

A RECURSIVE QR APPROACH TO ADAPTIVE EQUALIZATION OF TIME-VARYING MIMO CHANNELS*

S. Y. KUNG[†], XINYING ZHANG[‡], AND CHAD L. MYERS^{†‡}

Abstract. This paper presents a novel adaptive equalization algorithm for time-varying MIMO systems with ISI channel conditions. The algorithm avoids channel estimation before equalization and leads to a direct QR-based procedure for updating the equalizer coefficients to track the time-varying channel characteristics. Our approach does not require precise channel estimation and needs relatively few pilot symbols for satisfactory equalization. The theoretical foundations of the proposed algorithm are rooted in signal recovery results derived from the generalized Bezout identity and the finite alphabet property inherent in digital communication schemes. Concerning the convergence behavior of the algorithm, we address the following three issues: *existence* of fixed points, *exclusiveness* of fixed points, and *robustness* under noise disturbance and parameter selection. The equalizer demonstrates promising capability in achieving low symbol error rates for a very broad range of SNRs. Simulation results are presented confirming that this approach outperforms the more traditional recursive least squares (RLS) adaptive equalizer for this application and rivals the performance of MMSE equalizers requiring channel knowledge.

1. Introduction. The major challenge in signal recovery for mobile communication systems is mitigating the ISI (Inter-Symbol-Interference), due to multi-path propagation and channel fading and distortion, as well as ICI (Inter-Channel-Interference), a result of multiple transmitters sharing the communication medium. When signals are transmitted over multiple independent fading channels, a variety of diversities is often available. For non-dispersive channels, static diversity-combining techniques are effective because simultaneous deep fades in all sub-channels are highly improbable. For example, the signal from each of the q diversity branches is weighted (in proportion to SNR) to produce the optimal combiner output in Maximum Ratio Combining (MRC). When channels are dispersive, however, it is necessary to resort to a more flexible FIR (finite-impulse-response) filtering, which is the approach used here. In general, equalization schemes can be classified as one of two types: equalizers that attempt to estimate the channel impulse response before equalization, or those which directly equalize the received signal with no attention to channel estimation [2]. The first of these categories has received a great deal of attention and several successful schemes have been developed (i.e. [3, 4, 5]). However, we focus on an equalizer of the second type because channel estimation is never perfect due to time-varying characteristics and noise, and this imprecision often leads to significant performance degradation in estimation and recovery. This assumption of an imprecise or unknown

*This research was supported in part by a grant from MERL, Mitsubishi Electric Research Laboratories. A brief version of this paper can be found in [1].

[†]The authors are with the Electrical Engineering Department, Princeton University, Princeton, NJ 08544. E-mail: kung@ee.princeton.edu, xinying@ee.princeton.edu, cmyers@princeton.edu

[‡]To whom all correspondence should be addressed.

channel is the motivation for our equalizer. Such an approach bears some resemblance to the well-known class of blind equalizers for MIMO systems ([6, 7, 8, 9, 10, 11, 12]). However, for the proposed equalizer, training is required but only when the algorithm fails to track a satisfactory equalizer. The required frequency and size of pilot symbol blocks is significantly less than that of traditional non-blind equalizers, implying a drastic reduction in bandwidth used for training.

The organization of this paper is as follows. In Section 1.1, a model for MIMO ISI channels is introduced. Sections 1.2 and 1.3 briefly reviews the basis of zero-forcing Bezout equalizer. Section 2 proposes our iterative equalization algorithm. Concerning the convergence of the iterative algorithm, we address three issues *existence*, *exclusiveness*, and *robustness* in Sections 3.1, 3.2, and 3.3 respectively. Specifically, we find that existence of a fixed point is inherently tied to the well-known perfect recoverability condition for MIMO systems with additional dependence on equalizer order and system delay. We also show that exclusiveness is guaranteed by the key "finite-alphabet exclusiveness" (FAE) property. In Section 4, we develop a computationally efficient recursive QR scheme for adaptive equalization in time-varying MIMO systems and in Section 5, present supporting simulation results.

1.1. MIMO Channel Model. We consider a frequency-selective fading MIMO model with t inputs and r outputs. Let $s_j(k)$ denote the signal sequence from input j and $h_{ij}(k)$ the channel response from input j to output i . Accordingly, the observation sequence $x_i(k)$ at i -th output can be expressed as a convolution:

$$(1) \quad x_i(k) = \sum_{j=1}^t \sum_{l=0}^L h_{ij}(l) s_j(k-l) + n_i(k),$$

where L is the maximal degree among all the channels (the ISI length is $L+1$) and $n_i(k)$ is the noise sequence at output i . Each of the transmit channels can be modelled as an FIR filter of order L or less. We denote L channel coefficient matrices of size $r \times t$: $\mathbf{H}(l) = [h_{ij}(l)]_{i,j}$.

We also introduce D -transform notation for the ISI MIMO model, where D denotes a unit time delay. In D -transform domain, the channel in (1) can be equivalently expressed as:

$$(2) \quad \underline{\mathbf{x}}(D) = \mathbf{H}(D)\underline{\mathbf{s}}(D) + \underline{\mathbf{n}}(D),$$

where $\underline{\mathbf{s}}(D)$ and $\underline{\mathbf{x}}(D)$ are $t \times 1$, $r \times 1$ respectively. The $r \times t$ polynomial matrix $\mathbf{H}(D)$ is referred to as the *transfer function* of the MIMO system.

1.2. Zero-Forcing Equalizer for Flat MIMO Channels. For flat-fading MIMO channels with $L=0$, if the transfer function H has full column rank, then a left inverse matrix G exists such that:

$$(3) \quad GI = I_t$$

An identity system I_t means that after equalization, the total system becomes interference-free. That is, there is a (virtual) direct communication path from each input stream to its estimate, with neither ISI nor CCI. In short, only the effect of thermal noise remains. The criterion in (3) is well-known as the zero-forcing constraint [13] [14] because, in the absence of noise, perfect symbol recovery can be accomplished by extracting one desired input stream while eliminating all the others through a linear combination of the receiver observation data.

There are infinitely many solutions that can satisfy the above zero-forcing (ZF) condition. A more intriguing question is how to design a noise-resilient ZF equalizer in the presence of the inevitable channel noise. For noisy channels, we have $\hat{\underline{s}}(k) = G\mathbf{H}\underline{s}(k) + G\underline{n}(k) = \underline{s}(k) + G\underline{n}(k)$. A commonly accepted solution is one which yields the minimum postprocessing noise power $\mathcal{E}[\|G\underline{n}(k)\|^2]$, or equivalently, maximum postprocessing SNR. For example, it is well-known that such an optimal equalizer for SIMO channels is the maximum ratio combining (MRC) equalizer. This notion may be generalized for MIMO channels as $G = \mathbf{H}^\dagger$, which yields a maximum postprocessing SNR among all the possible ZF equalizers. Here the notation $(\cdot)^\dagger$ denotes the Moore-Penrose inverse of a matrix. Unfortunately, such a simple formula works only for flat MIMO channels and is not naturally extendible to ISI MIMO channels, a subject to be treated in the subsequent discussion.

