetilde{\min_{(\mathbf{s},\mathbf{L}) \in \mathcal{A}_M^B \times \mathcal{L}}} N_R \log \det \mathbf{E}_s + \log \det \bar{\mathbf{C}}_{qam}(k)$$
$$+ \mathcal{G}(\bar{\mathbf{C}}_{qam}(k), \mathbf{y}_s(k) - \bar{\mathbf{m}}_{ssk}(k)) \quad (29)$$

with $\mathbf{E}_s \overset{def}{=} \mathbf{SS}^H$, $\mathbf{y}_s(k) \overset{def}{=} \text{vec}\left(\mathbf{Y}(k)\mathbf{S}^H \mathbf{E}_s^{-1} \right)$, and

$$\bar{\mathbf{C}}_{qam}(k) \overset{def}{=} \sigma_z^2(\mathbf{E}_s^{-1} \otimes \mathbf{I}_{N_R}) + (\mathbf{L}^T \otimes \mathbf{I}_{N_R})\Big[\mathbf{I}_{N_R N_T} - \mathbf{A}(k)$$
$$\times \left(\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k) \right)^* \Big] \Phi(\mathbf{L}^* \otimes \mathbf{I}_{N_R}). \quad (30)$$

C. TWO-STAGE M-PSK SM DETECTOR
The ML detector (27) calls for an exhaustive search over the set of all candidate antenna index-modulated symbol pairs, which has a cardinality of $|\mathbb{L} \times \mathcal{A}_M|^B$. As mentioned before, low-complexity alternatives for rectangular QAM [13] and PSK [14] with perfect CSI have been proposed but similar reduction method is not applicable for our CEEA receivers. Instead, we consider a two-stage approach that detects the active antenna indices and then the transmitted symbols in each block. A similar approach with perfect CSIR assumption have been suggested [1].

We first notice that

$$P\Big[\mathbf{Y}(k)\Big|\mathbf{L}(k), \hat{\mathbf{H}}(k)\Big]$$
$$= \sum_{s_1(k)}\sum_{s_2(k)}\cdots\sum_{s_B(k)} P\Big[\mathbf{Y}(k)\Big|\mathbf{L}(k),\mathbf{S}(k),\hat{\mathbf{H}}(k)\Big]P\Big[\mathbf{S}(k)\Big]$$

where $\mathbf{L}(k)$ and $\mathbf{S}(k)$ are the kth block's SSK and symbol matrices. It follows that the estimate of the activated antenna indices is

$$\hat{\mathbf{L}}(k) = \arg \max_{\mathbf{L} \in \mathcal{L}} \frac{1}{M^B \sqrt{\det \varepsilon_s \bar{\mathbf{C}}_{psk}(k)}} \sum_{\mathbf{s} \in \mathcal{A}_M^B} \mathcal{F}(\mathbf{s}) \quad (31)$$

where

$$\mathcal{F}(\mathbf{s}) \overset{def}{=} \exp\left[\frac{-\mathcal{G}(\bar{\mathbf{C}}_{psk}(k), \bar{\mathbf{m}}_{ssk}(k))}{2} - \frac{\mathbf{s}^T \mathbf{J}(k)\mathbf{s}^*}{2\varepsilon_s^2} \right.$$
$$\left. + \frac{\Re\{\mathbf{b}^H(k)\mathbf{s}^*\}}{\varepsilon_s} \right], \quad (32a)$$
$$\mathbf{J}(k) = \mathcal{G}\left(\bar{\mathbf{C}}_{psk}(k), \text{DIAG}(\mathbf{y}_1(k),\cdots,\mathbf{y}_B(k)) \right), \quad (32b)$$
$$\mathbf{b}(k) = \text{DIAG}(\mathbf{y}_1(k),\cdots,\mathbf{y}_B(k))^H \bar{\mathbf{C}}_{psk}^{-1}(k)\bar{\mathbf{m}}_{ssk}(k). \quad (32c)$$

For each candidate SSK matrix $\mathbf{L}$, we have the approximation

$$\sum_{\mathbf{s}} \mathcal{F}(\mathbf{s}) \approx \mathcal{F}(\bar{\mathbf{s}}(\mathbf{L})), \quad \bar{\mathbf{s}}(\mathbf{L}) \overset{def}{=} \mathcal{Q}_{\mathcal{A}_M}\left(\tilde{\mathbf{s}}(\mathbf{L}) \right), \quad (33)$$

where $\mathcal{Q}_{\mathcal{A}_M}(\cdot)$ quantizing the enclosed items to the nearest constellation points in $\mathcal{A}_M$ and $\tilde{\mathbf{s}}(\mathbf{L}) = \varepsilon_s(\mathbf{J}^{-1}(k)\mathbf{b}(k))^*$ being the solution of $\partial\mathcal{F}(\mathbf{s})/\partial\mathbf{s} = \mathbf{0}_B$. $\mathbf{b}(k)$ defined in (32c) is an implicit function of $\mathbf{L}$ since as mentioned in the previous

subsection, $\bar{\mathbf{C}}_{psk}(k)$ and $\bar{\mathbf{m}}_{ssk}(k)$ are both functions of $\mathbf{L}$. The approximation (33) thus depends on $\mathbf{L}$ and the associated tentative demodulation decision $\bar{\mathbf{s}}(\mathbf{L})$. A simpler antenna indices estimate is then given by

$$\hat{\mathbf{L}}(k) \approx \arg\min_{\mathbf{L}\in\mathcal{L}} \log\det(\varepsilon_s \bar{\mathbf{C}}_{psk}(k)) + \mathcal{G}(\bar{\mathbf{C}}_{psk}(k), \bar{\mathbf{m}}_{ssk}(k))$$
$$+ \frac{\bar{\mathbf{s}}^T(\mathbf{L})\mathbf{J}(k)\bar{\mathbf{s}}^*(\mathbf{L})}{\varepsilon_s^2} - \frac{2\Re\left\{\mathbf{b}^H(k)\bar{\mathbf{s}}^*(\mathbf{L})\right\}}{\varepsilon_s}. \quad (34a)$$

Once the active SSK matrix is determined, we output the decision $\bar{\mathbf{s}}(\mathbf{L})$ associated with $\hat{\mathbf{L}}(k)$ which has been determined in the previous stage

$$\hat{\mathbf{s}}(k) = \bar{\mathbf{s}}(\hat{\mathbf{L}}(k)). \quad (34b)$$

(34b) will reappear later repeatedly, each with different antenna index detection rule $\hat{\mathbf{L}}(k)$. Obviously, the search complexity of the detector (34a) is much lower than that of (27). For systems employing QAM constellations, the approximation (33) is not directly applicable as the nonconstant-modulus nature of QAM implies that $\mathcal{F}(\mathbf{s})$ has an additional amplitude-dependent term in the exponent.

IV. DD CE-AIDED CEEA-ML DETECTORS
A. GENERAL MIMO SIGNAL DETECTION WITH IMPERFECT CSI
To derive the *CEEA-ML detector* for general MIMO signals using a DD LS channel estimate, we again appeal to *Lemma 1* with $\mathbf{z}_1$ defined by (20a) and $\mathbf{z}_2 \overset{def}{=} \text{vec}(\hat{\mathbf{H}}(k-1)) = \text{vec}(\mathbf{Y}(k-1)\hat{\mathbf{X}}^\dagger(k-1))$. For $k_p < k < k_p + N$, the signal block which maximizes the likelihood function $P\left[\text{vec}(\mathbf{Y}(k))\middle|\text{vec}(\mathbf{X}(k)), \text{vec}(\hat{\mathbf{H}}(k-1))\right]$ is given by

$$\hat{\mathbf{X}}_{dd}^{\text{ML}}(k) = \arg\min_{\mathbf{X}\in\mathcal{A}_{M+}^{N_T\times B}} \log\det\mathbf{C}_{dd}$$
$$+ \mathcal{G}(\mathbf{C}_{dd}, \text{vec}(\mathbf{Y}(k)) - \mathbf{m}_{dd}). \quad (35)$$

To have more compact expressions we use $\hat{\mathbf{H}}$ and $\hat{\mathbf{H}}(k-1)$ interchangeably for the DD *CEEA-ML detectors*. This is justifiable as the conditional mean and covariance of $\mathbf{Y}(k)$ given $\mathbf{X}(k)$ and $\hat{\mathbf{H}}(k-1)$ are

$$\mathbf{m}_{dd} = \left(\mathbf{X}^T(k)\otimes\mathbf{I}_{N_R}\right)\mathbf{A}\text{vec}(\hat{\mathbf{H}}(k-1)), \quad (36a)$$
$$\mathbf{C}_{dd} = \sigma_z^2\mathbf{I}_{BN_R} + \left(\mathbf{X}^T(k)\otimes\mathbf{I}_{N_R}\right)\left[\mathbf{I}_{N_RN_T} - \rho_T(1)\mathbf{A}\right]$$
$$\cdot\Phi\left(\mathbf{X}^*(k)\otimes\mathbf{I}_{N_R}\right) \quad (36b)$$

with $\mathbf{A}\overset{def}{=}\rho_T(1)\,\Phi\,(\Phi + \sigma_z^2\mathbf{I}_{N_RN_T})^{-1}$.

B. CEEA-ML SM SIGNAL DETECTORS
Following the derivation of Section III-B and using the decomposition $\mathbf{X} = \mathbf{L}\mathbf{S}$, we summarize below the resulting DD *CEEA-ML* detectors using PSK or QAM signal.

1) M-PSK CONSTELLATION
For a PSK-SM MIMO system, the *CEEA-ML detector* is

$$\hat{\mathbf{X}}_{dd}^{\text{ML}}(k) = \arg\widetilde{\min_{s_j\in\mathcal{A}_M,\,\ell_j\in\mathbb{L}}}\ \log\det\tilde{\mathbf{C}}_{psk}$$
$$+ \mathcal{G}(\tilde{\mathbf{C}}_{psk}, \mathbf{y}_s(k) - \tilde{\mathbf{m}}_{ssk}) \quad (37)$$

where $\tilde{\mathbf{m}}_{ssk} = (\mathbf{L}^T\otimes\mathbf{I}_{N_R})\mathbf{A}\text{vec}(\hat{\mathbf{H}})$, $\tilde{\mathbf{C}}_{psk}\overset{def}{=}\frac{\sigma_z^2}{\varepsilon_s}\mathbf{I}_{BN_R}+ (\mathbf{L}^T\otimes\mathbf{I}_{N_R})\left[\mathbf{I}_{N_RN_T} - \rho_T(1)\mathbf{A}\right]\Phi(\mathbf{L}^*\otimes\mathbf{I}_{N_R})$, and $\mathbf{y}_s(k) = \text{vec}\left(\mathbf{Y}(k)\mathbf{S}^H\right)/\varepsilon_s$.

