# Robust Receiver Design for MIMO Single-Carrier Block Transmission over Time-Varying Dispersive Channels Against Imperfect Channel Knowledge

Chih-Yuan Lin, *Member, IEEE,* Jwo-Yuh Wu, *Member, IEEE,* and Ta-Sung Lee, *Senior Member, IEEE*

***Abstract*—We consider MIMO single-carrier block transmission over time-varying multipath channels, under the assumption that the channel parameters are not exactly known but are estimated via the least-squares training technique. While the channel temporal variation is known to negate the tone-by-tone frequency-domain equalization facility, it is otherwise shown that in the time domain the signal signatures can be arranged into groups of orthogonal components, leading to a very natural yet efficient group-by-group symbol recovery scheme. To realize this figure of merit we propose a constrained-optimization based receiver which also takes into account the mitigation of channel mismatch effects caused by time variation and imperfect estimation. The optimization problem is formulated in an equivalent unconstrained generalized-sidelobe-canceller setup. This enables us to directly model the channel mismatch effect into the system equations through the perturbation technique and, in turn, to further exploit the statistical assumptions on channel temporal variation and estimation errors for deriving a closed-form solution. Within the considered framework the proposed robust equalizer can be combined with the successive interference cancellation mechanism for further performance enhancement. Flop count evaluation and numerical simulation are used to evidence the advantages of the proposed scheme.**

***Index Terms*—MIMO, single carrier block transmission, time-varying multipath channels, channel estimation, constrained optimization, generalized sidelobe canceller, perturbation analysis.**

## I. INTRODUCTION

MULTI-INPUT multi-output (MIMO) single-carrier (SC) block transmission with cyclic prefix (CP) has been identified as one key technique for supporting high data rates over frequency selective fading channels [1]; such a system configuration can be found in the uplink mode of the next-generation wireless communication standards like 3GPP-LTE [2], [3]. One particularly attractive feature unique to MIMO-SC systems is the low-complexity per-tone frequency-domain equalization (FDE) scheme, which facilitates the development of several efficient (batch or adaptive) receiver implementations [1], [4]. Such a figure of merit, however, hinges crucially on the time-invariant channel assumption. When the channel is otherwise subject to fast temporal variation, orthogonality among the signal components in the frequency domain will no longer be preserved, and tone-by-tone signal recovery is then rendered impossible. Moreover, in a time-varying environment channel parameter mismatch due to the mere availability of the out-dated channel estimate will become another detrimental factor dominating the system performance. For MIMO-SC block transmission over time-varying dispersive channels, robust receiver designs which can efficiently tackle the joint impacts due to channel time variation and estimation errors, to the best of our knowledge, are hardly found in the literature.

This paper proposes a robust receiver design scheme for MIMO-SC systems when the multipath channels undergo time selectivity, and are not exactly known but instead estimated via the least-squares (LS) training technique [5], [6]. In lieu of performing signal recovery in the frequency domain, the proposed approach relies on received data processing in the time domain. Specifically, by exploiting the cyclic shifting property of the time-domain channel matrix it is shown that the signal signatures can be arranged into groups of orthogonal columns. In case that the inter-group interference can be effectively mitigated, the orthogonality structure in conjunction with space-time matched filtering will lead to a low-complexity intra-group symbol recovery scheme. Toward realizing such a time-domain processing facility, a linear weighting matrix is designed for each group based on the constrained optimization formulation [7], [8]. To further mitigate channel mismatch due to time variation and estimation errors, we leverage the generalized side-lobe canceller (GSC) principle [9], [10], [11] to transform the constrained optimization problem into an equivalent unconstrained setup. The unconstrained GSC formulation enables us to leverage the perturbation analysis technique [12], [13] to explicitly model the channel mismatch effect into the system equations, in turn providing a unique two-fold advantage. First, this allows a very natural cost function for weighting matrix design against channel mismatch. Second, we can then exploit the underlying statistical assumptions on the channel errors, due to both temporal variation and imperfect estimation, to derive an *analytic* solution.

We note that constrained optimization based designs resistant to signal/channel parameter uncertainty has been ad-

Manuscript received April 2, 2007; revised September 18, 2007 and February 3, 2008; accepted June 24, 2008. The associate editor coordinating the review of this paper and approving it for publication was S. Aissa. This work is sponsored jointly by the National Science Council under grants NSC-97-2221-E-009-101-MY3, NSC-97-2221-E-009-056-MY2 and NSC-96-2628-E-009-003-MY3, by the Ministry of Education of Taiwan under the MoE ATU program, and by the MediaTek research center at National Chiao Tung University, Taiwan.

The authors are with the Department of Communication Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C. (e-mail: cylin.cm90g@nctu.edu.tw; jywu@cc.nctu.edu.tw; tslee@mail.nctu.edu.tw).

Digital Object Identifier 10.1109/T-WC.2008.070348

dressed in other contexts such as beamforming [14], [15], [16], and multiuser detection [17], [18]. The adopted approaches therein are however significantly different from the proposed method in this paper. In [14], [15], and [16], the parameter mismatch is considered instead to be deterministic, the design criteria are of a min-max type, and the solutions are obtained via the convex optimization techniques. In [17] and [18] the parameter errors are modeled as random variables with a Gaussian distribution; the problem is then solved by exploiting the Gaussinality assumption and an associated linear or nonlinear programming setup. The proposed GSC based formulation combined with perturbation analysis in this paper relies on a stochastic error modeling strategy similar to that in [17] and [18]. However, as we will show it merely calls for the knowledge of the first- and second-order error statistics but, unlike [17] and [18], is free from any priori assumptions on the error distributions. For high-rate MIMO-OFDM transmission over channels with long delay spreads, such an approach has been adopted in [19] for robust receiver design against channel estimation error. The rest of this paper is organized as follows. Section II presents some preliminary results. Section III introduces the motivation behind the proposed scheme, proposes a constrained-optimization based solution framework, and briefly reviews the equivalent unconstrained GSC formulation, all under the perfect channel knowledge assumption. Section IV shows the proposed robust equalization scheme. Section V compares the algorithm complexity. Section VI contains the simulation results. Finally, Section VII concludes this paper.

*Notation List:* Let $\mathbb{C}^m$ and $\mathbb{C}^{m\times n}$ be the sets of $m$-dimensional complex vectors and $m \times n$ complex matrices. Denote by $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$, respectively the transpose, the complex conjugate, and the complex conjugate transpose. $\mathbf{I}_m$ and $\mathbf{0}_m$ denote the $m \times m$ identity and zero matrices; $\mathbf{0}_{m\times n}$ is the $m \times n$ zero matrix. The symbols $\otimes$ and $\odot$, respectively, stand for the Kronecker product [20, p-242] and Hadamard product [20, p-298]. For $\mathbf{x} \in \mathbb{C}^m$ and $\mathbf{X} \in \mathbb{C}^{m\times m}$, let $diag\{x_1, ..., x_m\}$ be an $m \times m$ diagonal matrix with $x_n$, $1 \le n \le m$, on the main diagonal, and let $Diag\{\mathbf{X}\}$ be an $m \times m$ diagonal matrix with the diagonal entries of $\mathbf{X}$ on the main diagonal. The notation $E\{y\}$ stands for the expected value of the random variable $y$. For $\mathbf{x} \in \mathbb{C}^m$ and $\mathbf{X} \in \mathbb{C}^{m\times n}$, denote by $\|\mathbf{x}\|$ the vector two-norm and $\|\mathbf{X}\|$ the matrix Frobenius norm.

## II. Preliminary

### A. Signal Model

Consider the discrete-time baseband model of a MIMO-SC system with *N* transmit antennas, *M* receive antennas, CP length *G*, and symbol block size *Q*. Let $h_{m,n}(k,l)$, $0 \le l \le L$, be the *l*th tap of the channel between the *n*th and *m*th transmit-receive antenna pair at time instant *k*, where *L* denotes the delay spread assumed common to all *MN* subchannels. Assuming $G \ge L$, after CP removal the received symbol block at the *m*th receive antenna can be expressed as

$$\mathbf{r}_m(t) = \sum_{n=1}^{N} \mathbf{H}_{m,n}(t)\mathbf{s}_n(t) + \mathbf{v}_m(t), \qquad (1)$$

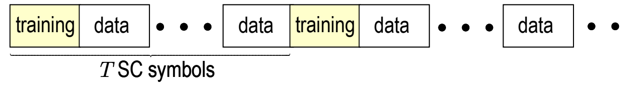


Fig. 1. Adopted frame structure.

where $s_n(t) := [s_{n,1}(t), ..., s_{n,Q}(t)]^T \in \mathbb{C}^Q$ is the *t*th symbol block sent from the *n*th transmit antenna, $\mathbf{v}_m(t) \in \mathbb{C}^Q$ is the channel noise vector, and $\mathbf{H}_{m,n}(t) \in \mathbb{C}^{Q\times Q}$ is the channel matrix whose *i*th column, denoted by $\mathbf{c}^{(i)}_{m,n}(t)$, is

$$\mathbf{c}^{(i)}_{m,n}(t) = \mathbf{J}^{i-1}\left[\mathbf{h}^{(i)T}_{m,n}(t)\ 0\ \cdots\ 0\right]^T,\ 1 \le i \le Q, \qquad (2)$$

in which $\mathbf{h}^{(i)}_{m,n}(t) := [h_{m,n}(\tilde{t} + (i-1)_Q, 0)\ \ h_{m,n}(\tilde{t} + (i)_Q, 1) \cdots h_{m,n}(\tilde{t} + (L+i-1)_Q, L)\ ]^T$ with $(\cdot)_Q$ denoting the modulo-$Q$ operation and $\tilde{t} := (t-1)(Q+G) + G$, and

$$\mathbf{J} := \begin{bmatrix} \mathbf{0}_{1\times(Q-1)} & 1 \\ \mathbf{I}_{Q-1} & \mathbf{0}_{(Q-1)\times 1} \end{bmatrix}. \qquad (3)$$

In case that the channel is time-invariant, i.e., $h_{m,n}(k,l) = h_{m,n}(l)$ for some $h_{m,n}(l)$, $\mathbf{H}_{m,n}(t)$ is a circulant matrix and symbol recovery can be done via the tone-by-tone FDE technique [1]. In the considered time-varying channel environment, $\mathbf{H}_{m,n}(t)$ is no longer circulant and the FDE facility is negated. Since there are no specific advantages of processing the data in the frequency domain, in this paper we will instead focus on the time-domain signal model (1) for receiver design; as will be shown next this can lead to an effective framework for addressing the robust signal recovery problem against imperfect channel knowledge. The following assumptions are made throughout the paper.

1) The number of receive antennas is equal to or greater than the number of transmit antennas, i.e., $M \ge N$.
2) The source symbols of each transmit antenna $s_n(t)$ is zero mean, unit-variance, and $E\{s_{n_1,q_1}(t_1)s_{n_2,q_2}(t_2)^*\} = \delta(n_1 - n_2)\delta(t_1 - t_2)\delta(q_1 - q_2)$, where $\delta(\cdot)$ is the Kronecker delta.
3) The elements of $\mathbf{v}_m(t)$'s are i.i.d. complex circular Gaussian with zero mean and variance $\sigma_v^2$.