1.3. FIR Zero-Forcing Equalizer for ISI MIMO Channels. For ISI MIMO Channels, as illustrated in Figure 1, an equalizer can be viewed as a bank of space-time filters at the receiver side, all with FIR (finite impulse response) and linear characteristics. In other words, the r received signals pass through an FIR array and the filtered outputs are linearly combined to obtain an estimate of an individual source signal. Let the matrix $\mathbf{G}(D)$ denote the FIR coefficient matrix of the equalizer, where each row $\underline{\mathbf{g}}_i(D)$ represents the FIR array for the estimate of a particular input. Then the zero-forcing constraint requires that

$$(4) \quad \mathbf{V}(D) - \mathbf{G}(D)\mathbf{H}(D) = I_t,$$

where $\mathbf{V}(D)$ is the combined channel-equalizer transfer function. For simplicity, we focus on the recovery of one input stream, say, source i . Then it requires

$$(5) \quad \mathbf{v}_i(D) - \underline{\mathbf{g}}_i(D)\mathbf{H}(D) = \underline{\mathbf{e}}_i,$$

where $\underline{\mathbf{e}}_i$ is a $t \times 1$ vector with all entries zero except 1 at position i . For noise free channels, the i -th source $s_i(k)$ can be recovered as

$$(6) \quad \hat{s}_i(D) = \underline{\mathbf{g}}_i(D)\mathbf{H}(D)\underline{\mathbf{s}}(D) = s_i(D).$$

On the other hand, for noisy channels, we have

$$(7) \quad \begin{aligned} \hat{s}_i(D) &= \underline{\mathbf{g}}_i(D)\mathbf{H}(D)\underline{\mathbf{s}}(D) + \underline{\mathbf{g}}_i(D)\underline{\mathbf{n}}(D) \\ &= s_i(D) + \underline{\mathbf{g}}_i(D)\underline{\mathbf{n}}(D). \end{aligned}$$

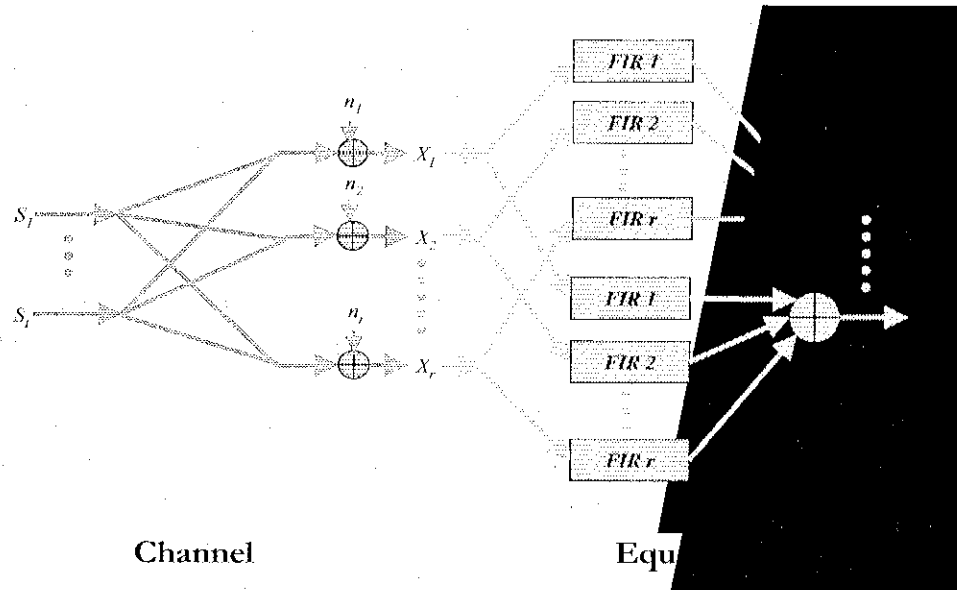


FIG. 1. An FIR zero-forcing equalizer for ISI MIMO channels

An optimally robust FIR ZF filter should yield a maximum postprocessing SNR. Its design is much more complex than its flat MIMO counterpart, and it involves two additional design parameters arising from the ISI channels: (1) the equalizer order; and (2) the optimal chosen system delay, which will be discussed in greater detail in Section 3.

2. Iterative Signal Recovery Algorithm. When channel state information (CSI) is available at the receiver, signal recovery can be accomplished in a straightforward manner by finding the polynomial inverse of $\mathbf{H}(D)$. However, in practice, precise CSI may not be readily or reliably available at the receiver for the following reasons: first, pilot symbols used to estimate the channel could be too expensive, especially when the channel is changing very fast; secondly, even if sufficient pilot symbols are available, the channel estimate could still be very imprecise, making the channel appear totally or partially blind to the receivers. The lack of CSI makes it impossible to find a closed-form expression for the equalizer coefficients.

There are in general two plausible approaches to handling the problem caused by the imprecise or blind channels: (1) methods based on an optimization formulation (e.g. CMA [15]) and (2) iterative updating (e.g. EM type) schemes. In an iterative scheme, a *joint estimate* is derived by iterating between the two spaces (equalizer taps and source symbols) until it reaches a fixed point. We focus on the latter and outline the key steps of our iterative approach below. Without loss of generality, assume the objective is the recovery of the first stream $s_1(k)$:

1. Predefine equalizer order and system delay from the feasible region (cf. Sec-

tion 3.3);

2. Obtain a proper initial estimate on the signal sequence $\hat{s}_1^{(0)}(k)$ with a block of symbols drawn from the digital constellation;
3. After the initial setting, iterate on the following two sub-steps until the convergence is reached:
 - (a) Fix the estimate of $\hat{s}_1^{(i)}(k)$ (as if it is correct) and find the equalizer coefficients $\hat{g}(k)$ that minimize the detection error¹:

$$(8) \quad \hat{g}^{(i+1)}(k) = \arg \min_{\hat{g}(k)} \|\hat{s}_1^{(i)}(k) - \hat{g}(k) * \underline{x}(k)\|;$$

- (b) To drive $\hat{s}_1(k)$ towards a valid finite alphabet sequence, the new estimate $\hat{s}_1^{(i+1)}(k)$ is updated to the nearest symbols in the constellation:

$$(9) \quad \hat{s}_1^{(i+1)}(k) = \arg \min_{\hat{s}_1(k) \in \mathcal{FA}} \|\hat{s}_1(k) - \hat{g}^{(i+1)}(k) * \underline{x}(k)\|,$$

where \mathcal{FA} is the set of all finite alphabet sequences.

The idea of iterating between (8) and (9) was proposed earlier by [16, 17]. The two equations in (8) and (9) can be represented in a matrix-vector form, which plays an instrumental role for the robustness analysis in Section 3.3. For any two positive integers ρ and N , we introduce a $(\rho+1) \times N$ block Toeplitz matrix $\Gamma^{\rho,N}[\underline{x}]$ of output signals:

$$(10) \quad \Gamma^{\rho,N}[\underline{x}] \equiv \begin{bmatrix} \underline{x}(N-1) & \underline{x}(N-2) & \dots & \underline{x}(0) \\ \underline{x}(N) & \underline{x}(N-1) & \dots & \underline{x}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \underline{x}(N+\rho-1) & \underline{x}(N+\rho-2) & \dots & \underline{x}(\rho) \end{bmatrix},$$

where $\underline{x}(k) = [x_1(k) \ x_2(k) \ \dots \ x_r(k)]^T$. The equalizer filters $\underline{g}(k)$ with order ρ (or, $\mathbf{g}(D)$ in D -transform domain) can also be represented in an extended vector form:

$$(11) \quad \Upsilon^\rho[\underline{g}] = [\underline{g}(\rho) \ \underline{g}(\rho-1) \ \dots \ \underline{g}(0)].$$

Similarly, we stack the N consecutive symbols of the signal estimate into a data vector:

$$(12) \quad \Gamma^{0,N}[\hat{s}_1] = [\hat{s}_1(N+\rho-1) \ \hat{s}_1(N+\rho-2) \ \dots \ \hat{s}_1(\rho)].$$

Note that the convolution of the equalizer coefficients with the data sequences $\underline{x}(k)$ is equivalent to the multiplication of the corresponding Toeplitz-structured matrices. More precisely, if the channel is noise-free, (6) can be written as

$$(13) \quad \Gamma^{0,N}[\hat{s}_1^{(i)}] = \Upsilon^\rho[\hat{g}] \Gamma^{\rho,N}[\underline{x}].$$

¹Here the notation $\|\cdot\|$ denotes the standard 2-norm over two sequences, indexed by (k) , i.e. two infinite-dimensional vectors.