2) M-QAM CONSTELLATION
When M-QAM is used, (35) reduces to

$$\hat{\mathbf{X}}_{dd}^{\text{ML}}(k) = \arg\widetilde{\min_{s_j\in\mathcal{A}_M,\,\ell_j\in\mathbb{L}}}\ N_R\log\det\mathbf{E}_s + \log\det\tilde{\mathbf{C}}_{qam}$$
$$+ \mathcal{G}(\tilde{\mathbf{C}}_{qam}, \mathbf{y}_s(k) - \tilde{\mathbf{m}}_{ssk}) \quad (38)$$

where $\tilde{\mathbf{C}}_{qam}\overset{def}{=}(\mathbf{L}^T\otimes\mathbf{I}_{N_R})\left[\mathbf{I}_{N_RN_T} - \rho_T(1)\mathbf{A}\right]\Phi(\mathbf{L}^*\otimes\mathbf{I}_{N_R}) + \sigma_z^2(\mathbf{E}_s^{-1}\otimes\mathbf{I}_{N_R})$, $\mathbf{y}_s(k)\overset{def}{=}\text{vec}\left(\mathbf{Y}(k)\mathbf{S}^H\mathbf{E}_s^{-1}\right)$, and $\mathbf{E}_s\overset{def}{=}\mathbf{S}\mathbf{S}^H$.

Remark 3: Compared with the MB-CE-aided detectors (cf. (27) and (29)), (37) and (38) require much less storage. The MB channel estimation is performed every two frames and thus $2N - 2$ estimated channel matrices have to be updated and saved for subsequent signal detection. The associated $\mathbf{A}(k)$'s can be precalculated but they have to be stored as well. The DD-CE aided detectors, on the other hand, need to store two matrices, $\mathbf{A}$ and $\hat{\mathbf{H}}$, only. However, as shown in Section VIII, the latter suffers from inferior performance. More detailed memory requirement comparison is provided in Section VII. □

C. TWO-STAGE DETECTOR FOR M-PSK SM SIGNALS
We can reduce the search range of (37) from the Bth Cartesian power of $\mathbb{L}\times\mathcal{A}_M$ to $\mathbb{L}^B$ by adopting the two-stage approach of (34) that detects the antenna indices and then the transmitted symbols. The associated detector is derived from maximizing the likelihood function

$$P\left[\mathbf{Y}(k)\middle|\mathbf{X}(k), \hat{\mathbf{H}}(k-1)\right]$$

with respect to (w.r.t.)$\mathbf{X}(k)$:

$$\max_{\mathbf{L}\in\mathcal{L},\,\mathbf{s}\in\mathcal{A}_M^B} (\det\varepsilon_s\tilde{\mathbf{C}}_{psk})^{-\frac{1}{2}}\exp\left[\frac{-\mathcal{G}(\tilde{\mathbf{C}}_{psk},\tilde{\mathbf{m}}_{ssk})}{2}\right.$$
$$\left.-\frac{\mathbf{s}^T\mathbf{J}(k)\mathbf{s}^*}{2\varepsilon_s^2} + \frac{\Re\{\mathbf{b}^H(k)\mathbf{s}^*\}}{\varepsilon_s}\right]$$
$$\approx \max_{\mathbf{L}\in\mathcal{L}} (\det\varepsilon_s\tilde{\mathbf{C}}_{psk})^{-\frac{1}{2}}\exp\left[\frac{-\mathcal{G}(\tilde{\mathbf{C}}_{psk},\tilde{\mathbf{m}}_{ssk})}{2}\right]$$
$$\cdot\max_{\mathbf{s}\in\mathcal{A}_M^B}\exp\left[-\frac{1}{2\varepsilon_s^2}\mathbf{s}^T\mathbf{J}(k)\mathbf{s}^* + \frac{1}{\varepsilon_s}\Re\{\mathbf{b}^H(k)\mathbf{s}^*\}\right]$$

$$= \max_{\mathbf{L} \in \mathcal{L}} \left(\det \varepsilon_s \tilde{\mathbf{C}}_{psk} \right)^{-\frac{1}{2}} \exp \left[\frac{-\mathcal{G}(\tilde{\mathbf{C}}_{psk}, \tilde{\mathbf{m}}_{ssk})}{2} \right.$$

$$\left. - \frac{\bar{\mathbf{s}}^T(\mathbf{L})\mathbf{J}(k)\bar{\mathbf{s}}^*(\mathbf{L})}{2\varepsilon_s^2} + \frac{\Re\{\mathbf{b}^H(k)\bar{\mathbf{s}}^*(\mathbf{L})\}}{\varepsilon_s} \right], \quad (39)$$

where $\mathbf{J}(k)$, $\mathbf{b}(k)$, and $\bar{\mathbf{s}}(\mathbf{L})$ are defined similarly to those in (32b), (32c) and (33) except that now $\hat{\mathbf{H}} = \hat{\mathbf{H}}(k-1)$. Replacing the likelihood function by its logarithm version, we obtain a *two-stage detector* similar to (34)

$$\hat{\mathbf{L}}(k) = \arg\min_{\mathbf{L} \in \mathcal{L}} \log \det(\varepsilon_s \tilde{\mathbf{C}}_{psk}) + \mathcal{G}(\tilde{\mathbf{C}}_{psk}, \tilde{\mathbf{m}}_{ssk})$$

$$+ \frac{\bar{\mathbf{s}}^T(\mathbf{L})\mathbf{J}(k)\bar{\mathbf{s}}^*(\mathbf{L})}{\varepsilon_s^2} - \frac{2\Re\{\mathbf{b}^H(k)\bar{\mathbf{s}}^*(\mathbf{L})\}}{\varepsilon_s}, \quad (40a)$$

$$\hat{\mathbf{s}}(k) = \bar{\mathbf{s}}(\hat{\mathbf{L}}(k)). \quad (40b)$$

V. ML DETECTION WITHOUT SPATIAL CORRELATION INFORMATION AT EITHER SIDE

The detector structures presented so far have assumed complete knowledge of the channel's spatial correlation Φ. As pointed out in [16], when both sides of a link are richly scattered, the corresponding spatial statistics can be assumed separable, yielding the Kronecker spatial channel model

$$\mathbf{H}(k) = \Phi_R^{\frac{1}{2}} \mathbf{H}_w(k) \Phi_T^{\frac{1}{2}} \quad (41)$$

with the spatial correlation matrix Φ given by [16]

$$\Phi = \Phi_T \otimes \Phi_R, \quad (42)$$

the Kronecker product of the spatial correlation matrix at the transmit side $\Phi_T = \mathbb{E}\{\mathbf{H}^T(k)\mathbf{H}^*(k)\}/\text{tr}(\Phi_T)$ and that at the receive side $\Phi_R = \mathbb{E}\{\mathbf{H}(k)\mathbf{H}^H(k)\}/\text{tr}(\Phi_R)$. When the latter is not available, we assume that $\Phi_R = \mathbf{I}_{N_R}$ hence $\mathbf{H}(k) = \mathbf{H}_w(k)\Phi_T^{\frac{1}{2}}$. The assumption of uncorrelated receive antennas also implies

$$P\left[\mathbf{Y}(k)|\mathbf{X}(k), \hat{\mathbf{H}}(k)\right] = \prod_{n=1}^{N_R} P\left[\underline{\mathbf{y}}_n(k)|\mathbf{X}(k), \hat{\underline{\mathbf{h}}}_n(k)\right], \quad (43)$$

where $\underline{\mathbf{y}}_n(k)$ is the sample (row) vector received by antenna n at block k and $\hat{\underline{\mathbf{h}}}_n(k)$ the estimated channel vector between the nth receive antenna and transmit antennas. As has been mention in Section I, we refer to a detector based on (43) as the *ZRC detector*. An expression similar to (43) for the case of uncorrelated transmit antennas can be used to derive the ZTC detector.

A. MB-CE-AIDED ZRC AND ZTC DETECTORS
1) GENERAL ZRC/ZTC DETECTORS
Define

$$\mathbf{z}_1^H = \underline{\mathbf{y}}_n(k) = \underline{\mathbf{h}}_n(k)\mathbf{X}(k) + \underline{\mathbf{z}}_n(k),$$

$$\mathbf{z}_2^H = \hat{\underline{\mathbf{h}}}_n(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)$$

$$\times \left[\tilde{\mathbf{y}}_{n1}(k_p), \tilde{\mathbf{y}}_{n2}(k_p), \cdots, \tilde{\mathbf{y}}_{nN_T}(k_p)\right]$$

where $\mathbf{Z}(k) \overset{def}{=} \left[\underline{\mathbf{z}}_1^T(k), \cdots, \underline{\mathbf{z}}_{N_R}^T(k)\right]^T$. We immediately obtain the mean and covariance of $\underline{\mathbf{y}}_n(k)$ conditioned on $\mathbf{X}(k)$ and $\hat{\underline{\mathbf{h}}}_n(k)$ as

$$\bar{\mathbf{m}}_n^T(k) = \hat{\underline{\mathbf{h}}}_n(k)\mathbf{A}_{zrc}(k)\mathbf{X}(k), \quad (44a)$$

$$\bar{\mathbf{C}}_{zrc}(k) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^H$$

$$= \sigma_z^2 \mathbf{I}_{N_T} + \mathbf{X}^H(k)\left[\mathbf{I}_{N_T} - \mathbf{A}(k)\left(\mathbf{t}^H(k)\mathbf{T}^{-1}\right.\right.$$

$$\left.\left. \times (k_p)\mathbf{q}(k)\right)^*\right] \times \Phi_T \mathbf{X}(k), \quad (44b)$$

where $\mathbf{A}_{zrc}(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\Phi_T \Sigma_{22}^{-1}$,

$$\Sigma_{11} = \mathbf{X}^H(k)\Phi_T \mathbf{X}(k) + \sigma_z^2 \mathbf{I}_{N_T},$$

$$\Sigma_{12} = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\mathbf{X}^H(k)\Phi_T,$$

$$\Sigma_{22} = \nu(k)\Phi_T + \sigma_z^2 \|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\|_F^2 \mathbf{I}_{N_T}$$

with $\mathbf{q}(k)$ and $\nu(k)$ being defined in (22).

The resulting *ZRC detector* is thus given by

$$\hat{\mathbf{X}}_{mb}^{\text{ZRC}}(k) = \arg\min_{\mathbf{X} \in \mathcal{X}} N_R \log \det \bar{\mathbf{C}}_{zrc}(k)$$

$$+ \text{tr}\left\{\mathcal{G}\left(\bar{\mathbf{C}}_{zrc}(k), \left(\mathbf{Y}(k) - \bar{\mathbf{M}}_{zrc}(k)\right)^H\right)\right\}, \quad (45)$$

where $\bar{\mathbf{M}}_{zrc}(k) = [\bar{\mathbf{m}}_1(k), \cdots, \bar{\mathbf{m}}_{n_R}(k)]^T = \hat{\mathbf{H}}(k)\mathbf{A}_{zrc}(k)\mathbf{X}(k)$.

(45) can also be derived by substituting $\Phi_R = \mathbf{I}_{N_R}$ into the *CEEA-ML* detection rule (24) and applying (42) with some algebraic manipulations.