### B. Time-Varying Channel Estimation & Equalization

This paper considers the burst-by-burst transmission such that *1)* each data burst consists of *T* symbol blocks, *2)* the leading block per burst serves as the training symbol for channel estimation, and *3)* the resultant channel estimate is used for receiver design to recover the subsequent *T*-1 source symbol blocks. A schematic description of the frame structure is shown in Figure 1. Since channel estimation and equalization is done on a burst-wise basis, in the sequel we shall focus on the initial burst. The time-varying channel estimation scheme adopted in this paper is briefly reviewed as below; the robust equalizer design which exploits the channel error characteristics will be discussed in Section IV. We assume that the MIMO channel is estimated by using the LS training technique [5], [6], which, in a time-varying environment, is known to yield the optimal estimate of the "averaged" channel impulse response within one symbol duration [21], namely,

$$\mathbf{h}_{m,n}^{(av)} := \left[\frac{1}{Q}\sum_{q=0}^{Q-1} h_{m,n}(G+q,0) \ \cdots \frac{1}{Q}\sum_{q=0}^{Q-1} h_{m,n}(G+q,L)\right]^T. \tag{4}$$

Associated with each $1 \le m \le M$ let us define $h_m^{(av)} = [h_{m,1}^{(av)T} \cdots h_{m,N}^{(av)T}]^T$, the resultant channel estimate is given by [21]

$$\widehat{\mathbf{h}}_m^{(av)} = \mathbf{h}_m^{(av)} + \underbrace{\mathbf{A}^+(\mathbf{i}_m + \mathbf{v}_m)}_{:=\Delta\mathbf{h}_m} \in \mathbb{C}^{N(L+1)\times 1}, 1 \le m \le M, \tag{5}$$

where $\mathbf{A} = [diag\{\mathbf{t}_1\}\mathbf{F}_L, ..., diag\{\mathbf{t}_N\}\mathbf{F}_L] \in \mathbb{C}^{Q\times N(L+1)}$ with $\mathbf{F}_L/\sqrt{Q}$ being the first $L$+1 columns of the $Q \times Q$ FFT matrix, $\mathbf{A}^+$ denotes the pseudoinverse of $\mathbf{A}$, $t_n \in \mathbb{C}^Q$ is the frequency-domain training sequence for the $n$th transmit antenna, and $\mathbf{i}_m = \sum_{n=1}^N \mathbf{i}_{m,n} \in \mathbb{C}^Q$ with the $q$th entry [21]

$$\left[\mathbf{i}_{m,n}\right]_q = \sum_{q_1=1}^{Q-1}\sum_{l=0}^{L+1}\left[\frac{1}{Q}\sum_{q_2=0}^{Q-1} h_{m,n}(G+q_2,l)e^{\frac{-j2\pi(q-q_1)l}{Q}}\right] \times [\mathbf{t}_n]_{(q-q_1)_Q}\, e^{\frac{-j2\pi q_1 q_2}{Q}}, 1 \le q \le Q. \tag{6}$$

Assuming that the channel variation is piecewise linear in time, $\widehat{\mathbf{h}}_{m,n}^{(av)}$ can also be treated as the optimal estimate of the channel parameters in the middle instant of the training period, i.e., $h_{m,n}(G+Q/2,l)$, $0 \le l \le L$ [21]. For fixed $1 \le m \le M$ and $1 \le q \le Q$ we assume that $h_{m,n}(q,l)$, $\forall n,l$, are independent circular complex Gaussian variables with zero-mean and variance $\sigma_l^2$. Then the channel estimation error $\Delta\mathbf{h}_m$ in (5) is zero mean with covariance matrix $\mathbf{R}_{\Delta\mathbf{h},m} := \mathbf{A}^+\left(E\{\mathbf{i}_m\mathbf{i}_m^H\} + \sigma_v^2\mathbf{I}_{N(L+1)}\right)(\mathbf{A}^+)^H \in \mathbb{C}^{N(L+1)\times N(L+1)}$, in which the $(q,q')$th entry of $E\{\mathbf{i}_m\mathbf{i}_m^H\}$, $1 \le q, q' \le Q$, can be directly computed from (6) as (7) (shown in the bottom of this page), where $R_h(q,l) = E\{h_{m,n}(k,l)h_{m,n}^*(k+q,l)\} = \sigma_l^2 J_0(2\pi f_d q T_s)$, with $J_0$, $f_d$, and $T_s$ respectively denoting the zeroth-order Bessel function, Doppler frequency, and sampling period. Moreover, assuming *i)* the taps among subchannels are mutually independent, and *ii)* the noise is spatially uncorrelated, we have

$$\mathbf{R}_{\Delta\mathbf{h},m_1,m_2} := E\{\Delta\mathbf{h}_{m_1}\Delta\mathbf{h}_{m_2}^H\} = \mathbf{0}_{N(L+1)}. \tag{8}$$

The results (7) and (8) will be used for robust equalizer design in Section IV.

## III. Group-Wise Symbol Detection: Perfect Channel Knowledge

### A. Motivation

This subsection highlights the motivation behind the proposed approach. Particularly, we will show that the time-domain channel matrix $\mathbf{H}_{m,n}(t)$ in (2) is imbedded with certain column-wise orthogonality structure; such an appealing feature will naturally lead to an inter-group interference cancellation framework followed by a low-complexity intra-group symbol recovery scheme. We shall first focus on the ideal case that the channel is perfectly known at the receiver, and will discuss more realistic situations in Section IV. To proceed, we first observe from (2) that, for $2 \le i \le Q$, the $i$th column $\mathbf{c}_{m,n}^{(i)}(t)$ of the channel matrix $H_{m,n}(t)$ is simply an $(i-1)$-step down-shifted version of the zero-padded channel impulse response vector $[\mathbf{h}_{m,n}^{(i)T}(t)\ 0\cdots 0]^T$. As a result, if the symbol block size $Q$ is chosen to be an integer multiple of $G$, i.e., $Q = PG$ for some positive integer $P$, then for each fixed $1 \le i \le G$ we have

$$\mathbf{c}_{m,n}^{(i+p_1G)}(t)^H\mathbf{c}_{m,n}^{(i+p_2G)}(t) = 0 \ for\ 0 \le p_1 \ne p_2 \le P-1, \tag{9}$$

since the locations of the respective nonzero entries never overlap; a schematic description of such an orthogonality relation is depicted in Figure 2. Equation (9) suggests that the $Q$ columns of $\mathbf{H}_{m,n}(t)$ can be divided into $G$ groups of orthogonal vectors as follows

$$\mathbf{H}_{m,n}^{(i)}(t) = \left[\ \mathbf{c}_{m,n}^{(i)}(t)\ \ \mathbf{c}_{m,n}^{(i+G)}(t) \cdots \mathbf{c}_{m,n}^{(i+(P-1)G)}(t)\ \right] \in \mathbb{C}^{Q\times P},\ 1 \le i \le G. \tag{10}$$

To further exploit the benefit from the orthogonality condition (9), for each fixed $1 \le i \le G$ let us stack $H_{m,n}^{(i)}(t)$ for all $1 \le m \le M$ and $1 \le n \le N$ to form

$$\mathbf{H}_n^{(i)}(t) := \left[\mathbf{H}_{1,n}^{(i)}(t)^T\ \ \mathbf{H}_{2,n}^{(i)}(t)^T\ \cdots\ \mathbf{H}_{M,n}^{(i)}(t)^T\right]^T \in \mathbb{C}^{MQ\times P},$$

and

$$\mathbf{H}^{(i)}(t) := \left[\ \mathbf{H}_1^{(i)}(t)\ \ \mathbf{H}_2^{(i)}(t) \cdots \mathbf{H}_N^{(i)}(t)\ \right] \in \mathbb{C}^{MQ\times NP},\ 1 \le i \le G, \tag{11}$$

Collecting the $M$ received signal blocks $\mathbf{r}_m$'s in (1) into a vector, the overall input-output relation can be rearranged as

$$\mathbf{r}(t) := \left[\mathbf{r}_1(t)^T \cdots \mathbf{r}_M(t)^T\right]^T = \sum_{i=1}^G \mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \underbrace{[\mathbf{v}_1(t)^T \cdots \mathbf{v}_M(t)^T]^T}_{:=\mathbf{v}(t)}, \tag{12}$$

where $\mathbf{s}^{(i)}(t) := \left[s_1^{(i)}(t)^T \cdots s_N^{(i)}(t)^T\right]^T \in \mathbb{C}^{NP}$, $\mathbf{s}_n^{(i)}(t) := \left[s_{n,i}(t)\ s_{n,i+G}(t)\ \cdots s_{n,i+G(P-1)}(t)\right]^T \in \mathbb{C}^P$. Toward symbol extraction based on (12) we propose to first design

$$\left[E\{\mathbf{i}_m\mathbf{i}_m^H\}\right]_{q,q'} = \frac{1}{Q^2}\sum_{n=1}^N \sum_{\substack{q_1=0,q_2=0\\ q_1\ne q,q_2\ne q'}}^{Q-1} \sum_{l=0}^{L}\sum_{q_3,q_4=0}^{Q-1} \mathbf{R}_h(q_3-q_4,l)e^{\frac{-j2\pi[(q-q_1)-(q'-q_2)]l}{Q}}[\mathbf{t}_n]_{(q-q_1)_Q}[\mathbf{t}_n]^*_{(q'-q_2)_Q}\, e^{\frac{-j2\pi(q_1-q_2)l}{Q}}. \tag{7}$$

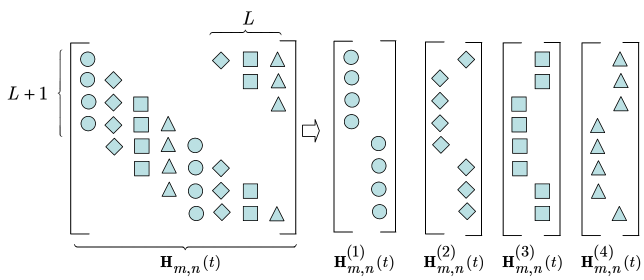


Fig. 2. A schematic description of the orthogonality condition (3.1) with $G = 4$.

an *inter-group interference suppression* matrix $\mathbf{W}^{(j)}(t) \in \mathbb{C}^{MQ\times NQ}$ such that

$$\begin{aligned}\mathbf{r}^{(j)}(t) &:= \mathbf{W}^{(j)}(t)^H\mathbf{r}(t) \approx \mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)\\ &= \begin{bmatrix} \mathbf{H}_1^{(j)}(t)^H\mathbf{H}_1^{(j)}(t) & \cdots & \mathbf{H}_1^{(j)}(t)^H\mathbf{H}_N^{(j)}(t)\\ \vdots & \ddots & \vdots \\ \mathbf{H}_N^{(j)}(t)^H\mathbf{H}_1^{(j)}(t) & \cdots & \mathbf{H}_N^{(j)}(t)^H\mathbf{H}_N^{(j)}(t)\end{bmatrix}\mathbf{s}^{(j)}(t),\end{aligned} \tag{13}$$

where the last equality in (13) follows directly from (11), and then recover $\mathbf{s}^{(j)}(t)$ via the signal model (13). Benefiting from the orthogonality condition (9), the matched-filtered channel matrix $\mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t)$ in (13) exhibits an appealing structure. Indeed, since the nonzero entries among the $P$ columns of $\mathbf{H}^{(i)}_{m,n}(t)$ in (10) do not overlap and, for each $1 \le i \le G$, $\mathbf{H}^{(i)}_n(t)$ is obtained by stacking $\mathbf{H}^{(i)}_{m,n}(t)$ over all $1 \le m \le M$, it is easy to check that

$$\mathbf{H}^{(i)}_{n_1}(t)^H\mathbf{H}^{(i)}_{n_2}(t) = \mathbf{D}^{(i)}_{n_1n_2}(t),\ 1 \le n_1, n_2 \le N, \tag{14}$$

where $\mathbf{D}^{(i)}_{n_1n_2}(t) \in \mathbb{C}^{P\times P}$ is diagonal. Equation (14) implies that, for subsequent *intra-group symbol recovery* through separating the $NP$ coupled streams in (13), the problem is reduced to solving a set of $P$ independent linear equations, each with dimension $N \times N$. Such a decoupled nature reduces computations, especially when the block length $Q$ (and hence $P = Q/G$) is large; it can also limit the error propagation effect in the symbol recovery stage. The key challenge of the proposed two-stage equalization scheme is the design of the interference suppression matrix $\mathbf{W}^{(j)}(t)$ for fulfilling (13); the underlying mathematical formulation is discussed next.