For noisy channels, (8) and (9) are rewritten in the following matrix-vector forms respectively:

$$(14) \quad \Upsilon^\rho[\hat{g}^{(i+1)}] = \arg \min_{\Upsilon^\rho[\hat{g}]} \|\Gamma^{0,N}[\hat{s}_1^{(i)}] - \Upsilon^\rho[\hat{g}]\Gamma^{\rho,N}[\underline{x}]\|;$$

$$(15) \quad \Gamma^{0,N}[\hat{s}_1^{(i+1)}] = \arg \min_{\Gamma^{0,N}[\hat{s}_1] \in \mathcal{FA}} \|\Gamma^{0,N}[\hat{s}_1] - \Upsilon^\rho[\hat{g}^{(i+1)}]\Gamma^{\rho,N}[\underline{x}]\|.$$

Fortunately, the matrix formulation admits a simple least square solution to (14) and non-linear mapping solution to (15):

$$(16) \quad \Upsilon^\rho[\hat{g}^{(i+1)}] = \Gamma^{0,N}[\hat{s}_1^{(i)}](\Gamma^{\rho,N}[\underline{x}])^\dagger;$$

$$(17) \quad \Gamma^{0,N}[\hat{s}_1^{(i+1)}] = T(\Upsilon^\rho[\hat{g}^{(i+1)}]\Gamma^{\rho,N}[\underline{x}]),$$

where $T(\cdot)$ denotes the mapping to the nearest valid symbol in the digital constellation. Furthermore, the two parts in each loop can be combined to yield an update formula for the estimate of the signal (e.g. $s_1(k)$):

$$(18) \quad \Gamma^{0,N}[\hat{s}_1^{(i+1)}] = T\{\Gamma^{0,N}[\hat{s}_1^{(i)}](\Gamma^{\rho,N}[\underline{x}])^\dagger\Gamma^{\rho,N}[\underline{x}]\},$$

by replacing $\Upsilon^\rho[\hat{g}^{(i+1)}]$ with its equivalent expression from (16) in (17).

In other words, each update from the previous estimate $\Gamma^{0,N}[\hat{s}_1^{(i)}]$ to the current estimate $\Gamma^{0,N}[\hat{s}_1^{(i+1)}]$ involves the following operations:

1. Least-Square Approximation (linear subspace projection):

The multiplication of $\Gamma^{0,N}[\hat{s}_1^{(i)}]$ with a weighting matrix

$$W = (\Gamma^{\rho,N}[\underline{x}])^\dagger\Gamma^{\rho,N}[\underline{x}]$$

implies a linear projection of the vector $\Gamma^{0,N}[\hat{s}_1^{(i)}]$ into the subspace spanned by the row vectors in $\Gamma^{\rho,N}[\underline{x}]$.

2. Nonlinear Decision Making (finite alphabet mapping):

The projected vector is mapped into the finite alphabet constellation to obtain the new estimate vector $\Gamma^{0,N}[\hat{s}_1^{(i)}]$.

A geometric illustration of the iterations for symbol-equalizer joint estimation is given in Figure 2.

3. Theoretical Foundation. For an iterative algorithm, the most critical concern lies in its *convergence* behavior. For this, the following three issues must be considered:

1. *Existence*: Will there exist correct fixed points for the signal recovery iterations using an FIR combiner? If so, what minimal FIR order and system delay are necessary?
2. *Exclusiveness*: In the noise free case, how do we assure that the system will not converge to a wrong solution?

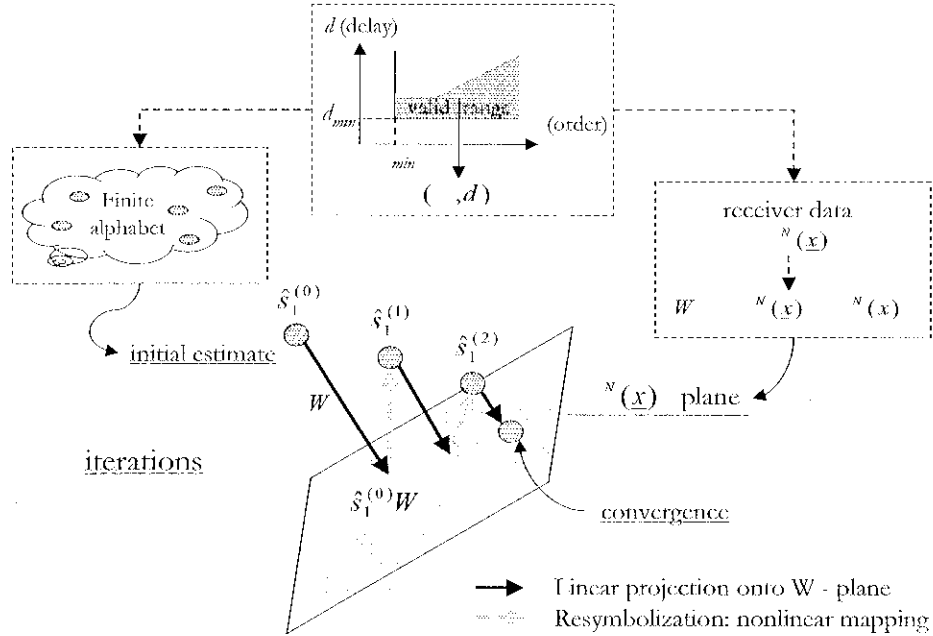


FIG. 2 Geometric illustration of the iterative signal recovery algorithm. With no noise, a symbol vector and some of its delayed versions lie on the row subspace spanned by $\Gamma^{p,N}[\underline{x}]$ (They correspond to fixed points of the iterative algorithm.) The initial conditions for the iterations are set by (1) preselecting valid parameters for the equalizer order and system delay; (2) forming the initial estimate vector on $s_1(k)$ from the digital constellation with the desired system delay; (3) with the equalizer order selected, forming the receiver data matrix $\Gamma^{p,N}[\underline{x}]$; (4) pre-determining the weighting matrix W from $\Gamma^{p,N}[\underline{x}]$. Given an initial signal estimate vector, a desirable solution for the equalizer $\Upsilon^p[\hat{g}^{(i+1)}]$ is chosen such that $\Upsilon^p[\hat{g}^{(i+1)}]\Gamma^{p,N}[\hat{s}_1^{(i)}]$ lies on the $\Gamma^{p,N}[\underline{x}]$ -plane and is the closest vector (relative to $\Gamma^{p,N}[\hat{s}_1^{(i)}]$). Geometrically, such a vector $\Upsilon^p[\hat{g}^{(i+1)}]\Gamma^{p,N}[\hat{s}_1^{(i)}]$ is none other than the projection of $\Gamma^{p,N}[\hat{s}_1^{(i)}]$ onto the $\Gamma^{p,N}[\underline{x}]$ -plane (shown by solid arrow). In the noisy case, $\Gamma^{p,N}[\hat{s}_1^{(i)}]$ is likely to be only an approximate symbol sequence. Therefore, we enact a nonlinear FA mapping process to replace an approximate symbol by its corresponding nearest valid symbol (shown as dashed arrow). If necessary, the projection and FA mapping iterations are repeated until convergence is reached. In practice, however, it appears one iteration suffices to perform an adequate updating.

3 *Robustness*: How do we make the algorithm’s convergence behavior robust against noise disturbance? How do we exercise an effective control of the search space (linearly proportional to the equalizer order)?

We elaborate on these three issues in the following sections.