On the other hand, when the spatial correlation at the transmit side is not available, we assume no *a priori* transmit spatial correlation, $\Phi_T = \mathbf{I}_{N_T}$ whence $\mathbb{E}\{\mathbf{H}(k)\mathbf{H}^H(k)\}/\text{tr}(\Phi_R) = \Phi_R$. Substituting $\Phi_T = \mathbf{I}_{N_T}$ into (42) we obtain the *ZTC* detector for a generic MB-CE-aided MIMO system with

$$\mathbf{m}_{mb}(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\text{vec}\left[\Phi_R\left(\nu(k)\Phi_R\right.\right.$$

$$\left.\left. + \sigma_z^2 \left\|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right\|_F^2 \mathbf{I}_{N_R}\right)^{-1} \hat{\mathbf{H}}(k)\mathbf{X}(k)\right], \quad (46)$$

$$\mathbf{C}_{mb}(k) = \sigma_z^2 \mathbf{I}_{N_R N_T} + \left[\mathbf{X}^T(k)\mathbf{X}^*(k)\right] \otimes \Phi_R\left[\mathbf{I}_{N_R}\right.$$

$$- |\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)|^2 \left(\nu(k)\mathbf{I}_{N_R}\right.$$

$$+ \sigma_z^2 \left\|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right\|_F^2 \Phi_R^{-1}\right)^{-1}\right]$$

$$\approx \mathbf{I}_{N_T} \otimes \bar{\mathbf{C}}_{ztc}(k), \quad (47)$$

where we use $\mathbf{X}(k)\mathbf{X}^H(k) \approx \varepsilon_s \mathbf{I}_{N_T}$ and define

$$\bar{\mathbf{C}}_{ztc}(k) \overset{def}{=} \sigma_z^2 \mathbf{I}_{N_R} + \varepsilon_s \Phi_R\left[\mathbf{I}_{N_R} - |\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)|^2\right.$$

$$\left. \times \Phi_R\left(\sigma_z^2 \left\|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right\|_F^2 \mathbf{I}_{N_R} + \nu(k)\Phi_R\right)^{-1}\right]. \quad (48)$$

to obtain (47). As a result, $\det \mathbf{C}_{ztc}(k) = \left(\det \bar{\mathbf{C}}_{ztc}(k)\right)^{N_T}$ and

$$\mathcal{G}(\mathbf{C}_{mb}(k), \text{vec}(\mathbf{Y}(k)) - \mathbf{m}_{mb}(k)) = \text{tr}\left\{\mathcal{G}(\bar{\mathbf{C}}_{ztc}(k), \mathbf{Y}(k)\right.$$

$$\left. - \bar{\mathbf{M}}_{ztc}(k))\right\},$$

where

$$\bar{\mathbf{M}}_{ztc}(k) = \mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k) \cdot \Phi_R\Big(\nu(k)\Phi_R$$
$$+ \sigma_z^2 \left\|\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\right\|_F^2 \mathbf{I}_{N_R}\Big)^{-1}\hat{\mathbf{H}}(k)\mathbf{X}(k). \quad (49)$$

The resulting *ZTC detector* is given by

$$\hat{\mathbf{X}}_{mb}^{ZTC}(k) = \arg \min_{\mathbf{X}\in\mathcal{A}_{M+}^{N_T\times B}} \mathrm{tr}\left\{\mathcal{G}(\bar{\mathbf{C}}_{ztc}(k), \mathbf{Y}(k) - \bar{\mathbf{M}}_{ztc}(k))\right\}$$
$$+ N_T \log\det\bar{\mathbf{C}}_{ztc}(k) \quad (50)$$

whose complexity is similar to that of (45).

In case both the transmit and receive correlations are unavailable, the conditional covariance matrix in the detection metric degenerates to an identity matrix scaled by a factor that is a function of the noise variance, channel's temporal correlation and block index k. Both *ZRC* and *ZTC* receivers can be further simplified by the two-stage approach by following the derivation given in Section III-B, we focus on deriving reduced complexity *ZRC detectors* only; the *ZTC* counterparts can be similarly obtained.

2) TWO-STAGE ZRC M-PSK DETECTOR

Using the decomposition, $\mathbf{X} = \mathbf{LS}$, we express the *ZRC* receiver (45) for an M-PSK SM system as

$$\hat{\mathbf{X}}_{mb}^{ZRC}(k) = \arg \widetilde{\min_{(\mathbf{s},\mathbf{L})\in\mathcal{A}_M^B\times\mathcal{L}}} N_R \log\det(\varepsilon_s\bar{\mathbf{C}}_{ssk}(k))$$
$$+ \mathrm{tr}\left\{\mathcal{G}\Big(\bar{\mathbf{C}}_{ssk}(k), \big(\mathbf{Y}_s(k) - \bar{\mathbf{M}}_{ssk}(k)\big)^H\Big)\right\} \quad (51)$$

where $\mathbf{Y}_s(k) = \mathbf{Y}(k)\mathbf{S}^H/\varepsilon_s$, $\bar{\mathbf{M}}_{ssk}(k) = \hat{\mathbf{H}}(k)\mathbf{A}(k)\mathbf{L}$, and

$$\bar{\mathbf{C}}_{ssk}(k) = \frac{\sigma_z^2}{\varepsilon_s}\mathbf{I}_{N_T} + \mathbf{L}^H$$
$$\times \left[\mathbf{I}_{N_T} - \mathbf{A}(k)\mathbf{t}^H(k)\mathbf{T}^{-1}(k_p)\mathbf{q}(k)\right]\Phi_T\mathbf{L}. \quad (52)$$

We can show that the decision rule for separate antenna index and modulated symbol detection is given by

$$\hat{\mathbf{L}}(k) = \arg \min_{\mathbf{L}\in\mathcal{L}} N_R \log\det(\varepsilon_s\bar{\mathbf{C}}_{ssk}(k))$$
$$+ \mathrm{tr}\left\{\mathcal{G}\Big(\bar{\mathbf{C}}_{ssk}(k), \bar{\mathbf{M}}_{ssk}(k)^H\Big)\right\}$$
$$+ \frac{\bar{\mathbf{s}}^H(\mathbf{L})\mathbf{J}(k)\bar{\mathbf{s}}(\mathbf{L})}{\varepsilon_s^2} - \frac{2\Re\{\mathbf{b}^T(k)\bar{\mathbf{s}}(\mathbf{L})\}}{\varepsilon_s}, \quad (53a)$$
$$\hat{\mathbf{s}}(k) = \bar{\mathbf{s}}(\hat{\mathbf{L}}(k)), \quad (53b)$$

where the entries of $\mathbf{b}(k)$ are the diagonal terms of $\mathbf{Y}^H(k)\bar{\mathbf{M}}_{ssk}\bar{\mathbf{C}}_{ssk}^{-1}(k)$, $\bar{\mathbf{s}}(\mathbf{L}) = \mathcal{Q}_{\mathcal{A}_M}\left(\varepsilon_s(\mathbf{b}^T(k)\mathbf{J}^{-1}(k))^H\right)$, and $\mathbf{J}(k) = \bar{\mathbf{C}}_{ssk}^{-1}(k)\odot(\mathbf{Y}^H(k)\mathbf{Y}(k))^*$.

B. DD-CE-AIDED ZRC AND ZTC DETECTORS
1) GENERAL ZRC/ZTC DETECTORS

Based on the ZRC assumption and the fact that in a DD system $\mathbf{X}(k)$ is detected with the DD CEs obtained at block $k-1$, we follow the procedure presented in the previous

subsection with $\mathbf{z}_1^H = \underline{\mathbf{y}}_n(k)$ and $\mathbf{z}_2^H = \hat{\underline{\mathbf{h}}}_n \stackrel{def}{=} \hat{\underline{\mathbf{h}}}_n(k-1) = \underline{\mathbf{h}}_n(k-1)\mathbf{G}_1(k-1) + \mathbf{z}_n(k-1)\mathbf{G}_2(k-1)$, where $\mathbf{G}_1(k) \stackrel{def}{=} \mathbf{X}(k)\hat{\mathbf{X}}^\dagger(k)$ and $\mathbf{G}_2(k) \stackrel{def}{=} \mathbf{X}^\dagger(k)\mathbf{G}_1(k)$, to obtain the covariance matrices

$$\Sigma_{11} = \mathbf{X}^H(k)\Phi_T\mathbf{X}(k) + \sigma_z^2\mathbf{I}_B, \quad (54a)$$
$$\Sigma_{12} = \rho_T(1)\mathbf{X}^H(k)\Phi_T\mathbf{G}_1(k-1), \quad (54b)$$
$$\Sigma_{22} = \mathbf{G}_1^H(k-1)\left[\Phi_T + \sigma_z^2\left(\mathbf{X}(k-1)\mathbf{X}^H(k-1)\right)^{-1}\right]$$
$$\times\mathbf{G}_1(k-1)$$
$$\approx \mathbf{G}_1^H(k-1)\left(\Phi_T + \frac{\sigma_z^2}{\varepsilon_s}\mathbf{I}_B\right)\mathbf{G}_1(k-1). \quad (54c)$$

It follows immediately that, given $\mathbf{X}(k)$ and $\hat{\underline{\mathbf{h}}}_n(k-1)$, $\underline{\mathbf{y}}_n(k)$ has mean $\mathbf{z}_2^H\Sigma_{22}^{-1}\Sigma_{12}^H = \rho_T(1)\hat{\underline{\mathbf{h}}}_n(\Phi_T + \sigma_z^2/\varepsilon_s\mathbf{I}_B)^{-1}\Phi_T\mathbf{X}(k)$ and covariance matrix

$$\tilde{\mathbf{C}}_{zrc} \stackrel{def}{=} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^H$$
$$= \sigma_z^2\mathbf{I}_B + \mathbf{X}^H(k)\left[\mathbf{I}_{N_T} - \rho_T^2(1)\left(\mathbf{I}_{N_T} + \frac{\sigma_z^2}{\varepsilon_s}\Phi_T^{-1}\right)^{-1}\right]$$
$$\times\Phi_T\mathbf{X}(k) \quad (55)$$

where $\rho_T(1)$ is a prediction term used to alleviate the error propagation effect. The resulting *ZRC detector* is

$$\hat{\mathbf{X}}_{dd}^{ZRC}(k) = \arg \min_{\mathbf{X}\in\mathcal{X}} N_R \log\det\tilde{\mathbf{C}}_{zrc}$$
$$+ \mathrm{tr}\left[\mathcal{G}\left(\tilde{\mathbf{C}}_{zrc}, \mathbf{Y}^H(k) - \tilde{\mathbf{M}}_{zrc}^H\right)\right], \quad (56)$$

where $\tilde{\mathbf{M}}_{zrc} = \rho_T(1)\hat{\mathbf{H}}\left(\mathbf{I}_{N_T} + \sigma_z^2/\varepsilon_s\Phi_T^{-1}\right)^{-1}\mathbf{X}(k)$. This *ZRC detector* can also be derived from the *CEEA-ML detector* (35) by using $\Phi_R = \mathbf{I}_{N_R}$ and some algebraic manipulations.