### B. Solution Based on Constrained Optimization

Based on (13), the linear weighting matrix $\mathbf{W}^{(j)}(t)$ for recovering the *j*th symbol group should be designed to minimize the inter-group interference power, and then extract the desired signal component through space-time matched filtering. A typical technique toward fulfilling such a two-fold task is through constrained optimization [7], [8]; more precisely, we shall design $\mathbf{W}^{(j)}(t)$ by solving the following problem

$$\begin{aligned}&\min_{\mathbf{W}^{(j)}(t)} \ E\left\{\left\|\mathbf{W}^{(j)}(t)^H\left(\sum_{i=1,i\ne j}^{G}\mathbf{H}^{(i)}(t)s^{(i)}(t)+\mathbf{v}(t)\right)\right\|^2\right\},\\ &\quad s.t.\ \mathbf{W}^{(j)}(t)^H\mathbf{H}^{(j)}(t) = \mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t);\end{aligned} \tag{15}$$

the recovery of the entire symbol vector $\mathbf{s}(t)$ is then done group-wise via repeatedly solving (15) for $1 \le j \le G$. By using the standard Lagrangian multiplier technique [22, p-215], the solution to (15) is given by

$$\begin{aligned}\mathbf{W}^{(j)}(t) =& \mathbf{R}_I^{(j)}(t)^{-1}\mathbf{H}^{(j)}(t)\left(\mathbf{H}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)^{-1}\mathbf{H}^{(j)}(t)\right)^{-1}\\ &\cdot \mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t),\end{aligned} \tag{16}$$

where $\mathbf{R}_I^{(j)}(t) := \sum_{i=1,i\ne j}^{G}\mathbf{H}^{(i)}(t)\mathbf{H}^{(i)}(t)^H + \sigma_n^2\mathbf{I}_{MQ}$. An alternative approach to solving (15) lies in transforming the constrained optimization formulation into an equivalent *unconstrained* setup via the GSC principle [9]. This relies on the following decomposition of $\mathbf{W}^{(j)}(t)$:

$$\mathbf{W}^{(j)}(t) = \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}^{(j)}(t), \tag{17}$$

where $\mathbf{H}^{(j)}(t) \in \mathbb{C}^{MQ\times NP}$ represents the *non-adaptive* portion which verifies the desired space-time matched filtering constraint, $\mathbf{B}^{(j)}(t) \in \mathbb{C}^{MQ\times(MQ-NP)}$ is the signal blocking matrix with $\mathbf{B}^{(j)}(t)^H\mathbf{H}^{(j)}(t) = \mathbf{0}_{(MQ-NP)\times NP}$, and $\mathbf{U}^{(j)}(t) \in \mathbb{C}^{(MQ-NP)\times NP}$ is the *adaptive component* which forms the remaining free parameters to be determined. With (10), the equalized output becomes

$$\mathbf{z}^{(j)}(t) := \mathbf{W}^{(j)}(t)^H\mathbf{r}(t) = \mathbf{z}_d^{(j)}(t) - \mathbf{U}^{(j)}(t)^H\mathbf{z}_b^{(j)}(t), \tag{18}$$

where

$$\begin{aligned}\mathbf{z}_d^{(j)}(t) :=& \mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)\\ &+ \underbrace{\mathbf{H}^{(j)}(t)^H\left(\sum_{i=1,i\ne j}^{G}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t)+\mathbf{v}(t)\right)}_{:=\mathbf{i}^{(j)}(t)},\end{aligned} \tag{19}$$

and

$$\mathbf{z}_b^{(j)}(t) := \mathbf{B}^{(j)}(t)^H\left(\sum_{i=1,i\ne j}^{G}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t)+\mathbf{v}(t)\right). \tag{20}$$

Since the desired signal component, namely, $\mathbf{H}^{(j)}(t)^H\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$, in the matched filtered branch $\mathbf{z}_d^{(j)}(t)$ is contaminated by $\mathbf{i}^{(j)}(t)$, toward effective interference suppression equation (18) suggests that the matrix $\mathbf{U}^{(j)}(t)$ should be chosen so that $\mathbf{U}^{(j)}(t)^H\mathbf{z}_b^{(j)}(t)$ is as close to $\mathbf{i}^{(j)}(t)$ as possible; we can thus determine $\mathbf{U}^{(j)}(t)$ via

$$\min_{\mathbf{U}^{(j)}(t)} E\left\{\left\|\mathbf{i}^{(j)}(t) - \mathbf{U}^{(j)}(t)^H\mathbf{z}_b^{(j)}(t)\right\|^2\right\}. \tag{21}$$

By following the standard procedure [10], the solution to (21) is given by

$$\begin{aligned}\mathbf{U}_{opt}^{(j)}(t) =& \left(\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{B}^{(j)}(t)\right)^{-1}\\ &\cdot \mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t);\end{aligned} \tag{22}$$

the resultant optimal GSC weight is thus[1]

$$\begin{aligned}\mathbf{W}_{opt}^{(j)}(t) =&\mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}_{opt}^{(j)}(t)\\ =&\mathbf{H}^{(j)}(t)-\\ &\mathbf{B}^{(j)}(t)\left(\mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{B}^{(j)}(t)\right)^{-1}\\ &\cdot \mathbf{B}^{(j)}(t)\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t).\end{aligned} \tag{23}$$

We note that *i)* the GSC scheme solves a constrained optimization problem via a simple unconstrained formulation, *ii)* solutions (16) and (23) are obtained based on the crucial perfect channel knowledge assumption; when channel parameter mismatch occurs due to imperfect estimation and time variation, they are only suboptimal because the formulations do not take into account channel error mitigation, *iii)* as will be seen in the next section, the GSC framework is advantageous in that it allows us to directly model the channel mismatch effect into the system equations for facilitating robust equalizer design.

## IV. Proposed Robust GSC Equalizer

Given only the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$, $1 \le i \le G$, which is acquired through training at $t = 1$, one may simply choose to modify solution (16) via replacing $\mathbf{H}^{(i)}(t)$ by $\widehat{\mathbf{H}}_{av}^{(i)}$ to get the time-invariant equalizer

$$\widehat{\mathbf{W}}^{(j)} = (\widehat{\mathbf{R}}_I^{(j)})^{-1}\widehat{\mathbf{H}}_{av}^{(j)}\left(\widehat{\mathbf{H}}_{av}^{(j)H}(\widehat{\mathbf{R}}_I^{(j)})^{-1}\widehat{\mathbf{H}}_{av}^{(j)}\right)^{-1}\widehat{\mathbf{H}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)}, \tag{24}$$

where $\widehat{\mathbf{R}}_I^{(j)} = \sum_{i=1,i\neq j}^{G}\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H} + \sigma_n^2\mathbf{I}_{MQ}$, and use (24) once and for all toward subsequent symbol recovery. Such a strategy, even though quite simple, fails to combat channel mismatch and could incur very poor equalization performance. The GSC principle, on the contrary, provides a very natural and effective framework for robust time-varying equalizer design, as is shown in this section. We will first introduce the problem formulation in Section IV-A, and then derive the solution in Section IV-B. Some discussions regarding the proposed robust scheme are given in Section IV-C.

### A. Problem Formulation

Recall that the mechanism of GSC filter (17) basically involves space-time signal matched filtering and signal blocking followed by interference suppression through (21). While the signal combining and nulling components ($\mathbf{H}^{(j)}(t)$ and $\mathbf{B}^{(j)}(t)$, respectively) are immediately fixed upon the (possibly imperfect) knowledge of $\mathbf{H}^{(j)}(t)$, the adaptive portion $\mathbf{U}^{(j)}(t)$ can nonetheless be allowed to be time-varying, and is then designed to tackle the channel parameter mismatch effects caused by time variation and estimation error. It is such inherent channel tracking capability that makes GSC principle a promising approach in the considered scenario.

Specifically, when only a channel estimate $\widehat{\mathbf{H}}_{av}^{(j)}$ is available, exact signal matched filtering over $2 \le t \le T$ is impossible; the best one can do, however, is to linearly combine $\mathbf{H}^{(j)}(t)s^{(j)}(t)$ just with $\widehat{\mathbf{H}}_{av}^{(j)}$ to get the approximation $\widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$. This implies that the non-adaptive portion of the GSC weight should be set as $\widehat{\mathbf{H}}_{av}^{(j)}$ throughout $2 \le t \le T$ and, in turn, the blocking matrix is likewise fixed according to the relation $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$. Hence, given the channel knowledge $\widehat{\mathbf{H}}_{av}^{(j)}$ only, the signal matching and blocking matrices are restricted to be time-invariant. Nevertheless, to reliably recover the desired signal against background time-varying interference, the adaptive component must account for the temporal variation. This thus suggests the following modified GSC decomposition

$$\widehat{\mathbf{W}}^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)} - \widehat{\mathbf{B}}_{av}^{(j)}\overline{\mathbf{U}}^{(j)}(t). \tag{25}$$

With (25), the equalized output instead reads

$$\overline{\mathbf{z}}^{(j)}(t) = \widehat{\mathbf{W}}^{(j)}(t)^H\mathbf{r}(t) = \overline{\mathbf{z}}_d^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^H\overline{\mathbf{z}}_b^{(j)}(t), \tag{26}$$

where

$$\begin{aligned}\overline{\mathbf{z}}_d^{(j)}(t) :=&\widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)\\ &+ \underbrace{\widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{G}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{v}(t)}_{:=\overline{\mathbf{i}}_d^{(j)}(t)},\end{aligned} \tag{27}$$

is the (approximate) space-time matched filtered signal, and

$$\begin{aligned}\overline{\mathbf{z}}_b^{(j)}(t) :=&\widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)\\ &+ \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{G}\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{v}(t),\end{aligned} \tag{28}$$

is the corresponding blocking component. Due to the out-of-date channel knowledge, the signal of interest in the $\overline{\mathbf{z}}_d^{(j)}(t)$ is non-coherently combined and is corrupted by the interference $\overline{\mathbf{i}}_d^{(j)}(t)$; also, since the blocking matrix is determined via $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$, and hence $\widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t) \neq \mathbf{0}_{(MQ-NP)\times NP}$ in general, there is a signal leakage term $\widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{H}^{(j)}(t)\mathbf{s}^{(j)}(t)$ into the blocking branch $\overline{\mathbf{z}}_b^{(j)}(t)$. Toward signal recovery against interference, a natural strategy as suggested by the GSC principle is to treat $\overline{\mathbf{z}}_b^{(j)}(t)$ in (28) as an aggregate interference and to design $\overline{\mathbf{U}}^{(j)}(t)$ such that $\overline{\mathbf{U}}^{(j)}(t)^H\overline{\mathbf{z}}_b^{(j)}(t)$ is best close to $\overline{\mathbf{i}}_d^{(j)}(t)$.

For this we shall first note that in (27) and (28) only the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$'s (acquired through training at $t = 1$) are available but the true channel matrices $\mathbf{H}^{(i)}(t)$'s are unknown: the mismatches between $\widehat{\mathbf{H}}_{av}^{(i)}$'s and $\mathbf{H}^{(i)}(t)$'s are due to time variation as well as channel estimation errors. To facilitate subsequent analysis we must seek for explicit rules linking the unknown $\mathbf{H}^{(i)}(t)$ to the channel estimate $\widehat{\mathbf{H}}_{av}^{(i)}$. A commonly used model which specifies such channel parameter deviation can be found in [24], [25], and in terms of the current matrix formulation it reads

$$\mathbf{H}^{(i)}(t) = \mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)} + \delta\mathbf{H}^{(i)}(t), 1 \le i \le G, \tag{29}$$

where $\mathbf{D}_c^{(i)}(t)$ is an $MQ\times MQ$ diagonal matrix depending on the time-varying channel characteristics, and $\delta\mathbf{H}^{(i)}(t)$ models

[1]The resultant GSC based solution will coincide with the optimal one derived under the original constrained-optimization based formulation if $\mathrm{rank}\left(\mathbf{B}^{(j)}(t)\right) = MQ - NP$ [23]; this condition is fulfilled if the columns of $\mathbf{B}^{(j)}(t)$ form an orthonormal basis for the left null-space associated with $\mathbf{H}^{(j)}(t)$ .

the temporal channel variation whose entries are assumed to be zero-mean Gaussian random variables (explicit formulae for $\mathbf{D}_c^{(i)}(t)$ and the covariance of $\delta\mathbf{H}^{(i)}(t)$ are given in Appendix A). By using (29) we can rewrite (27) and (28) as (30) (shown in the bottom of this page), and

$$\overline{\mathbf{z}}_b^{(j)}(t) = \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\mathbf{s}^{(j)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{B}}_{av}^{(j)H}\mathbf{v}(t). \quad (31)$$

Based on (30) and (31), we specifically propose to design $\overline{\mathbf{U}}^{(j)}(t)$ by minimizing the following cost function

$$\overline{J}(t) = E\left\{\left\|\overline{\mathbf{i}}^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^H\overline{\mathbf{z}}_b^{(j)}(t)\right\|^2\right\}, \quad (32)$$

where the expectation is taken with respect to (w.r.t.) the source symbol, background noise, channel estimation error, and channel temporal variation (assuming all are mutually independent).