3.1. Existence of Fixed Points for Convergence in Noise-free Channels.

For simplicity, let us temporarily ignore the effect of noise. Then the necessary and sufficient condition for the existence of fixed points for the proposed iterative signal recovery algorithm is identical to the “perfect recoverability” condition of a MIMO

system, which is described below:

DEFINITION 1 (Perfect Recoverability (PR)). *Given a MIMO system with transfer function $\mathbf{H}(D)$, the i -th input is said to be perfectly recoverable (PR) of order ρ if and only if there exists a nonnegative integer τ and a polynomial (row) vector $\underline{\mathbf{g}}(D)$ with degree no larger than ρ such that*

$$(19) \quad \underline{\mathbf{g}}(D)\mathbf{H}(D) = \mathbf{e}_i D^\tau.$$

The filter array denoted by $\underline{\mathbf{g}}(D)$ is said to be a (i, ρ, τ) equalizer of the system. In general, a MIMO system is said to be PR if and only if all the t inputs are PR of a finite order. \square

The condition for perfect recoverability is inherently tied to three properties of a MIMO system: channel characteristics, system delay, and equalizer order. We discuss each of these in detail below.

3.1.1. Condition on the Channel. The effect of channel characteristics on the existence of fixed points is established through the (Generalized) Bézout Identity. When applied to the integer ring, the Bézout Identity [18] gives the famous result that, for any two coprime integers a and b , there always exist another two integers c and d such that the identity equation $ac + bd = 1$ holds. This is also valid on the polynomial ring, when a , b , c , d are replaced by the finite-order polynomials $a(D)$, $b(D)$, $c(D)$ and $d(D)$, respectively. It can also be expressed in an equivalent vector form as $\begin{bmatrix} c(D) & d(D) \end{bmatrix} \begin{bmatrix} a(D) \\ b(D) \end{bmatrix} = 1$.

The Bézout Identity and the test of coprimeness were further extended to the polynomial matrix case, cf. [19][20], which are instrumental to MIMO equalization applications. More precisely, given an FIR MIMO channel $\mathbf{H}(D)$, there exists a polynomial matrix $\mathbf{G}(D)$ satisfying

$$(20) \quad \mathbf{G}(D)\mathbf{H}(D) = I_t$$

if and only if $\mathbf{H}(D)$ is right coprime. This satisfies the PR criterion in (19) with zero delay.²

Note that the notion of perfect recoverability in Definition 1 relaxes the condition on the channel transfer function to allow for the possibility of a delayed recovery. For this, a notion of delay-permissive coprimeness is introduced. A polynomial matrix (of D) is said to be delay-permissive (right) coprime if and only if the determinant

²A polynomial matrix $\mathbf{C}(D)$ is said to be a right common divisor (red) of the rows in $\mathbf{H}(D)$ if a finite-order polynomial matrix $\mathbf{H}'(D)$ can be found to support the factorization $\mathbf{H}(D) = \mathbf{H}'(D)\mathbf{C}(D)$. A matrix $\mathbf{R}(D)$ is the gred of $\mathbf{H}(D)$ if and only if any red $\mathbf{R}'(D)$ of $\mathbf{H}(D)$ is also an red of $\mathbf{R}(D)$. The matrix $\mathbf{H}(D)$ is called right coprime if and only if its right common divisors are all unimodular matrices. (A polynomial matrix is unimodular if and only if its determinant is a constant, rather than a polynomial.)

of any its gcd (greatest right common divisor) has the form of a pure delay D^r . A MIMO channel is PR if and only if the transfer function $\mathbf{H}(D)$ is delay-permissive right coprime. A more elaborate theorem summarizing these observations is stated below:

THEOREM 1 (Perfect Recoverability Condition of MIMO Systems). *A t -in- r -out MIMO channel transfer function $\mathbf{H}(D)$ is perfect recoverable if and only if any of the following (equivalent) conditions are true:*

1. *There exists a $r \times t$ polynomial matrix $\mathbf{G}(D)$ with finite order satisfying*

$$(21) \quad \mathbf{G}(D)\mathbf{H}(D) = \text{diag}\{D^{\tau_i}\}_{i=1}^t;$$

2. *$\mathbf{H}(D)$ is delay-permissive right coprime;*
3. *$\mathbf{H}(\lambda)$ has full column rank for any complex value $\lambda \neq 0$.*

These three conditions are equivalent to each other.

Proof. Omitted here. See [19, 21, 22]. □

3.1.2. Condition on the System Delay. If a MIMO channel satisfies the PR requirement, there is in general a necessary minimum delay elapsed before an input signal can be reconstructed by an FIR equalizer. This notion of minimum system delay can be illustrated by the following example.

EXAMPLE 1 (System Delay of Perfect Signal Recovery). *We consider a 2-in-3-out MIMO system with the transfer function*

$$\mathbf{H}(D) = \begin{bmatrix} D+1 & D^2+D+1 \\ D+2 & 2D^2+D+2 \\ D-1 & -D^2+D-1 \end{bmatrix}.$$

A gcd of $\mathbf{H}(D)$ is $\mathbf{R}(D) = \begin{bmatrix} D & D \\ 1 & D^2+1 \end{bmatrix}$. Therefore $\mathbf{H}(D)$ is delay-permissive right coprime and the MIMO system is PR. More specifically, we have $\mathbf{G}(D)\mathbf{H}(D) = \begin{bmatrix} D^3 & 0 \\ 0 & D^3 \end{bmatrix}$ by setting $\mathbf{G}(D) = \begin{bmatrix} -D^2 + \frac{1}{4}D - 1 & D^2 - \frac{1}{2}D + 1 & D^2 + \frac{1}{4}D + 1 \\ \frac{D-1}{5} & \frac{D-1}{5} & \frac{2D+3}{5} \end{bmatrix}$. It can be further verified that the system delays of the two inputs upon recovery are both 3 and cannot be reduced any further. □

The minimum system delays for different input signals can be different. These delays are denoted by $\{\tau_j\}_{j=1}^t$, corresponding to each input. It can be shown that if $\mathbf{H}(D)$ is delay-permissive right coprime, then $\max_j\{\tau_j\} < \text{deg}|\mathbf{R}(D)|$, where $\mathbf{R}(D)$ is the gcd of $\mathbf{H}(D)$.

3.1.3. Condition on the Equalizer Order. In addition to the condition on the MIMO channel and recovery delay, the existence of fixed points is also related to the equalizer order. In other words, there is a minimal degree requirement on the matrix $\mathbf{G}(D)$ in (20). In general, different degrees are required for the recovery of

different input signals. As to the sufficient degrees of Bézout equalizers given $\mathbf{H}(D)$, a very comprehensive treatment on the degree bound is derived in [22]

THEOREM 2 (Sufficient Order of Bézout Equalizer for General MIMO). *Suppose that the MIMO transfer function $\mathbf{H}(D)$ is delay-permissive right coprime (i.e. it is PR) and that $\tau \geq \tau_i$, where τ_i is the minimum system delay for the corresponding source. There exists a (i, ρ, τ) PR Bézout equalizer if*

$$(22) \quad \rho \geq \max\{\nu - 1, \tau + L - 1\},$$

where ν is the degree of the null-space minimum basis for $\mathbf{H}(D)$.