To derive the *ZTC detector* for a DD-CE-aided MIMO system, we replace Φ by $\mathbf{I}_{N_T}\otimes\Phi_R$ in (36) and invoke

$$\mathbf{A} = \rho_T(1)(\mathbf{I}_{N_T}\otimes\Phi_R)(\mathbf{I}_{N_T}\otimes\Phi_R + \sigma_z^2\mathbf{I}_{N_RN_T})^{-1}$$
$$= \rho_T(1)\,\mathbf{I}_{N_T}\otimes\left(\Phi_R(\Phi_R + \sigma_z^2\mathbf{I}_{N_R})^{-1}\right) \quad (57a)$$

and $\mathbf{X}(k)\mathbf{X}^H(k) \approx \varepsilon_s\mathbf{I}_{N_T}$ to obtain

$$\mathbf{m}_{dd} = \mathrm{vec}\left(\rho_T(1)\left(\Phi_R(\Phi_R + \sigma_z^2\mathbf{I}_{N_R})^{-1}\right)\hat{\mathbf{H}}\mathbf{X}(k)\right)$$
$$\stackrel{def}{=} \mathrm{vec}(\tilde{\mathbf{M}}_{ztc}), \quad (57b)$$
$$\mathbf{C}_{dd} \approx \mathbf{I}_{N_T}\otimes\left(\sigma_z^2\mathbf{I}_{N_R}\right.$$
$$+ \varepsilon_s\left(\Phi_R - \rho_T^2(1)(\mathbf{I}_{N_R} + \sigma_z^2\Phi_R^{-1})^{-1}\Phi_R\right)\Big)$$
$$\stackrel{def}{=} \mathbf{I}_{N_T}\otimes\tilde{\mathbf{C}}_{ztc}. \quad (57c)$$

Thus, $\det\mathbf{C}_{dd} = (\det\mathbf{I}_{N_T})^{N_R}\cdot(\det\tilde{\mathbf{C}}_{ztc})^{N_T}$ and

$$\mathcal{G}(\mathbf{C}_{dd}, \mathrm{vec}(\mathbf{Y}(k)) - \mathbf{m}_{dd}) = \mathrm{tr}\left\{\mathcal{G}(\tilde{\mathbf{C}}_{ztc}, \mathbf{Y}(k) - \tilde{\mathbf{M}}_{ztc})\right\}.$$

and from (35) we have the *DD-ZTC detector*

$$\hat{\mathbf{X}}_{dd}^{ZTC}(k) = \arg \min_{\mathbf{X}\in\mathcal{A}_{M+}^{N_T\times B}} \mathrm{tr}\left\{\mathcal{G}(\tilde{\mathbf{C}}_{ztc}, \mathbf{Y}(k) - \tilde{\mathbf{M}}_{ztc})\right\}$$
$$+ N_T \log\det\tilde{\mathbf{C}}_{ztc} \quad (58)$$

2) TWO-STAGE ZRC *M*-PSK DETECTOR

We first define

$$\tilde{\mathbf{M}}_{ssk} = \rho_T(1)\hat{\mathbf{H}}\left(\mathbf{I}_{N_T} + \frac{\sigma_z^2}{\varepsilon_s}\Phi_T^{-1}\right)^{-1}\mathbf{L}, \tag{59a}$$

$$\tilde{\mathbf{C}}_{ssk} = \frac{\sigma_z^2}{\varepsilon_s}\mathbf{I}_B + \mathbf{L}^H\left[\mathbf{I}_{N_T} - \rho_T^2(1)\left(\mathbf{I}_{N_T} + \frac{\sigma_z^2}{\varepsilon_s}\Phi_T^{-1}\right)^{-1}\right] \\ \cdot \Phi_T\mathbf{L}. \tag{59b}$$

Using the above definitions and the decomposition $\mathbf{X} = \mathbf{LS}$, we obtain an alternate expression for (55) and rewrite (56) as

$$\hat{\mathbf{X}}_{dd}^{\text{ZRC}}(k) = \arg \widetilde{\min_{(\mathbf{s},\mathbf{L})\in\mathcal{A}_M^B\times\mathcal{L}}} N_R \log \det(\varepsilon_s\tilde{\mathbf{C}}_{ssk}) \\ + \text{tr}\left\{\mathcal{G}\left(\tilde{\mathbf{C}}_{ssk}, \left(\mathbf{Y}_s(k) - \tilde{\mathbf{M}}_{ssk}\right)^H\right)\right\}. \tag{60}$$

The corresponding *two-stage ZRC detector* can then be derived as

$$\hat{\mathbf{L}}(k) = \arg \min_{\mathbf{L}\in\mathcal{L}} N_R \log \det(\varepsilon_s\tilde{\mathbf{C}}_{ssk}) + \text{tr}\left\{\mathcal{G}\left(\tilde{\mathbf{C}}_{ssk}, \tilde{\mathbf{M}}_{ssk}^H\right)\right\} \\ + \frac{\bar{\mathbf{s}}^H(\mathbf{L})\mathbf{J}(k)\bar{\mathbf{s}}(\mathbf{L})}{\varepsilon_s^2} - \frac{2\Re\{\mathbf{b}^T(k)\bar{\mathbf{s}}(\mathbf{L})\}}{\varepsilon_s}, \tag{61a}$$

$$\hat{\mathbf{s}}(k) = \bar{\mathbf{s}}(\hat{\mathbf{L}}(k)) \tag{61b}$$

where $\bar{\mathbf{s}}(\mathbf{L}) = \mathcal{Q}_{\mathcal{A}_M}\left(\varepsilon_s(\mathbf{b}^T(k)\mathbf{J}^{-1}(k))^H\right)$, $\mathbf{b}(k)$ is the vector consists of the diagonal elements of $\mathbf{Y}^H(k)\tilde{\mathbf{M}}_{ssk}\tilde{\mathbf{C}}_{ssk}^{-1}$ and $\mathbf{J}(k) = \tilde{\mathbf{C}}_{ssk}^{-1} \odot (\mathbf{Y}^H(k)\mathbf{Y}(k))^*$.

As the dimension of $\mathbf{A}_{zrc}(k)$ for the *ZRC detectors* (45) and (51), $N_T \times N_T$, is much smaller than that for the *CEEA-ML detectors*, $N_RN_T \times N_RN_T$, the former class of detectors needs far less memory space. The two-stage ZTC detector can be similarly derived.

VI. PERFORMANCE ANALYSIS OF CEEA-ML AND RELATED DETECTORS

The bit error rate (BER) performance of the SM detectors depends on the channel estimation method used and, because of the assumed frame structure, is a function of the data block location index k; see Fig. 1. For the MB systems, the performance is better when $\min\{k, (N - k)\mod N\}$ is small while for the DD counterparts, the performance degrades with increasing k. Let BER(k) be the BER of the kth block then the average BER is

$$\text{BER}_{\text{MB}} = \frac{1}{2N - 2}\sum_{k=1, k\neq N}^{2N-1} \text{BER}_{\text{MB}}(k), \tag{62}$$

$$\text{BER}_{\text{DD}} = \frac{1}{N - 1}\sum_{k=1}^{N-1} \text{BER}_{\text{DD}}(k). \tag{63}$$

For simplicity, the subscript for the CE used shall be omitted unless necessary. It is straightforward to show that [10], [19]

for *CEEA-ML detectors*

$$\text{BER}(k) \leq \frac{1}{2^{mB}}\frac{1}{mB}\sum_{(\mathbf{s},\mathbf{L})\in\mathcal{A}_M^B\times\mathcal{L}} \sum_{\substack{(\mathbf{s}',\mathbf{L}')\in \\ \mathcal{A}_M^B\times\mathcal{L}\setminus\{(\mathbf{s},\mathbf{L})\}}} d_H\left(\mathbf{X}, \mathbf{X}'\right) \\ \cdot P_k\left\{\mathbf{X}\rightarrow\mathbf{X}'\right\} \tag{64}$$

where $\mathbf{X} = \mathbf{L}\cdot\text{Diag}(\mathbf{s})$, $\mathbf{X}' = \mathbf{L}'\cdot\text{Diag}(\mathbf{s}')$, $d_H(\mathbf{X}, \mathbf{X}')$ denotes the Hamming distance between the information bits carried by $\mathbf{X}$ and by $\mathbf{X}'$, and $P_k\{\mathbf{X}\rightarrow\mathbf{X}'\}$ the averaged pairwise error probability (PEP) of detecting the transmitted signal $\mathbf{X}$ as $\mathbf{X}'$. We derive the conditional PEP of the MB-CE-aided *CEEA-ML detectors* in the followings; that for DD-CE-aided detectors can be similarly obtained.

$$P_k\left\{\mathbf{X}\rightarrow\mathbf{X}'|\hat{\mathbf{H}}(k)\right\} \\ = P_k\Big\{\mathcal{G}(\mathbf{C}'_{mb}(k), \text{vec}(\mathbf{Y}(k)) - \mathbf{m}'_{mb}(k)) + \log\det\mathbf{C}'_{mb}(k) \\ < \mathcal{G}(\mathbf{C}_{mb}(k), \text{vec}(\mathbf{Y}(k)) - \mathbf{m}_{mb}(k)) + \log\det\mathbf{C}_{mb}(k)|\hat{\mathbf{H}}(k)\Big\} \\ = P_k\Big\{(\mathbf{y} - \mathbf{d})^H\mathbf{D}(\mathbf{y} - \mathbf{d}) - \epsilon^H\mathbf{D}^{-1}\epsilon \\ + \mathcal{G}\left(\mathbf{C}'_{mb}(k), \mathbf{m}'_{mb}(k)\right) - \mathcal{G}\left(\mathbf{C}_{mb}(k), \mathbf{m}_{mb}(k)\right) \\ + \log\det\left(\mathbf{C}'_{mb}(k)\mathbf{C}_{mb}^{-1}(k)\right) < 0|\hat{\mathbf{H}}(k)\Big\} \tag{65} \\ \stackrel{def}{=} P_k\left\{(\mathbf{y} - \mathbf{d})^H\mathbf{D}(\mathbf{y} - \mathbf{d}) < \eta(\mathbf{X}, \mathbf{X}', k)|\hat{\mathbf{H}}(k)\right\},$$

where $\mathbf{m}'_{mb}(k)$ and $\mathbf{C}'_{mb}(k)$ are respectively obtained from (23a) and (23b) with $\mathbf{X}$ replaced by $\mathbf{X}'$ and (65) is obtained by completing the square with $\mathbf{D} = (\mathbf{C}'_{mb}(k))^{-1} - \mathbf{C}_{mb}^{-1}(k)$, $\mathbf{y} = \text{vec}(\mathbf{Y}(k))$, $\epsilon = (\mathbf{C}'_{mb}(k))^{-1}\mathbf{m}'_{mb}(k) - \mathbf{C}_{mb}^{-1}(k)\mathbf{m}_{mb}(k)$, and $\mathbf{d} = \mathbf{D}^{-1}\epsilon$. Based on (26), we have

$$\mathbf{y} \sim \mathcal{CN}\left(\text{vec}\left(\hat{\mathbf{H}}(k)\mathbf{X}\right), \tilde{\Psi}\right), \tag{66}$$

where $\tilde{\Psi} \stackrel{def}{=} \sigma_z^2\mathbf{I}_{N_RN_T} + (\mathbf{X}^T \otimes \mathbf{I}_{N_R})\Psi_E(k)(\mathbf{X}^* \otimes \mathbf{I}_{N_R})$. Representing the Gaussian random vector $\mathbf{y}$ by $\tilde{\Psi}^{\frac{1}{2}}\tilde{\mathbf{y}} + \text{vec}(\hat{\mathbf{H}}(k)\mathbf{X})$ where $\tilde{\mathbf{y}} \sim \mathcal{CN}(\mathbf{0}_{N_RN_T}, \mathbf{I}_{N_RN_T})$, we obtain an alternate quadratic form

$$(\mathbf{y} - \mathbf{d})^H\mathbf{D}(\mathbf{y} - \mathbf{d}) \\ = \left(\tilde{\Psi}^{\frac{1}{2}}\tilde{\mathbf{y}} + \text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right)^H\mathbf{D}\left(\tilde{\Psi}^{\frac{1}{2}}\tilde{\mathbf{y}} + \text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right) \\ = \left(\tilde{\mathbf{y}} + \tilde{\Psi}^{-\frac{1}{2}}\left(\text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right)\right)^H \\ \cdot \tilde{\Psi}^{\frac{1}{2}}\mathbf{D}\tilde{\Psi}^{\frac{1}{2}}\left(\tilde{\mathbf{y}} + \tilde{\Psi}^{-\frac{1}{2}}\left(\text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right)\right) \\ = \left(\mathbf{U}^H\tilde{\mathbf{y}} + \mathbf{U}^H\tilde{\Psi}^{-\frac{1}{2}}\left(\text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right)\right)^H \\ \cdot \Lambda\left(\mathbf{U}^H\tilde{\mathbf{y}} + \mathbf{U}^H\tilde{\Psi}^{-\frac{1}{2}}\left(\text{vec}(\hat{\mathbf{H}}(k)\mathbf{X}) - \mathbf{d}\right)\right) \stackrel{def}{=} q_{nc},$$

where $\mathbf{U}\Lambda\mathbf{U}^H$ is the eigenvalue decomposition of $\tilde{\Psi}^{\frac{1}{2}}\mathbf{D}\tilde{\Psi}^{\frac{1}{2}}$ with orthonormal matrix $\mathbf{U}$ and diagonal matrix Λ containing respectively the eigenvectors and corresponding eigenvalues. Since $\mathbf{U}^H\tilde{\mathbf{y}}$ and $\tilde{\mathbf{y}}$ have the same distribution, q_{nc} is

a noncentral quadratic form in Gaussian random variables and the CDF

$$P_k\left\{q_{nc} < \eta(\mathbf{X}, \mathbf{X}', k)|\hat{\mathbf{H}}(k)\right\} \quad (67)$$

can be evaluated by the method proposed in [20] or the series expansion approach of [21, Ch. 29].