### B. Optimal Solution

Let us expand (32) into

$$\overline{J}(t) = \overline{\mathbf{U}}^{(j)}(t)^H E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) - \overline{\mathbf{U}}^{(j)}(t)^H E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} - E\left\{\overline{\mathbf{i}}^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) + E\left\{\overline{\mathbf{i}}^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}. \quad (33)$$

Since for a given matrix $\mathbf{A}$ we have

$$\partial Tr(\overline{\mathbf{U}}^{(j)}(t)^H\mathbf{A}\mathbf{U}^{(j)}(t))/\partial\overline{\mathbf{U}}^{(j)}(t) = \left(\overline{\mathbf{U}}^{(j)}(t)^H\mathbf{A}\right)^T,$$
$$\partial Tr(\mathbf{A}\overline{\mathbf{U}}^{(j)}(t))/\partial\overline{\mathbf{U}}^{(j)}(t) = \mathbf{A}^T,$$

and $\partial Tr(\overline{\mathbf{U}}^{(j)}(t)^H\mathbf{A})/\partial\overline{\mathbf{U}}^{(j)}(t) = \mathbf{0}$ [26], taking the first-order partial derivative of $\overline{J}(t)$ w.r.t. $\overline{\mathbf{U}}^{(j)}(t)$ yields

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}\overline{\mathbf{U}}^{(j)}(t) = E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}. \quad (34)$$

The optimal $\overline{\mathbf{U}}^{(j)}(t)$ can then be obtained by solving the matrix equation (34), provided that the covariance matrices $E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\}$ and $E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\}$ are available. With (30), (31), and by taking expectation w.r.t. source symbols, channel noise, and channel temporal variation characterized via (29), we can reach the following intermediate expressions (see Appendix A for detailed proof):

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\textstyle\sum_{i=1}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + N\mathbf{I}_M\otimes\mathbf{D}_Q(t) + \sigma_v^2\mathbf{I}_{MQ}\right)\widehat{\mathbf{B}}_{av}^{(j)}\right\}, \quad (35)$$

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = E_e\left\{\widehat{\mathbf{B}}_{av}^{(j)H}\left(\textstyle\sum_{i=1,i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + N\mathbf{I}_M\otimes\mathbf{D}_Q(t) + \sigma_v^2\mathbf{I}_{MQ}\right)\widehat{\mathbf{H}}_{av}^{(j)}\right\}, \quad (36)$$

where

$$\mathbf{D}_Q(t) := diag\left\{\textstyle\sum_{l=0}^{L}\sigma_l^2\left(1-\left|\rho(\overline{t},l)\right|^2\right), \ldots, \sum_{l=0}^{L}\sigma_l^2\left(1-\left|\rho(\overline{t}+Q-1,l)\right|^2\right)\right\}, \quad (37)$$

with $\rho(k,l) := \mathbf{R}_h(k,l)(\sigma_l\widehat{\sigma}_l)^{-1}$, $\widehat{\sigma}_l^2 = \sigma_l^2 + [\mathbf{R}_{\Delta h,m}]_{(n-1)(L+1)+l+1,(n-1)(L+1)+l+1}$, $\overline{t} := (t-1)(G+Q) - Q/2$, and $E_e\{\cdot\}$ denotes the expectation involving channel estimation error yet to be carried out. To explicitly determine the expectations in (35) and (36) we must further seek for tractable relations linking the channel estimates ($\widehat{\mathbf{H}}_{av}^{(i)}$ for $1 \le i \le G$ and $\widehat{\mathbf{B}}_{av}^{(j)}$) and the background estimation errors. While $\widehat{\mathbf{H}}_{av}^{(i)}$ can be directly modeled as the actual channel parameter corrupted by the errors, namely,

$$\widehat{\mathbf{H}}_{av}^{(i)} = \mathbf{H}_{av}^{(i)} + \Delta\mathbf{H}_{av}^{(i)}, 1 \le i \le G, \quad (38)$$

an exact expression of the blocking component $\widehat{\mathbf{B}}_{av}^{(j)}$ in terms of $\Delta\mathbf{H}_{av}^{(j)}$ remains formidable to characterize, since it is determined through $\widehat{\mathbf{B}}_{av}^{(j)H}\widehat{\mathbf{H}}_{av}^{(i)} = \mathbf{0}_{(MQ-NP)\times NP}$ and, thus, is obtained as an orthonormal basis of the left null space of $\widehat{\mathbf{H}}_{av}^{(i)}$. To resolve this difficulty, we will leverage the perturbation technique [12], [13] to get an approximate, but analytic, relation among $\widehat{\mathbf{B}}_{av}^{(j)}$ and the channel estimation error $\Delta\mathbf{H}_{av}^{(j)}$. The result is shown in the next lemma (the proof can be found in [12]).

*Lemma 4.1:* Let $\mathbf{H}_{av}^{(j)} = \mathbf{U}_h^{(j)}\Sigma_h^{(j)}\mathbf{V}_h^{(j)H}$ be a singular value decomposition of the actual channel matrix $\mathbf{H}_{av}^{(j)}$. The blocking matrix $\widehat{\mathbf{B}}_{av}^{(j)}$ can be approximated by

$$\widehat{\mathbf{B}}_{av}^{(j)} \approx \mathbf{B}_{av}^{(j)} - \underbrace{\mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}\Delta\mathbf{H}_{av}^{(j)H}\mathbf{B}_{av}^{(j)}}_{:=\Delta\mathbf{B}_{av}^{(j)}}. \quad (39)$$

□

By substituting $\widehat{\mathbf{B}}_{av}^{(j)}$ in (39) into (35) and (36), we have the following main results (see Appendix B for detailed derivations).

$$\overline{\mathbf{z}}_d^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\mathbf{s}^{(j)}(t) + \underbrace{\widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1,i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\mathbf{s}^{(i)}(t) + \widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{v}(t)}_{:=\overline{\mathbf{i}}^{(j)}(t)}. \quad (30)$$

*Proposition 4.1:* The covariance matrices involved in (34) can be expressed as

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\Big(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\Big)\mathbf{B}_{av}^{(j)}, \tag{40}$$

$$E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = \mathbf{B}_{av}^{(j)H}\Big(\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\Big)\mathbf{H}_{av}^{(j)}, \tag{41}$$

where

$$\mathbf{R}_L^{(j)}(t) := \mathbf{D}_c^{(j)}(t)\mathbf{H}_{av}^{(j)}\mathbf{H}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)^H, \tag{42}$$

$$\overline{\mathbf{R}}_I^{(j)}(t) := \sum_{i=1, i\neq j}^{G} \mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ}, \tag{43}$$

and $\mathbf{R}_{e,1}^{(j)}(t) := \sum_{i=1}^{4}\mathbf{X}_i(t)$ and $\mathbf{R}_{e,2}^{(j)}(t) := \sum_{k=1}^{3}\mathbf{Y}_k(t)$, in which the component matrices $\mathbf{X}_i(t)$'s and $\mathbf{Y}_k(t)$'s are defined in Table I.

□

Based on (34), (40), and (41), the robust solution of $\overline{\mathbf{U}}^{(j)}(t)$, which is *on average* optimal for mitigating channel temporal variation and estimation error, is thus

$$\overline{\mathbf{U}}_{opt}^{(j)}(t) = \left[\mathbf{B}_{av}^{(j)H}\left(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{B}_{av}^{(j)}\right]^{-1} \times \mathbf{B}_{av}^{(j)H}\left(\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{H}_{av}^{(j)}. \tag{44}$$

In the practical situation when only a channel estimate is available, the sampled-version of the overall robust GSC weighting matrix is accordingly given by

$$\widehat{\mathbf{W}}_r^{(j)}(t) = \widehat{\mathbf{H}}_{av}^{(j)} - \widehat{\mathbf{B}}_{av}^{(j)}\widehat{\mathbf{U}}_{opt}^{(j)}(t). \tag{45}$$

### C. Discussions

1) The proposed design formulation aims for joint mitigation of channel temporal variation and estimation error. Compared with (22) obtained under exact channel knowledge, the main distinctive feature of the proposed robust scheme (44) lies in replacing the two $\mathbf{R}_I^{(j)}(t)$'s in (22) respectively by the "composite" interference covariance matrices $\mathbf{R}_L^{(j)}(t)+\overline{\mathbf{R}}_I^{(j)}(t)+\mathbf{R}_{e,1}^{(j)}(t)+N\mathbf{I}_M\otimes \mathbf{D}_Q(t)$ and $\overline{\mathbf{R}}_I^{(j)}(t) + \mathbf{R}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)$. We should note that (i) the common term $\overline{\mathbf{R}}_I^{(j)}(t)$ can be regarded as the analogue of $\mathbf{R}_I^{(j)}(t)$ with the availability of $\mathbf{H}_{av}^{(i)}$'s only (rather than the exact $\mathbf{H}^{(j)}(t)$), (ii) the term $\mathbf{R}_L^{(j)}(t)$ arises from signal leakage into the blocking branch due to channel mismatch, (iii) the matrices $\mathbf{R}_{e,1}^{(j)}(t)$ and $\mathbf{R}_{e,2}^{(j)}(t)$ are due to the aggregate impacts caused by channel estimation errors, and (iv) the quantity $N\mathbf{I}_M \otimes \mathbf{D}_Q(t)$ accounts for the background channel temporal variation.
2) Another scheme for estimating time-varying channels is to treat the channel estimates for two consecutive data bursts as the end points, and further leverage linear interpolation for acquiring the channel information within the entire time frame [21]; this approach applies whenever the channel varies linearly w.r.t. time. In such a scenario, the impacts due to channel temporal variation would be largely reduced and channel estimation error, instead, becomes the dominant adverse factor to be combated. The proposed design strategy can be used for constructing an associated error-resistant GSC equalizer. Indeed, the cost function (32) instead involves only the averages w.r.t. source, noise, and channel estimation errors; by following the procedures shown in Section IV-B, the sampled-version of the resultant robust GSC weight can be obtained by (the derivations are highlighted in Appendix C)
$$\widehat{\mathbf{W}}_a^{(j)}(t) = \widehat{\mathbf{H}}^{(j)}(t) - \left(\widehat{\mathbf{B}}^{(j)}(t)^H\left(\widehat{\mathbf{R}}_I^{(j)}(t) + \widehat{\mathbf{R}}_e(t)\right)\widehat{\mathbf{B}}^{(j)}(t)\right)^{-1} \cdot \widehat{\mathbf{B}}^{(j)}(t)^H\widehat{\mathbf{R}}_I^{(j)}(t)\widehat{\mathbf{H}}^{(j)}(t), \tag{46}$$
where $\widehat{\mathbf{H}}^{(j)}(t)$ and $\widehat{\mathbf{R}}_I^{(j)}(t)$ respectively denote the estimates of $\mathbf{H}^{(j)}(t)$ and $\mathbf{R}_I^{(j)}(t)$ (through linear interpolation), and $\widehat{\mathbf{R}}_e(t)$ is the estimated version of $\mathbf{R}_e(t)$ defined in Table I.
3) The proposed group-wise symbol recovery scheme can be directly combined with the successive interference cancellation (SIC) mechanism for further performance improvement; the resultant solutions in each processing layer is essentially of the form (45), except that in $\overline{\mathbf{R}}_I^{(j)}(t)$, $\mathbf{R}_{e,1}^{(j)}(t)$, and $\mathbf{R}_{e,2}^{(j)}(t)$ the signature matrices corresponding to the previously detected signal components are removed. Detailed description of the SIC based implementation is listed in Table II.
4) We finally note that the analytic expression of $\widehat{\mathbf{B}}_{av}^{(j)}$ in Lemma 4.1 involves only the first-order error perturbation. Even though more accurate approximation to $\widehat{\mathbf{B}}_{av}^{(j)}$ can be obtained by incoPersonNamerporating higher order terms [27], the analysis would however become intractable.