Proof. For the proof and the definition of null-space minimum bases, see [22] and the Appendix. \square

Two useful corollaries follow the theorem above immediately:

COROLLARY 1 (Sufficient Equalizer Orders) *Given a t -in- r -out MIMO system with a coprime transfer function $\mathbf{H}(D)$, a sufficient degree to reconstruct any signal is*

$$(23) \quad \rho \geq \max\{\nu - 1, L - 1\},$$

where L is the maximum ISI length of the channel and ν is the degree of null-space minimum basis of the transfer function $\mathbf{H}(D)$. \square

COROLLARY 2 (Sufficient Equalizer Order for Column-Reduced Systems). *Note that if $\mathbf{H}(D)$ is already column-reduced, then tighter degree bounds can be derived:*

1. A sufficient degree to reconstruct $\mathbf{s}_i(D)$ is $\rho \geq \nu - 1$; and
2. A sufficient degree to reconstruct $\mathbf{s}_i(D)D^\tau$ (assuming that $\tau > \tau_i$) is [22]

$$(24) \quad \rho \geq \max\{\nu - 1, \tau\}.$$

\square

EXAMPLE 2 (An Example on the Equalizer Order). *We consider a 2-in-3-out MIMO system with the transfer function*

$$\mathbf{H}(D) = \begin{bmatrix} D+2 & D+1 \\ 1 & D \\ D & 1 \end{bmatrix}.$$

For this MIMO channel, a null basis (with minimal degree) is $\mathbf{N}(D) = [D^2 - 1 \quad -D^2 + 2 \quad -D^2 - D + 1]$. Thus, the minimal null basis has a degree $\nu = 2$, and consequently, any equalizer with order $\rho \geq \max\{\nu - 1, L - 1\} = 1$ should suffice. Indeed, an equalizer with the transfer function $\mathbf{G}(D) = \begin{bmatrix} 1 & 1 & 1 \\ -D & D & D+1 \end{bmatrix}$ can perfectly recover both input signals. The first input is recoverable with a zeroth-order equalizer, while the second input requires a first-order equalizer. Moreover, as the MIMO system above happens to be column-reduced, the tighter degree bounds in Corollary 2 also apply. For

example, a sufficient degree to reconstruct a one-unit delayed source signal $s_i(D)D$, is $\rho \geq \max\{\nu - 1, 1\} - 1$. This can be verified by noting that $\mathbf{C}'(D)\mathbf{H}(D) = DI_L$, where $\mathbf{C}'(D) = \begin{bmatrix} D & -D & D \\ -1 & 2 & 1 \end{bmatrix}$. \square

In practice, we generally do not have direct knowledge of ν , while we do have a reasonable estimate of the channel order L . Thus, the following statistical average value for ν may prove useful:

$$(25) \quad \nu \approx \frac{Lt}{r-t}$$

3.2. Exclusiveness of Fixed Points. In order to establish the exclusiveness of convergent points, we must exploit the pivotal *finite alphabet property* inherent in digital communication systems, where each valid symbol in an information stream is an element drawn from a finite set. A sequence is called a (valid) *finite alphabet sequence* if and only if each of its symbols is one of the M valid points in a digital constellation. For examples, BPSK signaling has $M = 2$ symbols, while the QPSK constellation has $M = 4$ symbols symmetrically located in the complex plane. For digital transmission, it is critical to make use of anFAE (Finite Alphabet Exclusiveness) property, originally proposed in [16], as stated below: Given t finite alphabet sequences $s_1(k), \dots, s_t(k)$ and t weighting factors v_1, \dots, v_t , it is then necessary that

$$v = \begin{bmatrix} v_1 & v_2 & \dots & v_t \end{bmatrix} = c\mathbf{e}_i$$

so the linear combination $\sum_{i=1}^t v_i s_i(k)$ is a valid finite alphabet sequence given any finite alphabet input patterns $s_1(k), \dots, s_t(k)$. Here c is a constant, e.g. $c = \pm 1$ in the case of BPSK constellation (or $c = e^{j\pi n/2}$ for any integer n in case of QPSK). For the original proof of the binary alphabet case, the reader is referred to [16].

Here we outline a proof for the general case with M distinct symbols $\{o_m\}_{m=1}^M$. Assume that $u(k) = \sum_{i=1}^t v_i s_i(k)$ is a valid finite alphabet sequence for any pattern in the input finite alphabet sequences $s_1(k), \dots, s_t(k)$. WLOG, let us assume $v_1 \neq 0$ and consider the following patterns in the input sequences:

1. $s_1(k) = s_1(M+k) = o_k$, for $k = 1, \dots, M$;
2. For each input sequence $s_i(k)$ ($i = 2, \dots, t-1$), all the $2M$ symbols $\{s_i(k)\}_{k=1}^{2M}$ have the same value;
3. $s_t(1) = s_t(2) = \dots = s_t(M) \neq s_t(M+1) = s_t(M+2) = \dots = s_t(2M)$.

It is then straightforward to show that for $k \neq l \in \{1, \dots, M\}$, $u(k) - u(l) = v_1(o_k - o_l) \neq 0$, therefore $\{u(k)\}_{k=1}^M$ are all distinct. By the same token, $\{u(k)\}_{k=M+1}^{2M}$ are all

distinct. Thus, we have

$$\begin{aligned}
 & \sum_{k=1}^M u(k) - \sum_{k=M+1}^{2M} u(k) \\
 \rightarrow & v_t \sum_{k=1}^M s_t(k) - v_t \sum_{k=M+1}^{2M} s_t(k) \\
 (26) \quad & \implies v_t = 0.
 \end{aligned}$$

Thus we conclude that $v_t = 0$. Similar procedures can be adopted to show $v_i = 0$ ($i = 2, \dots, t$). \square

The FAE property can be extended straightforwardly to the FIR filter case which is vital for FIR equalization of ISI MIMO channel.

THEOREM 3 (FAE Property for FIR Filters). *Let $\mathbf{s}(D)$ denote t finite alphabet sequences and suppose that the $1 \times t$ polynomial vector $\underline{\mathbf{v}}(D)$ represents a linear FIR system. For the output $\mathbf{v}(D)\mathbf{s}(D)$ to always be a valid symbol sequence given any input symbol sequence $\mathbf{s}(D)$, it is necessary that $\underline{\mathbf{v}}(D) = cD^r \mathbf{e}_i$. In other words, given t (sufficiently long and random) finite alphabet sequences, it is impossible to produce by linear FIR filters a FA output sequence other than one of the t original sequences or its scaled and/or delayed version.*

Proof. To extend the FAE to FIR filters, we note that any FIR convolution $\mathbf{v}(D)\mathbf{s}(D)$ can be expressed as a matrix operation: $[\underline{v}(L + \rho) \ \dots \ \underline{v}(0)] \Gamma^{L + \rho, N} \mathbf{s}$. Hence, the FAE property for flat MIMO channels can be extended to ISI MIMO channels if such channels and equalizers can be modelled as banks of FIR filters.

Failure Rate of the FAE Property.

the FAE Theorem relies on the assumption that the input finite alphabet sequence is sufficiently long. Thus, it is possible to construct examples violating the FAE property, but we argue that such conditions rarely occur in most practical applications. In fact, the failure rate for binary-symbol case was quantified in [16]. In short, the FAE property is a practical assumption as long as N is reasonably large.

Duality Between Channel Identifiability and Equalization Exclusiveness.

Without invoking the FAE property, it is well-known that a MIMO channel is identifiable only up to a transformation of a unimodular matrix. In other words, the distinction between $\mathbf{H}(D)$ and $\mathbf{H}(D)\mathbf{U}(D)$ is unresolvable for any unimodular matrix $\mathbf{U}(D)$. This is illustrated below. Note that

$$\mathbf{x}(D) = \mathbf{H}(D)\underline{\mathbf{s}}(D) = \mathbf{H}(D)\mathbf{U}(D)\mathbf{U}'(D)\underline{\mathbf{s}}(D) = \mathbf{H}'(D)\underline{\mathbf{s}}'(D),$$

where $\mathbf{U}'(D) = \mathbf{U}^{-1}(D)$ is itself a polynomial matrix. Moreover, $\underline{\mathbf{s}}'(D) = \mathbf{U}^{-1}(D)\underline{\mathbf{s}}(D)$ and $\mathbf{H}'(D) = \mathbf{H}(D)\mathbf{U}(D)$. Since $\underline{\mathbf{s}}'(D)$ is also a polynomial vector, it would qualify as a permissible input vector if no finite alphabet constraint is imposed. Consequently, either $\mathbf{H}'(D)$ or $\mathbf{H}(D)$ could represent the original MIMO system, i.e. they are not distinguishable.