The average PEP is thus given by

$$\begin{aligned}
P_k\left\{\mathbf{X} \to \hat{\mathbf{X}}\right\} &= \int P_k\left\{q_{nc} < \eta(\mathbf{X}, \mathbf{X}', k)|\hat{\mathbf{H}}(k)\right\} P(\hat{\mathbf{H}}(k))\,\mathrm{d}\hat{\mathbf{H}} \\
&= \int P_k\left\{q_{nc} < \eta(\mathbf{X}, \mathbf{X}', k)|\hat{\mathbf{H}}(k)\right\} \\
&\quad \times \frac{e^{-\mathcal{G}(\tilde{\Phi}(k), \mathrm{vec}(\hat{\mathbf{H}}(k)))}}{\pi^{N_R N_T} \det(\tilde{\Phi}(k))}\mathrm{d}\hat{\mathbf{H}}
\end{aligned} \quad (68)$$

where $\tilde{\Phi}(k) \overset{def}{=} \mathrm{E}\{\mathrm{vec}(\hat{\mathbf{H}}(k))\mathrm{vec}^H(\hat{\mathbf{H}}(k))\}$ is equal to (21c). As the both sides of the inequality in (67) depend on $\hat{\mathbf{H}}(k)$, the average over $\hat{\mathbf{H}}(k)$ (68) can only be computed by numerical integration. The BER performance of the *ZRC*, *ZTC* and general MIMO *CEEA-ML detectors* can also be analyzed by the same approach with the corresponding PEP derived from different spatial correlation structures.

VII. COMPUTING COMPLEXITIES AND MEMORY REQUIREMENTS OF VARIOUS SM DETECTORS

We now compare the computational complexities and memory requirements of the detectors derived so far. The memory space is used to store the required items involved in the detection metrics and is divided into two categories: i) fixed and ii) dynamic. The former stores the items that are independent of the received samples and/or updated channel estimates and can be calculated offline. The latter specifies those that vary with the received samples. Take the *ML detector* (24) for example, to achieve fast real-time detection, we pre-calculate and store $\mathbf{C}_{mb}^{-1}(k)$, $\det \mathbf{C}_{mb}(k)$ and $\mathbf{m}_{mb}(k)$ for all candidate signals ($\mathbf{X}$) and all k in two consecutive frames. These items are time-invariant, independent of $\mathbf{Y}(k)$ or $\hat{\mathbf{H}}(k)$. The dynamic part refers to the received samples (in two frames) which have to be buffered before being used to compute the MB channel estimate and detecting the associated signals. As $(\mathbf{X}^T(k) \otimes \mathbf{I}_{N_R})\mathbf{A}(k)$ for all candidate $\mathbf{X}$ can be precalculated and stored, the complexity to compute (23a) is only $\mathcal{O}(M^{N_T B}N_R^2 N_T^2)$ complex multiplications per block. The remaining complexity is that of computing $\mathcal{G}(\mathbf{C}_{mb}(k), \mathrm{vec}(\mathbf{Y}(k)) - \mathbf{m}_{mb}(k))$. Note that the complexities of the ML detectors for PAM, rectangular QAM and PSK signals, assuming perfect CSI, can be reduced by taking advantage of their regular constellation structure [13], [14]. The resulting performance is equal to that of the mismatched detector in the presence of CSI error and, as will seen in the next section, is worse than that of our detectors. The degradation ranges from small to very significant, depending on the CE used and the channel condition.

As the *CEEA-ML detector* degenerates to the *ZRC detector* by setting $\Phi_R = \mathbf{I}_{N_R}$, the dimensions of the correlation-related terms can be reduced by a factor of N_R.

This reduction directly affects both computing complexity and memory requirement. The complexity of the *two-stage detector* is only $1/M^B$ of its single-stage counterpart because the parallel search on both transmit antenna index and data symbol has been serialized. However, the memory required remains unchanged.

For the *mismatched detectors* (18), besides the dynamic memory to store $\hat{\mathbf{H}}$ and $\mathbf{Y}(k)$, the fixed memory, which consists mainly of those for storing the terms the CEs need, is relative small and usually dominated by the memory to store the statistics for the *CEEA-ML detectors*. A minimal complex multiplication complexity of $\mathcal{O}(M^{N_T}BN_RN_T)$ is called for, since without channel statistics data of each channel use can be detected separately, i.e., $\hat{\mathbf{X}}^{\mathrm{MM}}(k) \overset{def}{=} [\hat{\mathbf{x}}_1^{\mathrm{MM}}(k), \cdots, \hat{\mathbf{x}}_B^{\mathrm{MM}}(k)]$ with

$$\hat{\mathbf{x}}_j^{\mathrm{MM}}(k) = \arg \min_{\mathbf{x} \in \mathcal{A}_{M+}^{N_T}} \|\mathbf{y}_j(k) - \hat{\mathbf{H}}\mathbf{x}\|_F^2.$$

The detectors using DD channel estimates needs significantly less memory than those using MB ones as they do not jointly estimate the channel of several blocks and (35) indicates that $\mathbf{C}_{dd}^{-1}$ is independent of the block index. We summarize the computing complexities of various detectors in Tables 4.

We conclude that, among all the proposed detectors, the *ZRC* (or *ZTC*) detectors are the most desirable as they require the minimal computation and memory to achieve satisfactory detection performance. Although the *mismatched detector* is the least complex and requires only comparable memory as *ZRC* (or *ZTC*) detectors do, its performance, as shown in the following section, is much worse than that of the proposed detectors in some cases.

VIII. SIMULATION RESULTS

In this section, the BER performance of the detectors we derived is evaulated through computer simulations. We use the S-T channel model of [16] and assume that uniform linear arrays (ULAs) are deployed on both sides of the link. The spatial correlation follows the Kronecker model (41) so that (6) can be written as $\rho_S(i-m, j-n) = [\Phi_T]_{jn} \cdot [\Phi_R]_{im} = \rho_S(0, j-n) \cdot \rho_S(i-m, 0)$. When the angle-of-arrivals (AoAs) and angle-of-departures (AoDs) are uniformly distributed in $(0, 2\pi]$, we have [16]

$$\rho_S(\ell - \ell', 0) = \rho_S(0, \ell - \ell') = J_0(2\pi(\ell - \ell')\delta/\lambda), \quad (69)$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, δ the antenna spacing for both transmitter and receiver and λ the signal wavelength. On the other hand, if the AoAs and AoDs have limited angle spreads [22], the spatial correlation can be expressed as

$$[\Phi_R]_{ij} = r^{|i-j|}, \quad [\Phi_T]_{ij} = t^{|i-j|} \quad (70)$$

with $0 \le r, t < 1$. The above exponential model has been widely applied for MIMO system evaluation [23] and proven to be consistent with field measurements [24]. We use both

TABLE 4. Computational complexity (in number of real-valued multiplications) involved to detect a data block for various detectors and modulation schemes.

	General MIMO	SM	Two-stage
Mismatched	$2M^{N_T}N_RB(2N_T+1)$	$6MN_TN_RB$	N/A
CEEA-ML	$4M^{N_TB}N_RB(N_RB + N_TN_R+1)$	$4M^BN_T^BN_RB(N_RB + N_TN_R+1)$	$2N_T^BN_RB(4N_RB + 2N_TN_R+7)$
ZRC	$4M^{N_TB}N_TN_R(N_T+B+1)$	$4M^BN_T^{B+1}N_R(N_T+B+1)$	$2N_T^B[2N_TN_R(N_T+B+1) + 4B^2(N_R+1)+2N_T^2+B]$

TABLE 5. Size of memory (in numbers of complex values) of various detectors to store required items to detect a data block where $\gamma = M^{N_TB}$, N_T^B, and $M^BN_T^B$ for general MIMO, PSK-SM, and QAM-SM signals, respectively.

		MB channel estimates	DD channel estimates
Mismatched	Fixed	$\mathcal{O}(N)$	$\mathcal{O}(B)$
	Dynamic	$\mathcal{O}(NN_RN_T)$	$\mathcal{O}(N_RN_T)$
CEEA-ML	Fixed	$\mathcal{O}(\gamma NN_R^2N_T^2)$	$\mathcal{O}(\gamma NN_R^2N_T^2)$
	Dynamic	$\mathcal{O}(NN_RN_T)$	$\mathcal{O}(NN_RN_T)$
ZRC	Fixed	$\mathcal{O}(\gamma NN_T^2)$	$\mathcal{O}(\gamma N_T^2)$
	Dynamic	$\mathcal{O}(NN_RN_T)$	$\mathcal{O}(N_RN_T)$

(70) and the isotropic model (69). As for time selectivity, we assume that it is characterized by Jakes' model [25]

$$\rho_T(k-\ell) = J_0\left(2\pi f_D \left\lfloor \frac{k-\ell}{K_c} \right\rfloor K_cBT_s\right), \qquad (71)$$

where f_D is the maximum Doppler frequency, T_s the symbol duration, K_c the coherent time in blocks. In this section, $K_c = 1$ is assumed. The remainder of this section considers SM and SMX MIMO systems of different system parameters and transmission rates.

Remark 4: Note that although we consider only uncoded systems, most of the detection decision variables are proportional or approximately proportional to the likelihood function, hence for a coded SM system, one can use them to compute the log-likelihood ratio as the soft input to the decoder. Performance in a coded SM system can be improved not only through coding but by performing iterative soft message exchanges between the decoder and the data detector (or channel estimator) [27].

Remark 5: Recall that the SM systems considered, like most studies on SM systems [1], are narrowband ones and are proper model of subchannels (subcarriers) of a wideband multicarrier (e.g., OFDM) system. For an OFDM system with operating frequency 2.4 GHz and subcarrier spacing 15 KHz, $T_s = 66.67\ \mu s$ and $f_DT_s = 0.01$ means a modest vehicular speed of 67.5 km/hr.