## V. Algorithm Complexity

The computation of the proposed GSC weighting matrix basically involves solving $\widehat{\mathbf{B}}_{av}^{(j)H}\hat{\mathbf{H}}_{av}^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$ for the blocking matrix $\widehat{\mathbf{B}}_{av}^{(j)}$ followed by inverting an $(MQ - NP) \times (MQ - NP)$ matrix. Due to the sparse structure of $\hat{\mathbf{H}}_{av}^{(j)}$ (see (2) and (11)), the computation of $\widehat{\mathbf{B}}_{av}^{(j)}$ only requires $(4MN(ML-2N)(L+1)/G)\,Q$ flops (i.e., $O(Q)$). The complexity for matrix inversion can be further reduced by leveraging the conjugate gradient [28] based reduced-rank (RR) implementation [29] (listed in Table III). Table IV shows the flop counts (measured in terms of the number of complex-valued additions and multiplications) of the GSC method, GSC with RR implementation as well as two comparative

TABLE I
FORMULAE OF $\mathbf{X}^{(j)}(t)$, $\mathbf{X}^{(i,j)}(t)$, $\mathbf{X}_1(t)$~$\mathbf{X}_4(t)$, $\mathbf{Y}_1(t)$~$\mathbf{Y}_3(t)$, AND $\mathbf{R}_e(t)$.

| | | | |
|---|---|---|---|
| $\mathbf{X}^{(j)}(t)$ | $\mathbf{V}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{U}_h^{(j)H}(\mathbf{R}_L^{(j)}(t)+\overline{\mathbf{R}}_I^{(j)}(t)+N\mathbf{I}_M\otimes\mathbf{D}_Q(t))\times \mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}$ | $\mathbf{Y}_3(t)$ | $\sum_{i=1,i\neq j}^{G+1}\mathbf{D}_{i,j}\left(\mathbf{X}^{(i,j)}(t)^H\right)\mathbf{D}_c^{(i)}(t)^H$ |
| $\mathbf{X}^{(i,j)}(t)$ | $\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)\mathbf{U}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{V}_h^{(j)H}$ | $\mathbf{R}_e(t)$ | $\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))+\left(\overline{\mathbf{D}}_{j,j}(\mathbf{A}^{(j)}(t))-\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))\right)\mathbf{D}_T(t)$ $+\mathbf{D}_T(t)\left(\overline{\mathbf{D}}_{j,j}^H(\mathbf{A}^{(j)}(t))-\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))\right)$ $+\mathbf{D}_T(t)\left(2\mathbf{D}_{j,j}(\mathbf{A}^{(j)}(t))-\overline{\mathbf{D}}_{j,j}^H(\mathbf{A}^{(j)}(t))-\overline{\mathbf{D}}_{j,j}(\mathbf{A}^{(j)}(t))\right)\mathbf{D}_T(t)$ |
| $\mathbf{X}_1(t)$ | $\sum_{i=1}^{G+1}\mathbf{D}_c^{(i)}(t)\mathbf{D}_H^{(i)}\mathbf{D}_c^{(i)}(t)^H$ | | |
| $\mathbf{X}_2(t)$ | $\mathbf{D}_{j,j}\left(\mathbf{X}^{(j)}(t)\right)$ | $\mathbf{A}^{(j)}(t)$ | $\mathbf{V}_h^{(j)}\Sigma_h^{(j)-1}\mathbf{U}_h^{(j)H}\mathbf{R}_I^{(j)}(t)\mathbf{U}_h^{(j)}\Sigma_h^{(j)^{-1}}\mathbf{V}_h^{(j)H}$ |
| $\mathbf{X}_3(t)$ | $\sum_{i=1}^{G+1}\mathbf{D}_c^{(i)}\mathbf{D}_{i,j}\left(\mathbf{X}^{(i,j)}(t)\right)$ | $\left[\overline{\mathbf{D}}_{j,j}(\mathbf{Z}_1)\right]^{(m,m)}$ | $\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\overline{\mathbf{J}}_i\left(\mathbf{Z}_1^{(n_1,n_2)}\otimes Diag\left\{\overline{\mathbf{R}}_{\Delta\mathbf{h},m}^{(n_1,n_2)},\mathbf{0}_{L+1}\cdots\mathbf{0}_{L+1}\right\}\right)\overline{\mathbf{J}}_j^T$ |
| $\mathbf{X}_4(t)$ | $\mathbf{X}_3(t)^H$ | $\overline{\mathbf{R}}_{\Delta\mathbf{h},m}$ | $\mathbf{A}^+\mathbf{K}(\mathbf{A}^+)^H$ |
| $\mathbf{Y}_1(t)$ | $\sum_{i=1,i\neq j}^{G+1}\mathbf{D}_c^{(i)}(t)\mathbf{D}_H^{(i)}\mathbf{D}_c^{(i)}(t)^H$ | $[\mathbf{K}]_{q,q'}$ | $\frac{1}{Q^2}\sum_{n=1}^{N}\sum_{\substack{q_1=0,q_2=0\\ q_1\neq q,q_2\neq q'}}^{Q-1}\sum_{l=0}^{L}\sum_{q_3,q_4=0}^{Q-1}\mathbf{R}_h(T(G+Q)+q_3-q_4,l)e^{\frac{-j2\pi(q_1-q_2)l}{Q}}\cdot [\mathbf{t}_n]_{(q-q_1)_Q}[\mathbf{t}_n]^*_{(q'-q_2)_Q}e^{\frac{-j2\pi[(q-q_1)-(q'-q_2)]l}{Q}}$ |
| $\mathbf{Y}_2(t)$ | $\sum_{i=1,i\neq j}^{G+1}\mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}\mathbf{K}_{i,j}\left(\mathbf{D}_c^{(i)}(t)^H\right)$ | | |
| $\left[\mathbf{D}_H^{(i)}\right]^{(m,m)}$ | $\sum_{n=1}^{N}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)(G+1)-1}Diag\left\{\mathbf{R}_{\Delta\mathbf{h},m}^{(n,n)},\mathbf{0}_{L+1},\ldots,\mathbf{0}_{L+1}\right\}\left(\mathbf{J}^{i+(p-1)(G+1)-1}\right)^T$ | | |
| $\left[\mathbf{D}_{i,j}(\mathbf{Z}_1)\right]^{(m,m)}$ | $\sum_{n_1=1}^{N}\sum_{n_2=1}^{N}\overline{\mathbf{J}}_i\left(\mathbf{Z}_1^{(n_1,n_2)}\otimes Diag\left\{\mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)},\mathbf{0}_{L+1}\cdots\mathbf{0}_{L+1}\right\}\right)\overline{\mathbf{J}}_j^T$ | | |
| $\left[\mathbf{K}_{i,j}(\mathbf{Z}_2)\right]_{p_1,p_2}^{(n_1,n_2)}$ | $\sum_{m=1}^{M}vec\left(\left[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\right]_1^{(p_1,p_2)}\right)^T vec\left(\mathbf{R}_{\Delta\mathbf{h},m}^{(n_1,n_2)}\right)^*$ | | ($vec(.)$ stacks a matrix into a vector columnwise and $\left[\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\right]_1^{(p_1,p_2)}$ denotes the first $(L+1)\times(L+1)$ diagonal block submatrix of the $(p_1,p_2)$th block submatrix of $\overline{\mathbf{J}}_i^T\mathbf{Z}_2^{(m,m)}\overline{\mathbf{J}}_j\in\mathbb{C}^{PQ\times PQ}$) |

receivers, namely, the layered space-frequency equalization (LSFE) scheme [1] and the group-wise V-BLAST detector [30]. From the table we observe that the computation complexity of the three methods is all about $O(Q^3)$, while the adoption of the RR implementation can reduce the complexity of the GSC scheme to $O(Q^2)$. Specifically, for the system parameters $(N,M,Q,L,G,L_b,N_b,I)=(2,2,64,7,8,1,2,10)$ adopted in the simulation section the flop counts are, respectively, $FC_{GSC}\approx 1.46+\times 10^7$, $FC_{GSC}^{(RR)}\approx 3.2\times 10^6$, $FC_{LSFE}\approx 2\times 10^6$, and $FC_{GB}\approx 1.2\times 10^7$ ($L_b$ and $N_b$ respectively denote the tap order of feedback filter and the number of decision stages for the LSFE detector, and $I$ the number of iterations involved in the RR implementation).

## VI. SIMULATION RESULTS

This section illustrates the simulated performance of the proposed scheme. We consider a MIMO-SC system with carrier frequency 5 GHz, transmission bandwidth 20 MHz, $N=2$ transmit antennas, $M=2$ receive antennas, symbol block size $Q=64$, and CP length $G=8$. The velocity of the moving transmitter is set to be 120 Km/h. The source symbols are drawn from the QPSK constellation. The channels are characterized by the Jakes's model [31] with order $L=7$ and the impulse response is normalized such that $\sum_{n=1}^{2}\sum_{l=0}^{7}E\left\{\left|h^{(m,n)}(k,l)\right|^2\right\}=1$ for each $1\le m\le 2$ and $k\ge 0$. The input SNR at the $m$th receive antenna is defined as $\mathrm{SNR}:=\sigma_v^{-2}$, and the data burst length is set to be $T=15$. The number of iterations involved in the RR implementation is set to be $I=10$. After inter-group interference suppression the MMSE V-BLAST detector [32] is used for symbol recovery.

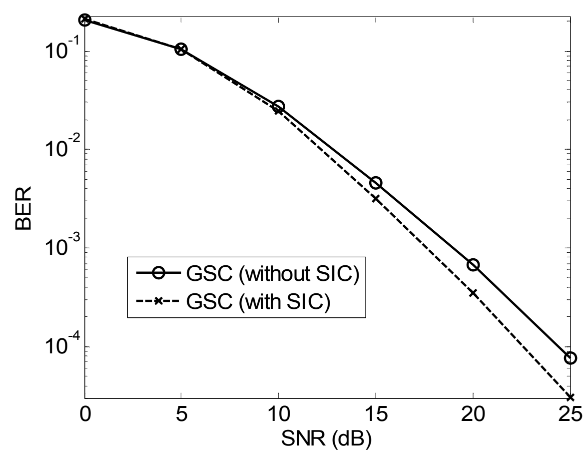


Fig. 3. BER performances of the proposed method with and without the SIC mechanism (perfect channel knowledge).

### A. Impacts of SIC Based Implementation

We first compare the performances of the proposed GSC equalizers with and without the SIC implementation. Figures 3 and 4, respectively, show the simulated bit-error-rate (BER) with perfect and imperfect channel knowledge (for the latter the channels are estimated via the LS training technique [5] with training power $P_t=32$). We note that, when channel is perfectly known, the robust GSC solution (45) reduces to (23). The results show that the proposed method combined with the SIC mechanism yields about a 2~3 dB gain with exact channel knowledge, and a 4~5 dB gain with LS channel estimate. This would benefit from the increased receive diversity attained by the SIC mechanism.

TABLE II
ALGORITHM SUMMARY OF THE SIC MECHANISM.

**Initialization:**

$\mathbf{r}^{(0)}(t) = \mathbf{r}(t);\ \mathbf{X}_1^{(0)}(t) = \widehat{\mathbf{X}}_1(t);\ \mathbf{X}_3^{(0)}(t) = \widehat{\mathbf{X}}_3(t);$

$\mathbf{Y}_2^{(0)} = \sum_{i=1}^{L+1} \mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\mathbf{K}_{i,j}(\mathbf{D}_c^{(i)}(t)^H);$

$\mathbf{R}_I^{(0)} = \sum_{i=1}^{L+1} \mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ};$

**Recursion:** for $1 \le j \le G+1$

1. if $j = 1$

   $\mathbf{X}_1^{(1)}(t) = \mathbf{X}_1^{(0)}(t);\ \mathbf{X}_3^{(1)}(t) = \mathbf{X}_3^{(0)}(t)$

   else

   $\mathbf{X}_1^{(j)}(t) = \mathbf{X}_1^{(j-1)}(t) - \mathbf{D}_c^{(j-1)}(t)\mathbf{D}_H^{(j-1)}\mathbf{D}_H^{(j-1)}(t)^H$