From an equalization perspective, if $\mathbf{G}(D)\mathbf{H}(D) = I_t$, then $\mathbf{G}'(D) = \mathbf{U}'(D)\mathbf{G}(D)$ would be a Bézout equalizer for $\mathbf{H}'(D)$, since $\mathbf{G}'(D)\mathbf{H}'(D) = \mathbf{U}'(D)\mathbf{G}(D)\mathbf{H}(D)\mathbf{U}(D) = I$. Note that $\mathbf{G}'(D)\underline{\mathbf{x}}(D) = \mathbf{G}'(D)\mathbf{H}'(D)\underline{\mathbf{s}}'(D) = \underline{\mathbf{s}}'(D)$. Thus, $\mathbf{G}'(D)$ would qualify as a fixed-point of the iterative algorithm if $\underline{\mathbf{s}}(D)$ is recognized as a valid symbol vector.

Fortunately, such ambiguity can be largely resolved once the FA constraint is imposed. Under the FA constraint, possible values of $\underline{\mathbf{s}}'(D)$ are much more limited. In fact, according to Theorem 3, in order for $\underline{\mathbf{s}}'(D)$ to be a valid FA input vector, its element must be a scaled and/or delayed version of the source signal. This implies that the matrix $\mathbf{U}'(D)$ must be restricted to the following form:

$$\mathbf{U}'(D) = P \text{diag}\{c_i D^{\tau_i}\}_{i=1}^L,$$

where P is a permutation matrix. In other words, with the imposition of the FAE property, the system becomes basically identifiable except for a scaling factor, a system delay, and/or a permutation on the sources.

Summary of Existence and Exclusiveness Results.

In the previous two sections, we have addressed two of the three concerns about our iterative signal recovery algorithm: the existence of fixed points, and the exclusivity of fixed points to correct solutions. For emphasis, we summarize these key results in the following two theorems.

THEOREM 4 (Existence and Exclusiveness of Fixed Points of MIMO Equalization Algorithm [22]).

1. **Existence of Fixed Points for MIMO Equalization Algorithm:** *Given a MIMO system, the proposed iterative algorithm prescribed by the equations (8) and (9) has a fixed point at $s_i(k - \tau)$ if*
 - (1) $\mathbf{H}(D)$ is perfectly recoverable.
 - (2) the estimated signal has incorporated an adequate system delay: for tracking of the i -th source sequence, $i = 1, \dots, L$:

$$(27) \quad \tau \geq \tau_i,$$

where τ_i is the minimal system delay the i -th source signal.

- (3) adopts an adequate equalizer order:

$$(28) \quad \rho \geq \max\{\nu - 1, L + \tau - 1\}.$$

2. **Exclusiveness of Fixed Points for MIMO Equalization Algorithm:** *According to Theorem 3, any fixed point must be a vector corresponding to one of the original source signals with system delay τ , where*

$$(29) \quad 0 \leq \tau \leq L + \rho - 1$$

Therefore, for noise-free channels, a sufficient equalizer order (as prescribed by Theorem 1) guarantees that, (1) at least one fixed point always exists; and (2) every fixed point must correspond correctly to one of the source EA sequences. \square

3.3. Robustness of Convergence. So far, we have only considered the algorithm's behavior under the idealistic assumptions that the system is noise-free and that existing fixed points will eventually be reached. In practice, thermal noise present at the receivers and the size of the search space, which is linearly proportional to equalizer order, can significantly affect the stability of the convergence. In light of these concerns, we pose and answer the following questions:

- How do we make the algorithm's convergence behavior robust against noise disturbance?
- How do we prevent an unnecessarily large search space?

3.3.1. Postprocessing SNR Gain and Min-Norm Solution. We first investigate the notion of *postprocessing SNR gain*. Assume that the thermal noise at receiving antennas is i.i.d. white Gaussian with covariance N . Note that, with more receivers than transmitters, the choice of $\mathbf{G}(D)$ satisfying the Generalized Bézout Identity in (20) is highly non-unique. From an SNR perspective, different choices of $\mathbf{G}(D)$ will lead to different noise amplification.

According to (7), after the application of an equalizer $g(k)$, the noise will be amplified (or attenuated) and the following noise power at the filter output will be produced: $\mathcal{E}[|g(k) * \underline{n}(k)|^2] = \sigma_n^2 \|\Upsilon^\rho[\underline{g}]\|^2$, where $\Upsilon^\rho[\underline{g}]$ denotes the expanded column vector as defined in (11). The postprocessing SNR gain is defined as the ratio of the SNR after equalization to the SNR before equalization. Because the original signal power is preserved in zero-forcing equalizers, it follows that the postprocessing SNR gain is $\frac{1}{\|\Upsilon^\rho[\underline{g}]\|^2}$. It is desirable to adopt an equalizer which yields the highest possible postprocessing SNR gain in order to obtain a more robust convergence in a noisy environment. It is obvious that a minimum-norm solution of the inverse system is the most desirable choice.

When the order and delay are fixed, the equalizer in (19) can be rewritten in a matrix form:

$$(30) \quad \Upsilon^\rho[\underline{g}]\Gamma^\rho[H] = \mathbf{e}_{j+(\rho+L-\tau)\nu}$$

where τ is the system delay and $\Gamma^\rho[H]$ is the $r\rho \times l(L + \rho)$ block Toeplitz *resultant matrix* of the MIMO channel:

$$(31) \quad \Gamma^\rho[H] = \begin{bmatrix} H(L) & H(L-1) & \dots & H(0) & \mathbf{0} & \dots & \mathbf{0} \\ 0 & H(L) & \dots & H(1) & H(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H(L) & \dots & H(1) & H(0) \end{bmatrix}$$

Recall that in the existence analysis, we established the condition on the order, ρ , and delay, τ , to guarantee the PR of the signal, and therefore, the existence of solutions to the above equation. The optimal SNR gain of a zero-forcing equalizer can now be quantified and the optimal SNR (or minimum 2-norm solution) to (30) is

$$(32) \quad \Upsilon^\rho[\underline{g}] = e_{j+(\rho+L-1-\tau)t}(\Gamma^\rho|H|)^\dagger.$$

The corresponding postprocessing SNR gain is denoted by

$$(33) \quad \gamma_j(\rho, \tau) = \frac{1}{\|e_{j+(\rho+L-1-\tau)t}(\Gamma^\rho|H|)^\dagger\|^2}.$$

3.3.2. Selection of Equalizer Order and Elapsed Delay. The postprocessing SNR gain can be improved by (1) adopting optimal recovering FIR order and (2) purposefully imposing some system delay. Improper parameters may adversely affect the bit-error-rate in signal recovery, so the order and delay should be carefully pre-selected.

A. Optimal Equalizer Order

The optimal equalizer order depends on the following trade-offs:

1. Given a precisely known channel, the 2-norm of the optimal equalizer vector $\underline{g}(D)$ is a monotonically decreasing function with respect to the FIR order ρ [23]. Therefore, theoretically, the SNR gain will improve with increasing FIR order.
2. When the channel is unknown, a higher ρ implies (a) a greater failure possibility of the FAP property, which could result in convergence to incorrect solution, and (b) an unnecessary expansion of search space, which could hamper the convergence process.

Thus, the equalizer order must be chosen in such a way that a desirable postprocessing SNR can be achieved while still maintaining a manageable search space.

B. Optimal System Delay

For SNR optimization, the system delay can be more or less treated as a free parameter. The latency due to a system delay corresponding to few symbol intervals should not be a serious problem for most practical applications. For example, a system delay of 10 symbols would amount only to a $10\mu s$ delay for a symbol rate of $1MHz$, which is quite negligible. Given the FIR order ρ , the possible range of the system delay (cf. (29)) is $j \in [0, 1, \dots, \tau + \rho - 1]$. Furthermore, the 2-norm, $\|\Upsilon^\rho[\underline{g}]\|$, displays a U-shape, when plotted as a function of system delay τ , i.e. they have lower values in the middle portion of the possible delay range $[0, 1, 2, \dots, L + \rho - 1]$. Hence the optimal system delay can be empirically chosen as $\tau^* = \frac{\rho+L}{2}$.