Remark 6: As the CE reliability is crucial to the system performance, a proper power splitting between pilots and data may bring about performance improvement. However, the splitting ratio, which is a function of the received SNR per transmit antenna, must be known for the transmitter to adjust the transmit power and for the receiver to calibrate the decision boundaries when detecting QAM or PAM signals. This is an overhead which is often neglected. The optimal power splitting issue is addressed only in some papers discussing MIMO CE algorithms when analytic expressions can be derived. The information of the (varying) pilot power level is also needed for the purpose of power control, handover, and sometimes for coarse link distance estimation as well.

A. MB-CE-AIDED SM DETECTORS

The performance of MB-CE-aided CEEA detectors is presented in Figs. 2 and 3 where we plot the BER performance of the *CEEA-ML*, (27), *mismatched*, (18), and suboptimal detectors as a function of $\bar{E}_b/N_0$, $\bar{E}_b$ being the average received bit energy per antenna. The suboptimal detectors include the *two-stage* (34), *ZRC* (45), *ZTC* (50), and the *two-stage ZRC detectors*, (53a) and (53b). Both the ML and suboptimal detectors outperform the *mismatched* one. With $N = 10$, the *two-stage detector*, which requires a much lower complexity, suffers only negligible degradation w.r.t. its *CEEL-ML* counterpart. The *ZTC detector* suffers slightly more performance degradation than the *ZRC* one does for it is obtained by using the extra approximation $\mathbf{X}(k)\mathbf{X}^H(k) \approx \varepsilon_s\mathbf{I}_{N_T}$.

The effect of spatial correlation can be found by comparing the curves corresponding to $N = 10$ (frame duration $40T_s$). When the spatial correlation follows (69) with $\delta = 0.5$ or 1λ, the correlation level is relatively low and the knowledge of this information gives limited performance gain. But if the correlation is high, as that described by (70) with $r = t = 0.8$ or 0.5, it becomes more difficult for an SM detector to resolve spatial channels (different $\mathbf{h}_j$'s) and the detector performance degrades accordingly. This holds for detectors with perfect CSI and those with imperfect CSI. Neglecting CSI error and spatial correlation cause more performance loss for channels with stronger correlations as can be found by comparing their mismatch losses. Higher spatial correlation also causes larger performance degradation for the *ZRC* and *ZTC detectors* which lack one side's spatial information. The effect of a shorter frame ($N = 5$) can be found in the same figures as well. As the CSI error is reduced, the mismatch
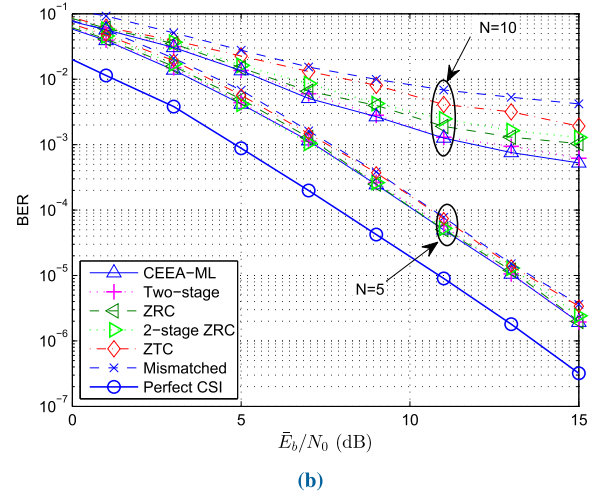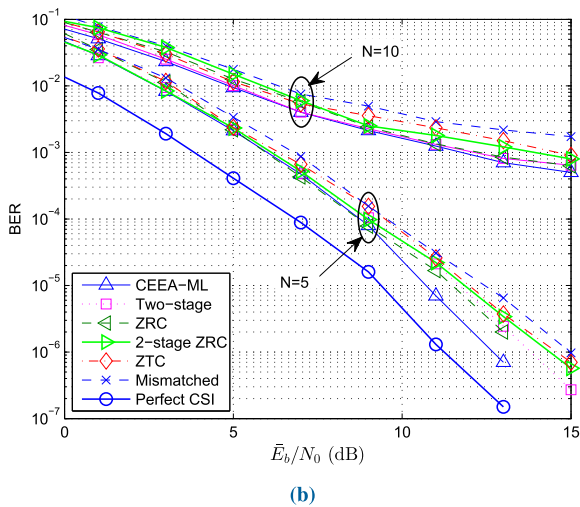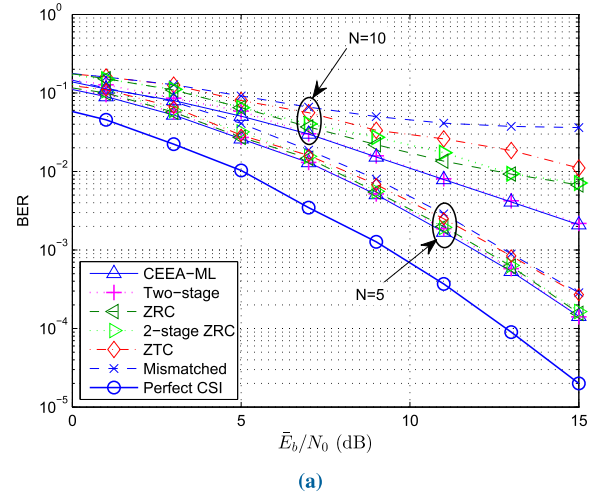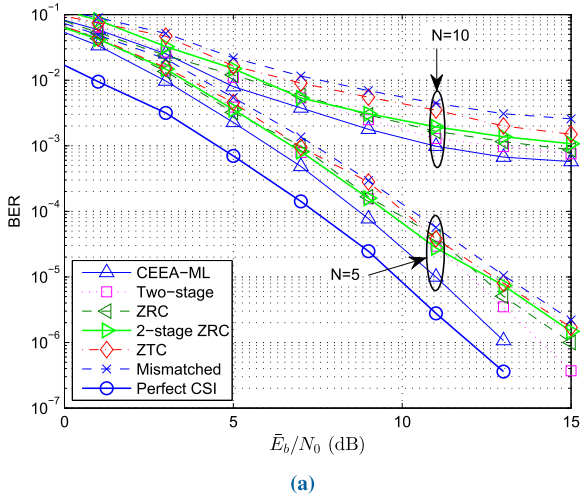
FIGURE 2. Impact of the frame size and antenna spacing on the performance of MB-CE-aided SM detectors in a Bessel correlated channel ($f_D T_s = 0.01$): (a) $\delta = 0.5\lambda$ (b) $\delta = 1\lambda$; $M = 4$, $N_T = N_R = B = 4$.

FIGURE 3. Impacts of frame size and spatial correlation on the performance of MB-CE-aided SM detectors in exponentially-correlated channels: (a) $r = t = 0.8$; (b) $r = t = 0.5$, $M = 4$, $N_T = N_R = B = 4$, $f_D T_s = 0.01$.

loss, which is proportional to the CSI error, becomes smaller accordingly.

The performance of the ML detector with perfect CSI (11) is insensitive to the time selectivity. For other detectors, the CSI error increases with a larger $f_D T_s$ and/or a sparser pilot density and so is their performance degradations. For example, from (71) we find that, for $f_D T_s = 0.01$, the 50%-coherence time is approximately $24.2\,T_s$ and thus with the frame size $N = 10$, each antenna receives a pilot symbol every other $39 T_s$ which is too sparse to track the channel's temporal variation. Although knowing the associated CSI error statistics helps reducing the performance loss, increasing the pilot density to obtain a more reliable CE is much more efficient-doubling the pilot density ($N = 5$) recovers most losses.

The above results assume that perfect spatial correlation information is available. We examine the impact of imperfect spatial information in Figs. 4 and 5 where Φ_R and Φ_T used by *ZTC* and *ZRC* detectors are estimated by first taking the

time averages over three consecutive pilot blocks

$$\bar{\Phi}_R \stackrel{def}{=} \frac{1}{3N_T} \sum_{k=0,N,2N} \hat{\mathbf{H}}(k)\hat{\mathbf{H}}^H(k), \tag{72a}$$

$$\bar{\Phi}_T \stackrel{def}{=} \frac{1}{3N_R} \sum_{k=0,N,2N} \hat{\mathbf{H}}^T(k)\hat{\mathbf{H}}^*(k) \tag{72b}$$

These initial estimates are then improved by using the exponentially-decay model

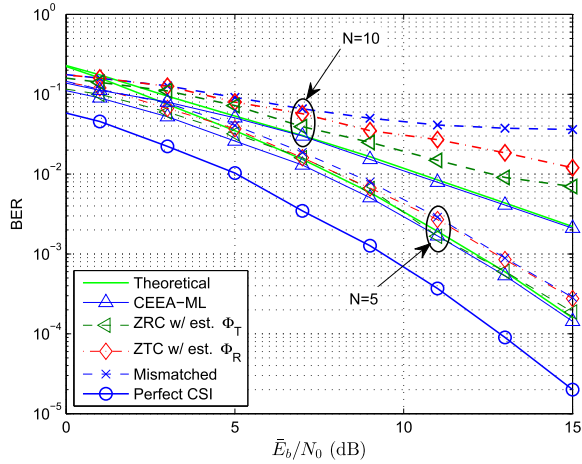$$[\hat{\Phi}_R]_{ij} = \hat{r}^{|i-j|}\mathrm{sgn}(\Re\{[\bar{\Phi}_R]_{ij}\}), \tag{73a}$$

$$[\hat{\Phi}_T]_{ij} = \hat{t}^{|i-j|}\mathrm{sgn}(\Re\{[\bar{\Phi}_T]_{ij}\}), \tag{73b}$$
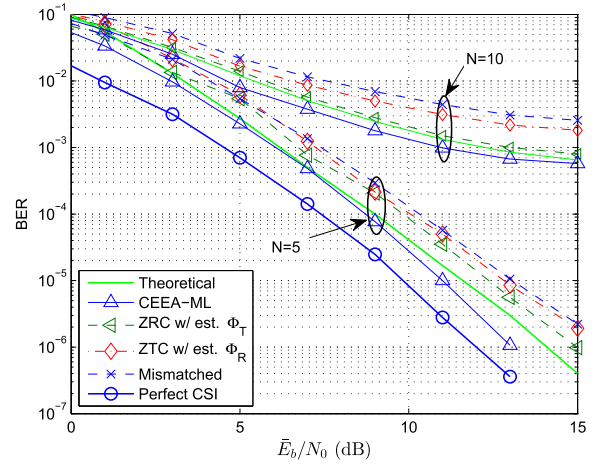
where

$$\hat{r} = \arg\min_{0 \le r < 1} \sum_{i,j} \left| r^{|i-j|} - \left|[\bar{\Phi}_R]_{ij}\right| \right|^2, \tag{74a}$$

$$\hat{t} = \arg\min_{0 \le t < 1} \sum_{i,j} \left| t^{|i-j|} - \left|[\bar{\Phi}_T]_{ij}\right| \right|^2. \tag{74b}$$
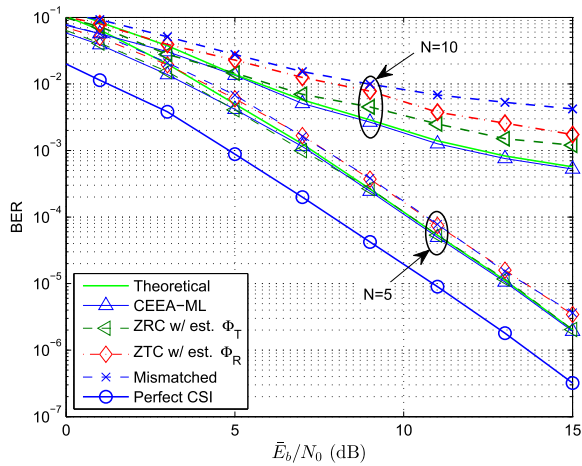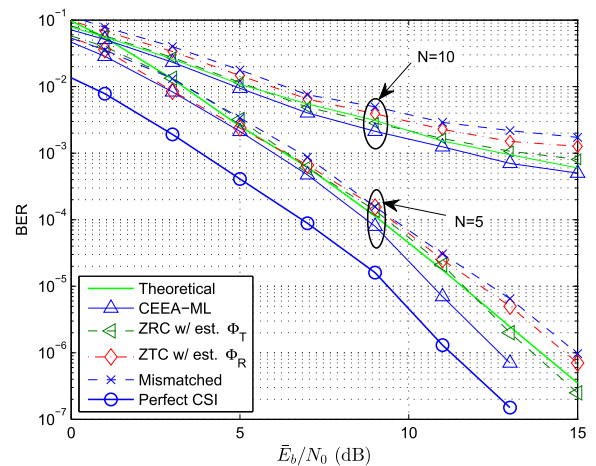
The temporal correlation can be similarly estimated.