   $\mathbf{X}_3^{(j)}(t) = \mathbf{X}_3^{(j-1)}(t) - \mathbf{D}_c^{(j-1)}(t)\mathbf{D}_{j-1,j}^{(j-1)}\left(\mathbf{X}^{(j-1,j)}(t)\right)$

   end

2. $\mathbf{R}_I^{(j)}(t) = \mathbf{R}_I^{(j-1)}(t) - \mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)^H;$
3. $\mathbf{R}_L^{(j)}(t) = \mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\widehat{\mathbf{H}}_{av}^{(j)H}\mathbf{D}_c^{(j)}(t)^H;$
4. $\mathbf{R}_{e,1}^{(j)}(t) = \mathbf{X}_1^{(j)}(t) + \mathbf{D}_{j,j}(\mathbf{X}^{(j)}(t)) + \mathbf{X}_3^{(j)}(t) + \mathbf{X}_3^{(j)}(t)^H;$
5. $\mathbf{R}_{e,2}^{(j)}(t) = \left[\mathbf{X}_1^{(j-1)}(t) - \mathbf{D}_c^{(j)}(t)\mathbf{D}_H^{(j)}\mathbf{D}_c^{(j)}(t)^H\right] + \left[\mathbf{Y}_2^{(j-1)}(t) - \mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\mathbf{K}_{j,j}(\mathbf{D}_c^{(j)}(t)^H)\right] + \left[\mathbf{X}_3^{(j-1)}(t)^H - \mathbf{D}_c^{(j)}(t)\mathbf{D}_{j,j}(\mathbf{X}^{(j,j)}(t)^H)\right]^H;$
6. Compute $\mathbf{W}_r^{(j)}$ using (4.21)
7. $\hat{\mathbf{s}}^{(j)}(t) = dec\left(\mathbf{W}_r^{(j)H}\mathbf{r}^{(j)}(t)\right)$
8. $\mathbf{r}^{(j+1)}(t) = \mathbf{r}^{(j)}(t) - \mathbf{D}_c^{(j)}(t)\widehat{\mathbf{H}}_{av}^{(j)}\hat{\mathbf{s}}^{(j)}(t)$

end

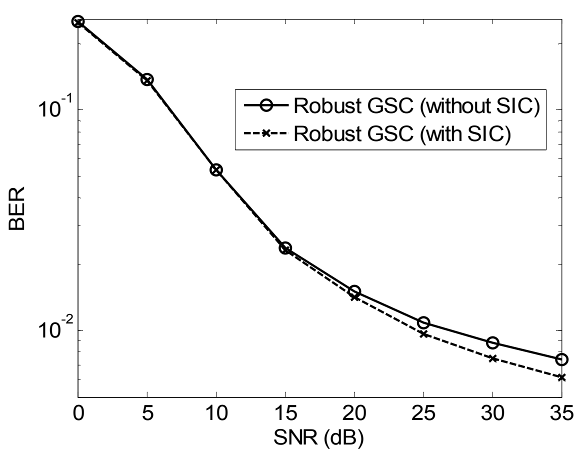


Fig. 4. BER performances of the proposed method with and without the SIC mechanism (LS channel estimate).

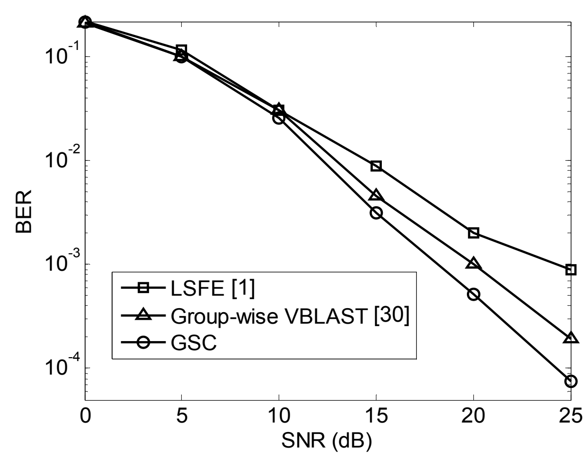


Fig. 5. BER performances of the three methods (perfect channel knowledge).

## *B. Comparison with Existing Works*

We compare the BER performance of the proposed method (combined with SIC) with the two alternative solutions group-wise V-BLAST [30] and LSFE [1]. The group size of the group-wise V-BLAST is set to be 16 (this is the same as the group size in our scheme $NP = 2 \times 8 = 16$). For the LSFE detector, the number of decision stages and the tap order of the feedback filter are respectively $N_b = 2$ and $L_b = 1$ ($N_b = 2$ is the suggested optimal choice [1] for two transmit antennas, and through simulation further increasing $L_b$ does not seem to improve performance). Figures 5 and 6, respectively, show the results with perfect and imperfect channel knowledges. As we can see, even in the ideal case the proposed approach can outperform the two comparative choices. When only a channel estimate is available, the performances of all equalizers degrade, but the robust solution (45) does yield the lowest BER. With a fixed SNR level (25 dB) Figure 7 illustrates the BER of all methods when the burst duration $T$ increases from 13 to 25. The results show that the proposed robust equalizer

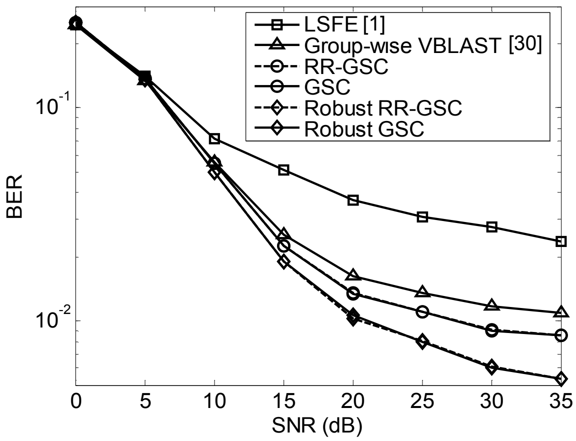


Fig. 6. BER performances of the three methods (LS channel estimate).

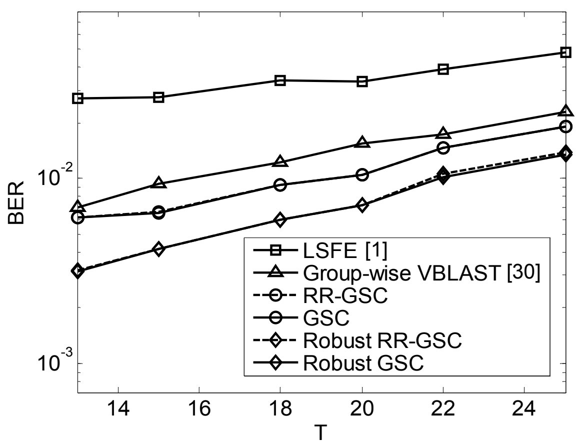


Fig. 7. BER performances of the three method w.r.t. various burst length (LS channel estimate).

(45) can significantly limit the BER penalty and is thus quite resistant to the increase of $T$. Due to space limitation, more performance comparisons regarding different mobile velocities and antenna configurations are relegated to a supplementary file available at [33].

### C. Channel Estimation with Linear Interpolation

Figure 8 further depicts the simulated BER when the channel information at each time instant is acquired through the linear interpolation technique [21]; in this case the proposed robust scheme is given by (46). Compared with Figure 6, for each receiver the incurred performance loss w.r.t. the exact channel knowledge case is less severe due to the availability of more timely channel information. The proposed robust solution (46), as expected, achieves the best performance since it is capable of combating channel estimation errors.

## VII. Conclusions

We study the robust receiver design problem for MIMO-SC systems when the multipath channels are time-varying and are estimated through the LS training technique. By exploiting certain group-wise orthogonality structure imbedded in the time-domain channel matrix, the proposed receiver aims for inter-group interference suppression followed by a low-complexity intra-group symbol recovery scheme.

The design of the interference rejection matrix is mathematically formulated via constrained optimization technique. To further tackle the adverse effects due to imperfect channel knowledge we leverage the GSC principle to reformulate the problem into an equivalent unconstrained setup. The constraint-free GSC formulation resorts to proper weighting matrix decomposition and has several unique advantages: 1) it provides a simple yet efficient channel tracking mechanism by setting the adaptive component to be time-varying, 2) it allows for a very natural cost function for weighting matrix design against channel uncertainty due to time variation and estimation errors, 3) it enables us to directly model the channel mismatch effect into the filtered signal model via the perturbation analysis and, accordingly, we can then exploit the channel error statistics to derive a closed-form solution. The proposed approach can be further implemented in an SIC fashion for performance enhancement; it can also be used for estimation-error resistant receiver design when the channel estimate at each time instant is acquired via linear interpolation. Compared with the existing works the proposed scheme yields significantly improved simulated BER, with comparable algorithm complexity.

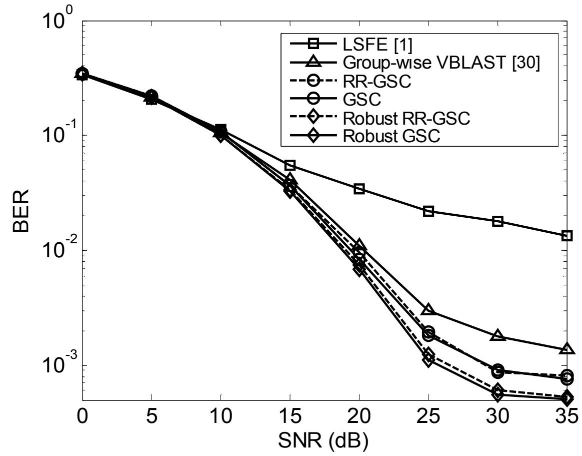


Fig. 8. BER performances of the three methods (interpolation-based channel estimate).

## Appendix A
## Derivations of (35) and (36)

With (30), (31) and by averaging over source symbols and channel noises, we have

$$
\begin{aligned}
&E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = \\
&E_e\left\{ \ \widehat{\mathbf{B}}_{av}^{(j)H}\left(\textstyle\sum_{i=1}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H\right.\right. \\
&\qquad\left.\left. +\sigma_v^2 I_{MQ} + E_h\{\mathbf{R}_{\delta\mathbf{H}}(t)\}\right)\widehat{\mathbf{B}}_{av}^{(j)} \ \right\},
\end{aligned} \tag{47}
$$

and

$$E\left\{\bar{\mathbf{z}}_b^{(j)}(t)\bar{\mathbf{i}}^{(j)}(t)^H\right\} = E_e\left\{ \widehat{\mathbf{B}}_{av}^{(j)H}\left(\sum_{i=1,i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sigma_v^2\mathbf{I}_{MQ} + E_h\left\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\right\}\right)\widehat{\mathbf{H}}_{av}^{(j)} \right\}, \quad (48)$$

where $E_e\{\cdot\}$ and $E_h\{\cdot\}$ denote the expectations taken with respect to the channel estimation error and channel temporal variation, respectively,

$$\mathbf{R}_{\delta\mathbf{H}}(t) = \sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sum_{i=1}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\delta\mathbf{H}^{(i)}(t)^H + \sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H, \quad (49)$$

and

$$\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t) = \sum_{i=1,i\neq j}^{G}\delta\mathbf{H}^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)^H + \sum_{i=1,i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\widehat{\mathbf{H}}_{av}^{(i)}\delta\mathbf{H}^{(i)}(t)^H + \sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H, \quad (50)$$

in which

$$\mathbf{D}_c^{(i)}(t) \in \mathbb{C}^{MQ\times MQ}$$
$$\mathbf{D}_c^{(i)}(t) := diag\left\{\mathbf{J}^{i-1}[\mathbf{a}_1(t)\cdots\mathbf{a}_P(t)]^T\right\}\otimes\mathbf{I}_M, \quad (51)$$

with

$$\mathbf{a}_p(t) = \Big[\ \rho\left(\bar{t} + ((p-1)(L+1)+i-1)_Q, 0\right)\frac{\sigma_0}{\widehat{\sigma}_0}, ..., \rho\left(\bar{t} + ((p-1)(L+1)+L+i-1)_Q, L\right)\frac{\sigma_L}{\widehat{\sigma}_L}\ \Big],$$
$$\rho(t,l) = \mathbf{R}_h(t,l)(\sigma_l\widehat{\sigma}_l)^{-1},$$
$$\widehat{\sigma}_l^2 = \sigma_l^2 + [\mathbf{R}_{\Delta h,m}]_{(n-1)(L+1)+l+1,(n-1)(L+1)+l+1},$$

where $\mathbf{a}_p(t) \in \mathbb{R}^{1\times(L+1)}$ and $\delta\mathbf{H}^{(i)}(t) \in \mathbb{C}^{MQ\times NP}$ has the $(m,n)$th $Q\times P$ block submatrix whose *p*th column $\delta\mathbf{c}_{m,n}^{(i,p)}(t) \in \mathbb{C}^Q$ is given by (52) (shown in the bottom of this page), where, for a fixed *t*, $\delta h_{m,n}(t,l)$, $0 \le l \le L$, are zero-mean Gaussian random variables with variance $\sigma_l^2\left(1 - |\rho(t,l)|^2\right)$. We claim that

$$E_h\{\mathbf{R}_{\delta\mathbf{H}}(t)\} = E_h\left\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\right\} = N\mathbf{I}_M\otimes\mathbf{D}_Q(t), \quad (53)$$

the result then follows from (47), (48), and (53).