C. Empirical Choice of Order and Delay

If the receiver diversity is at least twice as large as the transmitter diversity, say $r \approx 2L$, then it is likely that $\nu \leq L$. Empirically, this amounts to the choice of an equalizer order of L . It follows that the system delay can be set to $\tau = L$. Note that the order may need to be tuned slightly higher when the receiver diversity is less than twice the transmitter diversity. According to (25): $\nu \approx Lt/(r - t)$, assuming that the MIMO transfer function is column-reduced, then for a robust Bézout equalizer, the order can be empirically chosen as $\rho = \max\{L, Lt/(r - t)\}$ based on (24) and the delay is $\rho + L$.

4. Time-varying Channel Model and Adaptive Equalization Algorithms. In this section, we extend the iterative signal recovery algorithm in Section 2 to time-varying MIMO channels by developing recursive equalization schemes capable of fast computation with minor hardware overhead. First, we introduce the model used to simulate a time-varying MIMO ISI channel.

4.1. Time-Varying MIMO Channel Model. A time-varying MIMO system under the influence of ISI is modelled as before, referring to (1), but the channel coefficients $\{h_{ij}(l)\}_{l=0}^L$ are now replaced by the time-varying coefficients $\{h_{ij}^{(k)}(l)\}_{l=0}^L$ where k is the time index. As proposed by Kominakis in [3], each coefficient is the sum of a constant and time-varying part:

$$h_{ij}^{(k)}(l) = c_{ij}(l) + \bar{h}_{ij}^{(k)}(l),$$

where $h_{ij}^{(k)}(l)$ is a zero-mean, wide-sense-stationary, complex Gaussian process whose time-autocorrelation is related to the Doppler rate fT (Doppler frequency multiplied by the symbol period). More specifically,

$$(34) \quad E\{h_{ij}^{(k)}(l)h_{ij}^{(k-1)}(l)^*\} = J_0(2\pi f_{ij}T),$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind. The Doppler frequencies, f_{ij} , are computed using simulated relative velocities between the appropriate transmitters and receivers. Then the time-varying components of the channel coefficients, $\bar{h}_{ij}^{(k)}(l)$, are given by:

$$(35) \quad \bar{h}_{ij}^{(k)}(l) = J_0(2\pi f_{ij}T)h_{ij}^{(k-1)}(l) + g_{l,i,j}w(k),$$

where $w(k)$ is a zero-mean, i.i.d. complex Gaussian process with auto-correlation $R_w(j) = E\{w(k)w^*(k+j)\} = \delta(j)$, and the constant $g_{l,i,j}$ is specified by the quantity $K_{l,i,j}$ given by:

$$(36) \quad K_{l,i,j} = 10 \log \frac{|c_{ij}(l)|^2}{g_{l,i,j}^2/[1 - J_0(2\pi f_{ij}T)^2]}$$

$K_{l,i,j}$ is a physical characteristic of the channel representing the power of $c_{ij}(l)$ to the mean squared power of the random time-varying part $\bar{h}_{ij}^{(k)}(l)$.

4.2. Adaptive Equalization Algorithms. In order to apply the iterative signal recovery algorithm to a time-varying MIMO channel, some recursive tracking variants are proposed. We present two different approaches to adaptive equalization, each requiring a short block of pilot symbols at the start of each frame for initialization (semi-blind scenario). The first approach involves a modified recursive QR scheme which allows for efficient equalizer updates while enforcing finite alphabet constraints for improved signal recovery. For the purpose of comparison, we also considered a more traditional recursive least squares filter which, given training data, tracks the optimum equalizer coefficients in the least-square sense.

4.2.1. Recursive QR (RQR) Adaptive Equalizer. We now present our novel adaptive equalization algorithm— a QR-based recursive procedure incorporating the signal recovery results discussed earlier. The main goal of this scheme is to transform the observation space in such a way that updating the equalizer can be done efficiently while still forcing adaptation to changing channel characteristics. To this end, we introduce a modified recursive QR factorization of the data matrix introduced in (10), $\Gamma^{\rho,N}[x]$, which now takes the form $\Gamma^{\rho,N}[x]\Phi^N$ to allow multiplication by a forgetting factor ($\Phi^N = \text{diag}(\lambda^i)_{i=1}^N$ where $\lambda \in [0, 1]$).

Assume we have matrix R at time N such that $R \Gamma^{\rho,N}[x]\Phi^N = Q$ where R is lower-diagonal and Q is unitary. With the arrival of Δ additional symbols (or columns of $\Gamma^{\rho,N}[x]$), we wish to find \tilde{R} satisfying:

$$(37) \quad \tilde{R} \Gamma^{\rho,N+\Delta}[x]\Phi^{N+\Delta} = \tilde{Q}.$$

To expedite the computation of \tilde{R} , we compute an updating matrix C such that:

$$(38) \quad \tilde{R} = CR.$$

Note that

$$(39) \quad \tilde{Q} = \tilde{R} \Gamma^{\rho,N+\Delta}[x]\Phi^{N+\Delta} = CR \Gamma^{\rho,N+\Delta}[x]\Phi^{N+\Delta} = C \begin{bmatrix} Q_{\Delta} & \lambda^{\Delta} Q \end{bmatrix},$$

where Q_{Δ} denotes the sub-matrix formed by the Δ newly added columns to the data matrix, pre-multiplied by R . Then, by definition (\tilde{Q} is unitary):

$$(40) \quad \tilde{Q}\tilde{Q}^* = C[Q_{\Delta}Q_{\Delta}^* + \lambda^{2\Delta}I_{r(\rho+1)}]C^* = I_{r(\rho+1)}.$$

Thus, the updating matrix C can be computed via a Cholesky factorization.

With this modified recursive QR formulation, we outline the main components of our adaptive equalizer. As can be seen in Figure 3, each time a new block of data is received, the three procedures are performed: whitening, signal and equalizer tracking, and interference cancellation.

1. *Whitening:*

From a signal estimation perspective, the goal of the modified QR factorization described above is to “whiten” the data matrix, $\Gamma^{\rho,N+\Delta}[x]\Phi^{N+\Delta}$,

i.e. find \tilde{R} such that the rows of $\tilde{R}\Gamma^{\rho, N+\Delta_N}[x]\Phi^{N+\Delta_N} = \tilde{Q}$ are orthonormal. Equalizer tracking on this whitened space has empirically demonstrated more numerical stability and is quite practical given the recursive QR factorization. The whitening transformation can be recursively updated by the matrix C . According to (40), the updating matrix C can be computed via the following Cholesky factorization:

$$(41) \quad Q_{\Delta}Q_{\Delta}^* + \lambda^{2\Delta}I_{r(\rho+1)} = C^{-1}C^{*-1}.$$

2. *Signal and equalizer tracking:*

Letting $\Upsilon^{\rho}[\hat{g}_i^{(k)}] = \Upsilon^{\rho}[g_i^{(k)}]R^{-1}$ denote the transformed Bézout equalizer operating on the whitened data space at time k , we obtain an estimate of the block of the i -th user's newly arrived symbols, $\hat{s}_i(k)$, as follows:

$$[\hat{s}_i(N+\Delta-1) \ \dots \ \hat{s}_i(N)] = T[\Upsilon^{\rho}[\hat{g}_i^{(N-1)}]Q_{\Delta}]$$

where $\Upsilon^{\rho}[\hat{g}_i^{(N-1)}]$ is the equalizer updated during the previous block and $T[\cdot]$ is a transformation taking each estimate to the nearest valid symbol (finite alphabet constraint). For clarity, let $\hat{s}_i^{new} = [\hat{s}_i(N+\Delta-1) \ \dots \ \hat{s}_i(N)]$ and $\hat{s}_i^{old} = [\hat{s}_i(N-1) \ \dots \ \hat{s}_i(0)]$. After obtaining the signal estimate, we then update the equalizer using:

$$(42) \quad \Upsilon^{\rho}[\hat{g}_i^{(N)}] = [\hat{s}_i^{new} \ | \ \hat{s}_i^{old}] \tilde{Q}^*.$$