FIGURE 4. Effect of imperfect spatial correlation information on the MB-CE-aided ZRC and ZTC detectors' performance in exponentially-correlated channels; $N = 5$ or 10, $M = 4$, $N_T = N_R = B = 4$, $f_D T_s = 0.01$, and (a) $r = t = 0.8$, (b) $r = t = 0.5$.

FIGURE 5. Effect of imperfect spatial correlation information on the MB-CE-aided ZRC and ZTC detectors' performance in Bessel-correlated fading channels; $N = 5$, $M = 4$, $N_T = N_R = B = 4$, $f_D T_s = 0.01$, and (a) $\delta = 0.5\lambda$; (b) $\delta = 1\lambda$.

The performance of ZTC and ZRC detectors shown in these two figures indicates that the refined spatial correlation estimator (74a) and (74b) gives fairly accurate spatial correlation estimates. Both detectors keep their performance advantages over the *mismatched* counterparts when $N = 10$. But with a denser pilot $N = 5$ (thus smaller mismatch error), the *ZTC detector* (50) fails to offer noticeable gain due perhaps to additional approximation (47) used. In Fig. 5, the channel correlation follows (69) but the receiver still assumes (70) and uses the estimator (74a) and (74b). In spite of the correlation model discrepancy, the detectors still outperform the *mismatched* one. The theoretical performance upper bound of the *CEEA-ML detector* analyzed in Section VI is also shown in Figs. 4, 5 and 8–10. Except in the lower $\bar{E}_b/N_0$ region, the theoretical bounds are tight and give reliable numerical predictions. Similar accurate theoretical predictions on the effect of N are found in Figs. 6 and 8 for several 16-QAM SM detectors.

Fig. 7 show the effects of reducing the receive antenna number and increasing the pilot number for estimating the modeling coefficients in (14). As expected, the performance of the 4×2 system is much worse than that of its 4×4 counterpart (cf. Fig. 3b) due to the smaller receive diversity order. However, smaller N_R also lessens the impact of CSI error and thus reduces the mismatch loss. More pilots in the same modeling period (21 blocks) improves the system performance but the effect of a denser pilot ($N = 5$) is more impressive; see also Figs. 4-6. The MB-CE's error consists of the modeling error and that due to noise. The latter estimation error is reduced by having five pilots in 21 blocks but which seems to be insufficient to compensate for the larger modeling error introduced by increasing the modeling period from 11 blocks (with three pilots) to 21 blocks.

B. DD-CE-AIDED SM DETECTORS
Figs. 8 and 9 present the performance of the DD-CE-aided detectors. As expected, the proposed detectors outperform
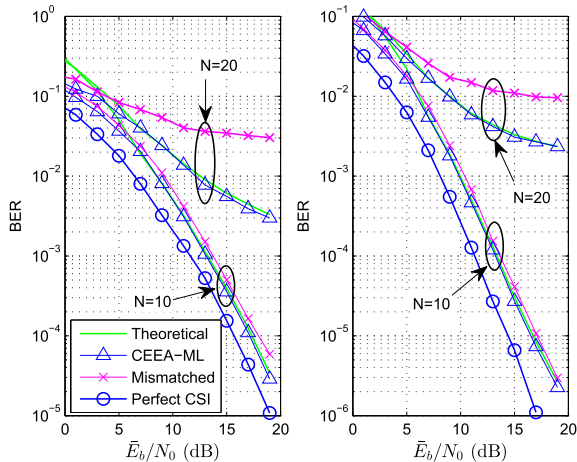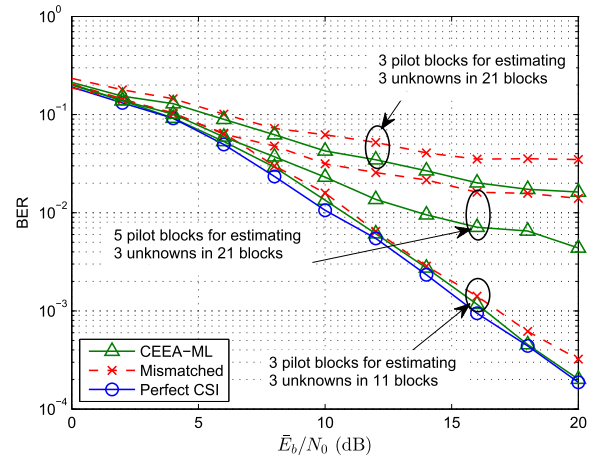
FIGURE 6. BER performance of MB-CE-aided 16-QAM SM systems (5-bit/transmission) with different frame sizes and spatial correlations; $N = 10$ or 20, $N_T = B = 2$, $N_R = 4$, $f_D T_s = 0.01$, and (left) $r = t = 0.8$; (right) $r = t = 0.5$.

FIGURE 7. Impacts of the number of pilot blocks per frame and the receive antenna number on the performance of MB-CE-aided SM detectors in exponentially-correlated fading channels; $r = t = 0.5$; $M = 4$, $N_T = B = 4$, $N_R = 2$, $f_D T_s = 0.01$.

the *mismatched* one. The effects of the pilot density, correlation level and other behaviors of these detectors are similar to those observed in MB-CE-aided detectors. But the DD-CE-aided detectors are more sensitive to the CSI error. This is because the way the CE is updated makes it quickly outdated in a fast fading environment and any decision error will propagate until the next pilot block is received.

In Figs. 6 and 8, we compare the effect of pilot density on both classes of detectors. With long frame size ($N = 20$), both MB- and DD-CE-aided detectors give unsatisfactory performance although the former is slightly better. But if we increase the pilot density to $N = 10$, the MB-CE-aided detector offers more significant improvement: doubling the pilot density gives a 3.5 dB gain at BER $= 1 \times 10^{-2}$ and $r = t = 0.8$ (or 2.9 dB at $r = t = 0.5$) for the ML-MB detector, in contrast to the 2.6 dB (2.4 dB) gain for the ML-DD detector. Obviously, the CSI accuracy is of great importance and the *CEEA-ML detectors* significantly outperform the *mismatched* ones when CSI is unreliable. The DD-CE is improved in a system using a smaller QAM constellation $\mathcal{A}_M$; see Fig. 9 where $M = 4$ is assumed. The *ZRC* and *ZTC detectors* ignore part of spatial correlation, hence, it is only natural that their performance becomes closer to that of the *CEEA-ML detector* as the spatial channel decorrelates (when r and t become smaller).

C. CEEA-ML DETECTION OF SMX SIGNALS
In Fig. 10, we show the performance of the SMX system in an S-T correlated fading channel using the *CEEA-ML detector* (24). The SMX system's parameter values are $B = N_T = N_R = 2$ and $M = 4$ so that it yields a rate of 4 bits/transmission. The figure reveals that both the *CEEA-ML detector* and its suboptimal variations also offer performance gain against the *mismatched* one. The SM system with 4 bits/transmission and the same frame size,
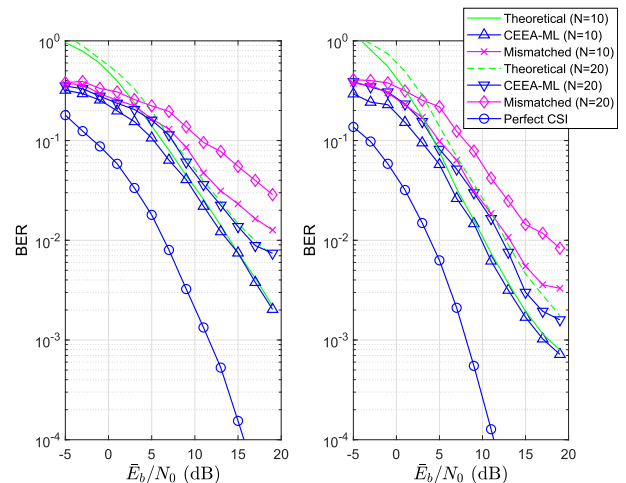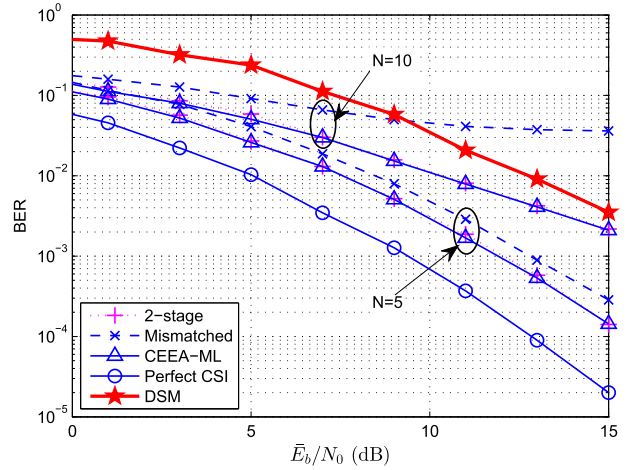