*Proof of (53):* Since the nonzero entries of $\delta H^{(i)}(t)$ are zero-mean Gaussian random variables, we have $E_h\{\delta\mathbf{H}^{(i)}(t)\} = \mathbf{0}_{MQ\times NP}$, for $1\le i\le G$, and from (49) and (50) we have

$$E_h\{\mathbf{R}_{\delta\mathbf{H}}(t)\} = E_h\left\{\mathbf{R}_{\delta\mathbf{H}}^{(j)}(t)\right\} = E_h\left\{\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right\}. \quad (54)$$

Let $Diag\{\mathbf{C}_1, ..., \mathbf{C}_m\} \in \mathbb{C}^{mn\times mn}$ be the block diagonal matrix with the diagonal block submatrices $\mathbf{C}_p \in \mathbb{C}^{n\times n}$, $1\le p\le m$. By the definition of $\delta\mathbf{H}^{(i)}(t) \in \mathbb{C}^{MQ\times NP}$ in (52), the $(m_1,m_2)$th $Q\times Q$ block submatrix of $\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H \in \mathbb{C}^{MQ\times MQ}$, $1\le m_1,m_2\le M$, is given by (55) (shown in the bottom of next page). Since the elements of $\delta\mathbf{h}_{m,n}^{(i,p)}(t)$, $\forall m,n$, are independent (see (52)), we have

$$E_h\left\{\delta\mathbf{h}_{m_1,n}^{(i,p)}(t)\delta\mathbf{h}_{m_2,n}^{(i,p)}(t)^H\right\} = \begin{cases}\mathbf{D}_p^{(i)}(t), & m_1 = m_2\\ \mathbf{0}_{L+1}, & m_1\neq m_2\end{cases}, \quad (56)$$

$\mathbf{D}_p^{(i)}(t) \in \mathbb{C}^{(L+1)\times(L+1)}$ is a diagonal matrix with the *l*th diagonal element $\left[\mathbf{D}_p^{(i)}(t)\right]_{l,l} = \sigma_{l-1}^2(1 - \left|\rho(\bar{t} + (i+(p-1)(G+1)+l-2)_Q, l-1)\right|^2)$ for $1\le l\le L+1$. Equation (56) directly implies that $E_h\left\{\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right\}$ is a diagonal matrix with the *m*th $Q\times Q$ diagonal block given by (57) (shown in the bottom of the next page). By some direct rearrangements and since *Q=PG*, (57) becomes

$$E_h\left\{\left[\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m,m)}\right\} = N\mathbf{D}_Q(t), \quad (58)$$

where

$$\mathbf{D}_Q(t) := diag\left\{\ \sum_{l=0}^{L}\sigma_l^2\left(1 - \left|\rho(\bar{t},l))\right|^2\right), ..., \sum_{l=0}^{L}\sigma_l^2\left(1 - \left|\rho(\bar{t}+Q-1,l))\right|^2\right)\ \right\} \in \mathbb{C}^{MQ\times MQ}.$$

Equation (53) directly follows from (58). □

## Appendix B
## Derivations of (40) and (41)

To derive (40) and (41), we have to determine the expectation quantities associated with the channel estimation error in (35) and (36). Based on (2) the $(m,n)$th block submatrices of $\mathbf{H}_{av}^{(i)} \in \mathbb{C}^{MQ\times NP}$ and $\Delta\mathbf{H}_{av}^{(i)} \in \mathbb{C}^{MQ\times NP}$ are respectively defined as

$$\left[\mathbf{H}_{av}^{(i)}\right]^{(m,n)} = \bar{\mathbf{J}}_i\left(\left[\mathbf{h}_{m,n}^{(av)T}, 0, ..., 0\right]^T\otimes\mathbf{I}_P\right) \in \mathbb{C}^{Q\times P}, \quad (59)$$

$$\delta\mathbf{c}_{m,n}^{(i,p)}(t) = \mathbf{J}^{i+(p-1)G-1}[\underbrace{\delta h_{m,n}(\bar{t} + ((p-1)(L+1)+i-1)_Q, 0), ..., \delta h_{m,n}(\bar{t} + ((p-1)(L+1)+i+L-1)_Q, L)}_{:=\delta\mathbf{h}_{m,n}^{(i,p)T}}, 0, ..., 0]^T. \quad (52)$$

and

$$\left[\Delta\mathbf{H}_{av}^{(i)}\right]^{(m,n)} = \overline{\mathbf{J}}_i \left( \left[\Delta\mathbf{h}_{m,n}^T, 0, ..., 0\right]^T \otimes \mathbf{I}_P \right) \in \mathbb{C}^{Q\times P}$$
$$, 1 \leq n \leq N,\ 1 \leq m \leq M, \quad (60)$$

where $\Delta\mathbf{h}_{m,n} \in \mathbb{C}^{L+1}$ is the $n$th subvector of $\Delta\mathbf{h}_m$ (see (5)) and $\overline{\mathbf{J}}_i \in \mathbb{R}^{Q\times PQ}$ denotes a matrix with the $p$th $Q\times Q$ block submatrix given by $\left[\overline{\mathbf{J}}_i\right]^{(1,p)} = \mathbf{J}^{i+(p-1)G-1}$ ($\mathbf{J}$ is defined in (3)). Assuming that the channel is slowly varying and SNR is high, the channel estimation error is thus small and $\Delta\mathbf{H}_{av}^{(i)}$ and $\Delta\mathbf{B}_{av}^{(j)}$ (see (39)) are close to zero matrices. Substituting (39) into (35) and using the fact that $E_e\{\Delta\mathbf{B}_{av}^{(j)H}\} = 0_{MQ\times(MQ-NP)}$ (cf. (39)) and the circularity condition of $\Delta\mathbf{h}_{m,n}^{(av)}$, by keeping only the first- and second-order terms of estimation error (35) can be expanded as

$$\begin{aligned}
&E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{z}}_b^{(j)}(t)^H\right\} = \\
&\mathbf{B}_{av}^{(j)H}\left(\mathbf{R}_L^{(j)}(t) + \overline{\mathbf{R}}_I^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{B}_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H} E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(j)}(t)\Delta\mathbf{H}_{av}^{(j)H}\right\}\mathbf{B}_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{G}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)^H\right)\mathbf{B}_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{G}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\mathbf{X}^{(i,j)}(t)\Delta\mathbf{H}_{av}^{(j)H}\right\}\right)\mathbf{B}_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1}^{G}E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(i,j)}(t)^H\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)^H\right)\mathbf{B}_{av}^{(j)},
\end{aligned} \quad (61)$$

where $\mathbf{R}_L^{(j)}(t)$ and $\overline{\mathbf{R}}_I^{(j)}(t)$ are defined in (42) and (43), respectively, and the matrices $\mathbf{X}^{(j)}(t)$ and $\mathbf{X}^{(i,j)}(t)$ are defined in Table I. Similarly, (36) also can be expanded as

$$\begin{aligned}
&E\left\{\overline{\mathbf{z}}_b^{(j)}(t)\overline{\mathbf{i}}^{(j)}(t)^H\right\} = \\
&\mathbf{B}_{av}^{(j)H}\left(\overline{\mathbf{R}}_I^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\mathbf{H}_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1, i\neq j}^{G}\mathbf{D}_c^{(i)}(t)E_e\left\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\right\}\mathbf{D}_c^{(i)}(t)\right)H_{av}^{(j)} + \\
&\mathbf{B}_{av}^{(j)H}\left(\sum_{i=1, i\neq j}^{G}\mathbf{D}_c^{(i)}(t)\mathbf{H}_{av}^{(i)}E_e\left\{\Delta\mathbf{H}_{av}^{(i)H}\mathbf{D}_c^{(i)}(t)H\Delta\mathbf{H}_{av}^{(j)}\right\}\right) \\
&+ \mathbf{B}_{av}^{(j)H}\Big(\textstyle\sum_{i=1, i\neq j}^{G}E_e\left\{\Delta\mathbf{H}_{av}^{(j)}\mathbf{X}^{(i,j)}(t)^H\Delta\mathbf{H}_{av}^{(i)H}\right\} \\
&\qquad \cdot\mathbf{D}_c^{(i)}(t)^H\Big)\mathbf{H}_{av}^{(j)}.
\end{aligned} \quad (62)$$

We observe that, in (61) and (62), each of the involved expectations admit one of the following forms: $E_e\{\Delta\mathbf{H}_{av}^{(i)}\Delta\mathbf{H}_{av}^{(i)H}\}$, $E_e\{\Delta\mathbf{H}_{av}^{(i)}\mathbf{Z}_1\Delta\mathbf{H}_{av}^{(j)H}\}$ and $E_e\{\Delta\mathbf{H}_{av}^{(i)H}\mathbf{Z}_2\Delta\mathbf{H}_{av}^{(j)}\}$, for some given matrices $\mathbf{Z}_1 \in \mathbb{C}^{NP\times NP}$ and $\mathbf{Z}_2 \in \mathbb{C}^{MQ\times MQ}$. Equation (40) and (41) can be obtained by exploiting the circularly Gaussian channel estimation error characteristic followed by some straightforward manipulations. The detailed derivation is referred to [34].

## APPENDIX C
## DERIVATION OF (46)

Assuming that the channel temporal variation is piecewise linear in time, the channel estimate at each time instant $k$ (within one burst) can be acquired via the following relation [21]

$$\widehat{h}_{m.n}(k,l) = \widehat{h}_{m,n}^{(av)}(l) + \widehat{\alpha}_l(k - (G+Q/2))$$
$$, G+Q \leq k \leq T(G+Q) - 1, \quad (63)$$

where $\widehat{h}_{m,n}^{(av)}(l)$ denotes the LS estimate of the $l$th channel tap, $\widehat{\alpha}_l = (\widehat{h}_{m,n}^{(av,next)}(l) - \widehat{h}_{m,n}^{(av)}(l))/T(G+Q)$ is the estimated variation slope with $\widehat{h}_{m,n}^{(av,next)}(l)$ being the channel estimate obtained in the next burst. Substituting $\widehat{h}_{m,n}^{(av)}(l) = h_{m,n}^{(av)}(l) + \Delta h_{m,n}(l)$ and $\widehat{h}_{m,n}^{(av,next)}(l) = h_{m,n}^{(av,next)}(l) + \Delta h_{m,n}^{(next)}(l)$, where $\Delta h_{m,n}(l)$ and $\Delta h_{m,n}^{(next)}(l)$ represent the estimation errors, into (63), we have

$$\widehat{h}_{m.n}(k,l) = h_{m.n}(k,l) + \Delta h_{m.n}(k,l)$$
$$, G+Q \leq k \leq T(G+Q) - 1, \quad (64)$$

$$\begin{aligned}
\left[\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m_1,m_2)} &= \sum_{i=1}^{G}\sum_{n=1}^{N}\left[\delta\mathbf{H}^{(i)}(t)\right]^{(m_1,n)}\left[\delta\mathbf{H}^{(i)}(t)\right]^{(m_2,n)H} \\
&= \sum_{i=1}^{G}\sum_{n=1}^{N}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)G-1}Diag\left(\delta\mathbf{h}_{m_1,n}^{(i,p)}(t)\delta\mathbf{h}_{m_2,n}^{(i,p)}(t)^H, \mathbf{0}_{L+1}, ..., \mathbf{0}_{L+1}\right)\left(\mathbf{J}^{i+(p-1)G-1}\right)^T
\end{aligned} \quad (55)$$

$$E_h\left\{\left[\sum_{i=1}^{G}\delta\mathbf{H}^{(i)}(t)\delta\mathbf{H}^{(i)}(t)^H\right]^{(m,m)}\right\} = N\sum_{i=1}^{G}\sum_{p=1}^{P}\mathbf{J}^{i+(p-1)G-1}Diag\{\mathbf{D}_p^{(i)}(t), \mathbf{0}_{L+1}, ..., \mathbf{0}_{L+1}\}\left(\mathbf{J}^{i+(p-1)G-1}\right)^T. \quad (57)$$

TABLE III
ALGORITHM SUMMARY OF THE RR IMPLEMENTATION.