FIGURE 8. Performance of 5-bit/transmission DD-CE-aided SM systems with different frame sizes and exponential spatial correlations; 16-QAM $\mathcal{A}_M$, $N = 10$ or 20, $N_T = B = 2$, $N_R = 4$, $f_D T_s = 0.01$, and (left) $r = t = 0.8$; (right) $r = t = 0.5$.
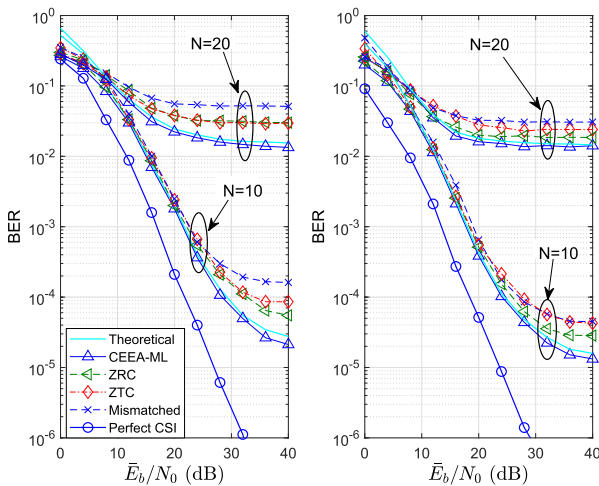
as shown in Fig. 4, achieves the same BER performance (say, at BER $= 10^{-3}$) with a much lower SNR. The SMX systems result in higher BER error floors whereas the SM counterparts yield performance that is much closer to that achieved with perfect CSI when the frame size is $20T_s$ ($N = 10$). This is because in the high-SNR regime where the ICI is the dominant deteriorating factor for an SMX system, a high spatial correlation may result in occasionally deep-fade across all spatial channels (all $|h_{ij}|$'s are small) and a burst of erroneous symbols. Since the SM systems do not suffer from ICI, a rare single-channel fade has less severe impact on its BER performance. Nevertheless, using more reliable CSI ($N = 10$) still helps to reduce an SMX system's error floor. We present the MB-CE-aided SMX detectors' performance only as the DD-CE-aided detectors give even worse performance.
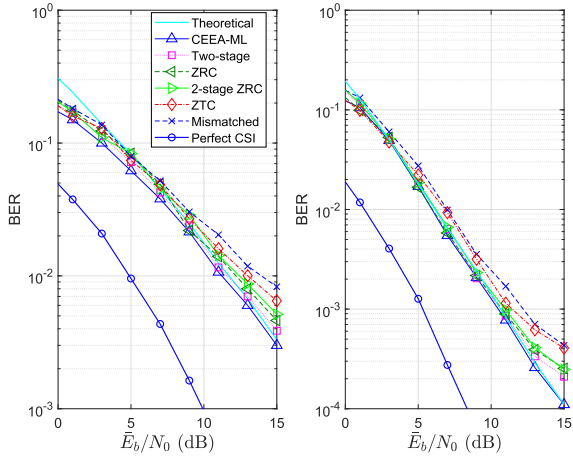
FIGURE 9. Performance of a 3-bit/transmission DD-CE-aided SM system in exponentially-correlated fading channels; $N = 10$, $M = 4$, $N_T = B = 2$, $N_R = 4$, $f_D T_s = 0.01$, and (left) $r = t = 0.8$; (right) $r = t = 0.5$.

FIGURE 10. Performance of the MB-CE-aided SMX systems in exponentially-correlated fading channels; $N = 10$ or 20, $M = 4$, $N_T = N_R = B = 2$, $f_D T_s = 0.01$, and (left) $r = t = 0.8$ or (right) $r = t = 0.5$.

D. DIFFERENTIAL SM

The differential spatial modulation system imposes limitations on antenna selection and is thus less spectral efficient. It is applicable for constant modulus signals (e.g., PSK) only and not suitable for QAM or PAM signals. Moreover, a basic assumption in applying the differential SM (DSM) scheme is that the channel remains static in at least two (usually much more than this in practice) consecutive frames (blocks) whence if the channel is turbulent enough to invalidate this assumption, the DSM system will not work properly. In Fig. 11, we show the performance comparison between the DSM scheme [11] and the proposed detectors with MB CE. Due to rate difference between DSM ($\frac{1}{4}\log_2(4!) + 2$ bits/transmission) and SM (4 bits/transmission), the SNR has been normalized for fair comparison. As can be seen, even when the channel varies slowly ($f_D T_s = 0.001$), the DSM system is still inferior to the 2-stage and CEEA-ML detectors

FIGURE 11. Performance of the MB-CE-aided SM and differential SM systems in exponentially-correlated channels; $N = 5$ or 10, $M = 4$, $N_T = N_R = B = 4$, $f_D T_s = 0.001$, and $r = t = 0.8$.

with sparse pilots ($N = 10$). For a higher pilot density ($N = 5$) the performance gap even increases drastically.

IX. CONCLUSION

We have derived ML and various suboptimal detector structures for general MIMO (including SM and SMX) systems that take into account practical design factors such as the channel's S-T correlations, the CE used and the corresponding estimation error. We also suggest a model-based spatial (and time) correlation estimator that yields quite accurate results. For SM systems, we show that the derived detectors are capable of reducing the mismatch loss and are especially effective in fast fading channels.

The suboptimal detectors are obtained by simplifying the ML detector's exhaustive search effort, the spatial correlation structure, the likelihood function, or a combination of these approximations. The complexities and memory requirements of the ML, suboptimal and mismatched detectors are analyzed. The effects of space and/or time selectivity and CSI error using MB- or DD-CEs on the system performance are studied via both analysis and computer simulations. Their performance is compared with that of perfect CSI detectors. It is found that the CEEA-ML and some suboptimal detectors suffer only minor degradation and, moreover, the class of two-stage detectors requires low complexities which are independent of the signal constellation size. The numerical results, including the performance of SMX system, enable us to verify the usefulness of our error rate analysis and demonstrate how the CSI uncertainty affects various detectors' BER performance, identifying the parameter ranges for which the fading channel's time or spatial selectivity has to be taken into consideration. They also help finding the channel conditions and performance requirements under which the low-complexity suboptimal detectors incur only minor performance degradation and become viable implementation choices.

REFERENCES

[1] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[2] S. K. Wilson, R. E. Khayata, and J. M. Cioffi, "16 QAM modulation with orthogonal frequency division multiplexing in a Rayleigh-fading environment," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Stockholm, Sweden, Jun. 1994, pp. 1660–1664.

[3] J. Dai, K. Niu, Z. Si, and D. Zhang, "Polar-coded spatial modulation," *IEEE Trans. Signal Process.*, vol. 69, pp. 2203–2217, 2021.

[4] E. Baccarelli and M. Biagi, "Performance and optimized design of space-time codes for MIMO wireless systems with imperfect channel estimates," *IEEE Trans. Signal Process.*, vol. 5, no. 10, pp. 2911–2913, Oct. 2004.

[5] K.-C. Lai, H.-Y. Su, D.-G. Peng, and S.-Z. He, "Detection algorithm for single-carrier spatial modulation signals in frequency-selective fading channels," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13341–13356, Nov. 2020.

[6] R. K. Mallik and P. Garg, "Performance of optimum and suboptimum receivers for space-time coded systems in correlated fading," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1237–1241, May 2009.

[7] J. Zhang, Y. V. Zakharov, and R. N. Khal, "Optimal detection for STBC MIMO systems in spatially correlated Rayleigh fast fading channels with imperfect channel estimation," in *Proc. Conf. Rec. 43rd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2009, pp. 1387–1391.

[8] M. Di Renzo and H. Haas, "Space shift keying (SSK) modulation with partial channel state information: Optimal detector and performance analysis over fading channels," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3196–3210, Nov. 2010.

[9] R. Mesleh, O. S. Badarneh, A. Younis, and H. Haas, "Performance analysis of spatial modulation and space-shift keying with imperfect channel estimation over generalized $\eta-\mu$ fading channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 88–96, Jan. 2015.

[10] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Performance of spatial modulation in the presence of channel estimation errors," *IEEE Commun. Lett.*, vol. 16, no. 2, pp. 176–179, Feb. 2012.

[11] Y. Bian, X. Cheng, M. Wen, L. Yang, H. V. Poor, and B. Jiao, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3262–3268, Jul. 2015.

[12] H. Chang, Y. Liu, and Y. T. Su, "Detection of spatial-modulated signals in the presence of CSI error and time-spatial correlation," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 82–86.

[13] P. Yang, Y. Xiao, Y. L. Guan, K. V. S. Hari, A. Chockalingam, S. Sugiura, and H. Haas, "Single-carrier SM-MIMO: A promising design for broadband large-scale antenna systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1687–1716, 3rd Quart., 2016.

[14] H. Men and M. Jin, "A low-complexity ML detection algorithm for spatial modulation systems with MPSK constellation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1375–1378, Aug. 2014.

[15] M.-X. Chang and Y. T. Su, "Model-based channel estimation for OFDM signals in Rayleigh fading," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 540–544, Apr. 2002.

[16] A. Giorgetti, P. J. Smith, M. Shafi, and M. Chiani, "MIMO capacity, level crossing rates and fades: The impact of spatial/temporal channel correlation," *J. Commun. Netw.*, vol. 5, no. 2, pp. 104–115, Jun. 2003.

[17] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing* (Estimation Theory), vol. 1. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[19] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1124–1144, Mar. 2012.

[20] T. Y. Al-Naffouri, M. Moinuddin, N. Ajeeb, B. Hassibi, and A. L. Moustakas, "On the distribution of indefinite quadratic forms in Gaussian random variables," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 153–165, Jan. 2016.

[21] N. L. Johnson and S. Kotz, *Distributions in Statistics: Continuous Univariate Distributions*, vol. 2. Hoboken, NJ, USA: Wiley, 1970.

[22] S. Büyükçorak and G. K. Kurt, "Spatial correlation and MIMO capacity at 2.4 GHz," *Procedia Technol.*, vol. 3, pp. 9–17, Jan. 2012.

[23] (Jul. 2015). *Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society (METIS); METIS Channel Models*. METIS Deliverable D1.4 V3. [Online]. Available: https://www.metis2020.com/documents/deliverables/

[24] D. Chizhik, J. Ling, P. W. Wolniansky, R. A. Valenzuela, N. E. Costa, and K. Huber, "Multiple-input-multiple-output measurements and modeling in Manhattan," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 321–331, Apr. 2003.

[25] W. C. Jakes, *Microwave Mobile Communications*. Hoboken, NJ, USA: Wiley, 1974.

[26] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and Modulation*, 3GPP, document TR 36.211 V11.0.0, Oct. 2012. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36211.htm

[27] B. Zheng, X. Wang, M. Wen, and F. Chen, "Soft demodulation algorithms for generalized spatial modulation using deterministic sequential Monte Carlo," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3953–3967, Jun. 2017.

HSUAN-CHENG CHANG received the B.S. and M.S. degrees in communications engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2011 and 2013, respectively. He was a Protocol Software Engineer of Asus, Taipei, Taiwan, from 2013 to 2017. Since 2017, he has been with MediaTek Inc., Hsinchu, first as an LTE/5G PHY Software Designer and focused on 5G and mmW software validation with cooperations from Ericsson and Huawei, since 2020.

YEN-CHENG LIU received the B.S. and Ph.D. degrees in communications engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2008 and 2017, respectively. He joined MediaTek, Taiwan, in 2017, where he is currently a Technical Manager involving in 5G NR transceiver algorithm design and standardization. His research interests include data demodulation and decoding, CSI tracking and prediction, and control signal design.

YU T. SU (Life Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Southern California, USA, in 1983. From 1983 to 1989, he was a Corporate Scientist at LinCom Corporation, Los Angeles, USA, where he was involved in various satellite communication and measurement systems design. Since 1989, he has been with National Chiao Tung University (now National Yang-Ming Chiao Tung University), Hsinchu, Taiwan, and is now a Retired Professor of the electrical and computer engineering. He was an Associate Dean of the College of Electrical and Computer Engineering, from 2004 to 2007, Head of the Communications Engineering Department from 2001 to 2003. He was also affiliated with the Microelectronic and Information Systems Research Center of the same university and served as the Deputy Director from 1997 to 2000. From 2005 to 2008, he was the Area Coordinator of Taiwan's National Science Counsel's Telecommunications Program and was the Director of the Special B5G/6G Program, from 2017 to 2021. His main research interests include communication systems, networking, and statistical signal processing.

• • •