**Initialization:**

$\mathbf{U}_0^{(j)} = \mathbf{0}_{(MQ-NP)\times NP}$

$\mathbf{Y}_0 = \widehat{\mathbf{B}}_{av}^{(j)H}\left(\widehat{\overline{\mathbf{R}}}_I^{(j)}(t) + \widehat{\mathbf{R}}_{e,2}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\widehat{\mathbf{H}}_{av}^{(j)}$

**Recursion:** for $1 \le k \le I$

1. if $k = 1$

   $\mathbf{P}_1 = \mathbf{Y}_0$

   else

   $\mathbf{L}_{1,k} = Diag\left(\mathbf{Y}_{k-2}^H\mathbf{Y}_{k-2}\right)^{-1} Diag\left(\mathbf{Y}_{k-1}^H\mathbf{Y}_{k-1}\right)$

   $\mathbf{P}_k = \mathbf{Y}_{k-1} + \mathbf{P}_{k-1}\mathbf{L}_{1,k}$

   end

2. $\mathbf{L}_{2,k} = Diag\left(\mathbf{P}_k^H\widehat{\mathbf{B}}_{av}^{(j)H}\left(\widehat{\mathbf{R}}_L^{(j)}(t) + \widehat{\overline{\mathbf{R}}}_I^{(j)}(t) + \widehat{\mathbf{R}}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\widehat{\mathbf{B}}_{av}^{(j)}\mathbf{P}_k\right)^{-1} Diag\left(\mathbf{Y}_{k-1}^H\mathbf{Y}_{k-1}\right)$

3. $\mathbf{Y}_k = \mathbf{Y}_{k-1} - \widehat{\mathbf{B}}_{av}^{(j)H}\left(\widehat{\mathbf{R}}_L^{(j)}(t) + \widehat{\overline{\mathbf{R}}}_I^{(j)}(t) + \widehat{\mathbf{R}}_{e,1}^{(j)}(t) + N\mathbf{I}_M \otimes \mathbf{D}_Q(t)\right)\widehat{\mathbf{B}}_{av}^{(j)}\mathbf{P}_k\mathbf{L}_{2,k}$

4. $\mathbf{U}_k^{(j)} = \mathbf{U}_{k-1}^{(j)} + \mathbf{P}_k\mathbf{L}_{2,k}$

end

$\mathbf{U}^{(j)} = \mathbf{U}_I^{(j)}$

TABLE IV
FLOP COUNT COMPARISON.

| | |
|---|---|
| $FC_{GSC}$ | $\frac{2}{3}GM^3Q^3 + \left(M^2G(MG-N) + M^2G(G-1)(2M+N) - 2(G-1)M^2NP\right)Q^2$ $+\left(M^2N(L+1)\left(\frac{G(G-1)}{2} - 3GP\right) - (G-1)MN(MG-N)P + 2(G-1)MN^2P^2\right)Q$ $+\left(4(G-1)MN(L+1)(M(L+1)-2N) + (G-1)N^2(M(L+1)+2P)P - \frac{2}{3}(G-1)N^3P^2\right)P$ |
| $FC_{GSC}^{(RR)}$ | $\left(M^3G(L+1) + (G-1)M^2N\right)Q^2 + \left(\frac{G(G-1)}{2}M^2N(L+1) - (G+1)M^2N(L+1)P - (G-1)MN^2P\right)Q$ $+\left(4(G-1)MN(L+1)(M(L+1)-2N) + MNP(L+1) + IN(MQ-NP)^2\right)P$ |
| $FC_{LSFE}$ | $\left(\frac{2}{3}M^3 + N\right)Q^3 + \left(\frac{1}{2}(M^4 + M^3 + M^2) + N_bL_b\right)Q^2 + \left(\frac{1}{2}M^3 + M^2 + \frac{5}{6}M\right)Q$ |
| $FC_{GB}$ | $\frac{5}{3}(G-1)M^3Q^3 + \left((G-1)M^2(M(L+1)+2) - \frac{5}{2}(G-1)(G-2)M^2NP\right)Q^2$ $+\left(\frac{5}{6}(G-1)(G-2)(2G-3)MN^2P^2 - (G-1)(G-2)MNP\right)Q$ $+\left(\frac{1}{6}(G-1)(G-2)(2G-3)N^2P^2 - \frac{1}{4}(G-1)^2(G-2)^2N^3P^3\right)$ |

in which $\Delta h_{m.n}(k,l) = \Delta h_{m,n}(l) + \Delta\alpha_l(k - (G + Q/2))$ is the composite channel estimation error with $\Delta\alpha_l = \left(\Delta h_{m,n}^{(next)}(l) - \Delta h_{m,n}(l)\right)/T(G+Q)$. Based on (63) and (11), the corresponding estimated signature matrix is then expressed as

$$\widehat{\mathbf{H}}^{(i)}(t) = \mathbf{H}^{(i)}(t) + \Delta\mathbf{H}^{(i)}(t),\ 1 \le i \le G, \tag{65}$$

where

$$\Delta\mathbf{H}^{(i)}(t) = \Delta\mathbf{H}_{av}^{(i)} + \mathbf{D}_T(t)\left(\Delta\mathbf{H}_{av,next}^{(i)} - \Delta\mathbf{H}_{av}^{(i)}\right),$$
$$\mathbf{D}_T(t) = (T(G+Q))^{-1}\mathbf{I}_M \otimes diag\{\overline{t}, ..., \overline{t} + Q - 1\}$$

and $\Delta\mathbf{H}_{av,next}^{(i)}$ can be formed by simply replacing $\Delta\mathbf{h}_{m,n}$ with $\Delta\mathbf{h}_{m,n}^{(next)} = \left[\Delta h_{m,n}^{(next)}(l), ..., \Delta h_{m,n}^{(next)}(L)\right]^T$ in the definition of $\Delta\mathbf{H}_{av}^{(i)}$ (see (60)). By following the same procedure in [19] and Appendix B, the optimal robust GSC weight for the *j*th group is given by $\mathbf{W}_a^{(j)}(t) = \mathbf{H}^{(j)}(t) - \mathbf{B}^{(j)}(t)\mathbf{U}_a^{(j)}(t)$ with

$$\mathbf{U}_a^{(j)}(t) = \left(\mathbf{B}^{(j)}(t)^H\left(\mathbf{R}_I^{(j)}(t) + \mathbf{R}_e(t)\right)\mathbf{B}^{(j)}(t)\right)^{-1} \cdot \mathbf{B}^{(j)}(t)^H\mathbf{R}_I^{(j)}(t)\mathbf{H}^{(j)}(t), \tag{66}$$

where $\mathbf{R}_e(t) \in \mathbb{C}^{MQ\times MQ}$ is defined in Table I. Based on (66), the sampled version of the overall GSC weight is shown in (46).

## REFERENCES


[1] X. Zhu and R. D. Murch, "Layered space-frequency equalization in a single-carrier MIMO system for frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 701-708, May 2004.

[2] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G long-term evolution," *IEEE Commun. Mag.*, vol. 44, no. 3, pp. 38-45, Mar. 2006.

[3] H. G. Myung, L. J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 30-38, Sept. 2006.

[4] J. Coon, S. Armour, M. Beach, and J. McGeehan, "Adaptive frequency-domain equalization for single-carrier multiple-input multiple-output wireless transmissions," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3247-3256, Aug. 2005.

[5] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO-OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1615-1624, June 2003.

[6] J. Coon, M. Beach, and J. McGeehan, "Optimal training sequences for channel estimation in cyclic-prefix-based single-carrier systems with transmit diversity," *IEEE Signal Processing Lett.*, vol. 11, no. 9, pp. 729-732, Sept. 2004.

[7] J. B. Schodorf and D. B. Williams, "A constrained optimization approach to multiuser detection," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 258-262, Jan. 1997.

[8] Z. Tian, K. L. Bell, and H. L. Van Trees, "Robust constrained linear receivers for CDMA wireless systems," *IEEE Trans. Signal Processing*, vol. 49, no. 7, pp. 1510–1522, July 2001.

[9] L. J. Griffiths, and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, no. 1, pp. 27-34, Jan. 1982.

[10] J. B. Schodorf and D. B. Williams, "A constrained optimization approach to multiuser detection," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 258-262, Jan. 1997.

[11] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, pp. 4-24, Apr. 1988.

[12] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170-1184, Oct. 1993.

[13] G. W. Stewart, "Error and perturbation bounds for subspaces associated with certain eigenvalue problems," *SIAM Review*, 15-33, pp. 1413-1416, 1973.

[14] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[15] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1684-1696, May 2005.

[16] S. Shahbazpanahi and A. B. Gershman, Z. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2257–2269, Nov. 2004.

[17] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "A robust linear receiver for multi-access space-time block-coded MIMO systems based on probability-constrained optimization," in *Proc. IEEE VTC*, vol. 1, pp. 118-122, May 2004.

[18] Y. Rong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block coded MIMO systems: a probabilistically constrained approach," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1560-1570, Aug. 2006.

[19] C. Y. Lin, J. Y. Wu, and T. S. Lee, "Robust constrained-optimization based linear receiver for high-rate MIMO OFDM against channel estimation errors," *IEEE Trans. Signal Processing,* vol. 55, no. 6, pp. 2628-2645, June 2007.

[20] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1995.

[21] Y. Mostofi and D. C. Donald, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765-774, Mar. 2005.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[23] B. R. Breed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Lett.*, vol. 9, no. 6, pp. 168-169, June 2002.

[24] N. Khaled, G. Leus, C. Desset, and H. De Man, "A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels," in *Proc. IEEE ICCASP*, vol. 4, pp. iv.485-iv.488, May 2004.

[25] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Select. Areas Commun.*, vol. 16, no. 18, pp. 1423-1436, Oct. 1998.

[26] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 2002.

[27] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2820-2830, Nov. 2002.

[28] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Cambridge University Press, third-edition, 1996.

[29] C. L. Ho, G. J. Lin, and T. S. Lee, "An OSIC based reduced-rank MIMO equalizer using conjugate gradient algorithm," *IEICE Trans. Commun.*, vol. E86-B, no. 9, pp. 2656-2664, Sept. 2003.

[30] X. N. Tran and T. Fujino, "Groupwise successive ICI cancellation for OFDM systems in time-varying channels," in *Proc. IEEE International Symposium on Signal Processing and Information Theory*, pp. 489-494, Dec. 2005.

[31] W. C. Jakes, *Microwave PlaceMobile Communications*. IEEE Press, 1974.

[32] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over rich scattering wireless channels," in *Proc. IEEE ISSSE-98*, Italy, pp. 295-300, Sept. 1998.

[33] http://140.113.144.102/documents/supplementary_results.pdf.

[34] http://mimo.cm.nctu.edu.tw/Professor/Papers/Derivation.pdf.


**Chih-Yuan Lin** was born in Yilan, Taiwan, R.O.C., in 1979. He received the B.S. degree from National Chung Hsing University in 2000, in electrical engineering, and the Ph.D. degree from National Chiao Tung University in 2007, in communication engineering. His current research interests include space-time signal processing for wireless communications and statistical signal processing.

**Jwo-Yuh Wu** (M'04) received the B.S. degree in 1996, the M.S. degree in 1998, and the Ph.D. degree in 2002 from the National Chiao Tung University, Taiwan, all in Electrical and Control Engineering. During 2003 and 2007 he was a post doctor research fellow in the Department of Communications Engineering in National Chiao Tung University, where he is an assistant professor starting from Feb. 2008. His research interests are in signal processing, wireless communications and networking, and control systems.

**Ta-Sung Lee** (M'94-SM'05) received the B.S. degree from National Taiwan University in 1983, the M.S. degree from University of Wisconsin, Madison, in 1987, and the Ph.D. degree form Purdue University, W. Lafayette, IN, in 1989, all in electrical engineering. In 1990, he joined the Faculty of National Chiao Tung University (NCTU), Hsinchu, Taiwan, where he holds a position as Professor of Department of Communication Engineering. From 2007 to 2008, he was Dean of Student Affairs, and from 2005 to 2007, he was Chairman of Department of Communication Engineering of NCTU. His other past positions include Technical Advisor at Information and Communications Research Labs. of Industrial Technology Research Institute (ITRI), Managing Director of MINDS Research Center, College of ECE, NCTU, and Managing Director of Communications & Computer Training Program, NCTU. In August 2008, he was appointed commissioner of National Communications Commission (NCC) of Taiwan. He is active in research and development in advanced techniques for wireless communications, such as smart antenna and MIMO technologies, sensor networks, cross-layer design, and hardware/software prototyping of advanced communication systems. He has co-led several National research programs, such as Program for Promoting Academic Excellence of Universities Phases I and II and 4G Mobile Communications Research Program initiated by Taiwan Government. Dr. Lee has won several awards for his research, engineering and teaching contributions; these include two times National Science Council (NSC) Superior Research Award, Young Electrical Engineer Award of the Chinese Institute of Electrical Engineers, and NCTU Teaching Award.