At first glance, (42) seems to suggest that we must keep track of both \hat{s}_i^{old} and \tilde{Q} for equalizer updating. However, neither is necessary, and in fact all we need to update $\Upsilon^{\rho}[\hat{g}_i^{(N)}]$ are $\Upsilon^{\rho}[\hat{g}_i^{(N-1)}]$, R , and the newly received data. Note that:

$$(43) \quad \tilde{Q} = \tilde{R}\Gamma^{\rho, N+\Delta}[x]\Phi^{N+\Delta} = CR[\Gamma^{\rho, \Delta}[x] \ | \ \Gamma^{\rho, N}[x]]\Phi^{N+\Delta} = [CQ_{\Delta} \ | \ \alpha CQ]$$

where $\alpha = \lambda^{\Delta}$. Substituting (43) into (42), we have:

$$(44) \quad \Upsilon^{\rho}[\hat{g}_i^{(N)}] = \{\hat{s}_i^{new}Q_{\Delta}^* + \alpha\hat{s}_i^{old}Q^*\}C^* = \{\hat{s}_i^{new}Q_{\Delta}^* + \alpha\Upsilon^{\rho}[\hat{g}_i^{(N-1)}]\}C^*$$

Thus, our equalizer update procedure takes on a relatively simple and efficient form

3. *Interference cancellation:*

If we assume the source signals are white and uncorrelated, then the channel parameters corresponding to the extracted source signal (i th user) can be estimated as follows:

$$\hat{h}_i = \mathcal{E}\left\{\frac{\Gamma^{L, N}[x]\hat{s}_i^*}{\|\hat{s}_i\|^2}\right\}$$

where L is the ISI length, \hat{s}_i is a $N \times 1$ vector of users i 's most recently recovered symbols, and N is sufficiently long to justify the whiteness and correlation assumptions. Once the channel is estimated, then the interference caused by that user can be accordingly cancelled as illustrated in Figure 3.

4.2.2. Decision-Directed Recursive Least Squares (RLS) Adaptive Equalizer. For the purpose of comparison, we now introduce a traditional least squares equalizer which is largely based on work presented in [24]. The structure of the equalizer is identical to that shown in Figure 1, with a bank of FIR filters whose outputs are summed to obtain a signal estimate. Let $\underline{g}_i^{(k)}$ denote the vector of equalizer coefficients for recovering the i th user at time index k arranged in space major order. We define the state space model as

$$\underline{g}_i^{(k+1)} = \lambda_k^{-1/2} \underline{g}_i^{(k)}, \quad \hat{s}_i(k) = \underline{g}_i^{(k)} X(k),$$

where $\hat{s}_i(k)$ is the estimate of the i th user's signal at time k and $X(k) = [x_1(k) \dots x_m(k), \dots, [x_1[k-\rho] \dots x_m[k-\rho]]^T]$, which is simply an appropriately chosen column of $P^{p,N}[x]$. The least squares solution can be formulated recursively and results in the following update equations [24]:

$$(45) \quad \begin{aligned} \underline{g}_i^{(k+1)} &= \underline{g}_i^{(k)} + K_p \left(s_i(k) - \underline{g}_i^{(k)} X(k) \right) \\ K_p &= (R^{-1}(k) X^*(k) P(k-1)) / \lambda_k, \quad R(k) = (X^*(k) P(k) X(k)) / \lambda_k + 1 \\ P(k+1) &= (P(k) - P(k) X(k) K_p) / \lambda_k \end{aligned}$$

where the initial states $P(0)$ and $\underline{g}_i^{(0)}$ can be arbitrarily chosen. After the initial block of pilot symbols, the decision-directed RLS equalizer coefficients are adjusted by mapping the filter output to the nearest valid symbols in the signal constellation and assuming the resulting difference is the estimation error. We make two modifications to the traditional RLS algorithm to improve its performance for time-varying channels. The first is to limit how often the equalizer is updated. The updates following the training symbol phase are only performed after the accumulated error (summed norm of the difference between outputs and valid symbols) exceeds a threshold, which is empirically chosen through several simulations. Furthermore, when updates are performed, λ_k is tuned optimally such that the equalization error is an uncorrelated noise sequence, as is described in [25].

4.2.3. Comparison of RQR and RLS algorithms. The formulation and computation needs for the RQR (Recursive QR) approach can be directly compared to those of the RLS algorithm. Note that the RQR equalizer updating effort is proportional to Δ (block size of newly arriving data). Of particular interest is when $\Delta = 1$, the case when the equalizer is updated upon arrival of each new observation.

For the RQR algorithm, the equalizer update equation (a simplification of (44))

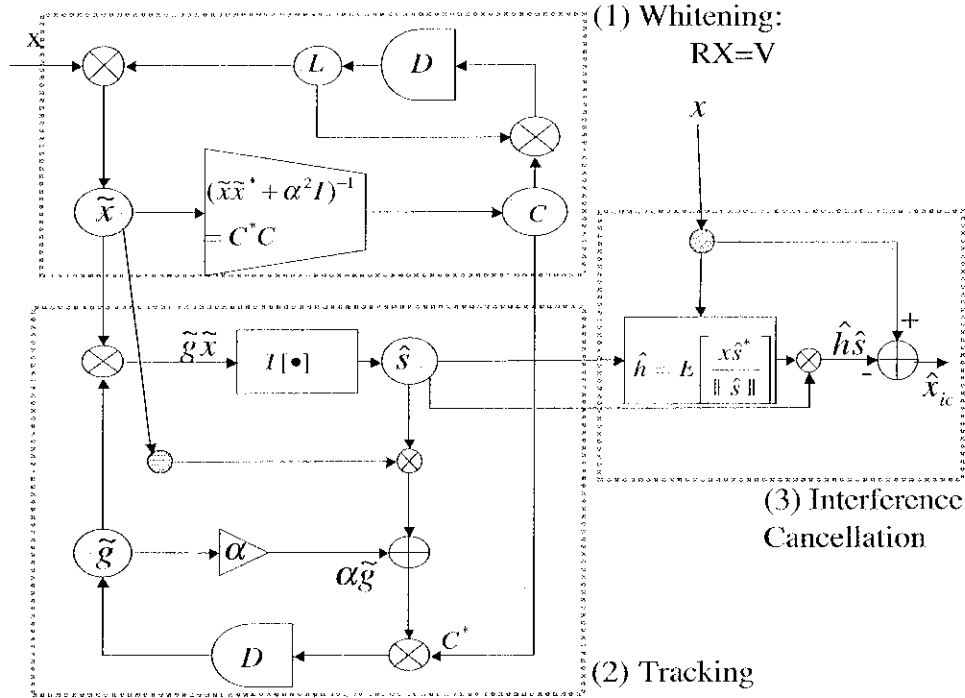


FIG. 3. Three subsystems of Recursive QR tracking system: (1) Prewhitening; (2) equalizer tracking and signal detection; and (3) decision-feedback for interference cancellation. For simplicity and without loss of generality, the flat MIMO case is illustrated here.

is:

$$(46) \quad \Upsilon^\rho[\hat{g}_i^{(k)}] = \alpha \Upsilon^\rho[\hat{g}_i^{(k-1)}]C^{*+} + \hat{s}_i(k)Q_\Delta^*C^{*+}$$

In comparison, when the RLS equalizer is operating in the post pilot symbol phase (estimation error is the difference between output and nearest the valid symbol), a minor rearrangement of the equalizer update equation in (45) yields:

$$(47) \quad \underline{g}_i^{(k)} = \lambda_k^{-1/2} \underline{g}_i^{(k-1)} \{I - X(k)K_p\} + \hat{s}_i(k)\lambda_k^{-1/2}K_p$$

The similarity in the two update equations suggests an interesting parallel between the two algorithms. We claim, and provide supporting simulations in the following section, that the RQR algorithm provides better equalizer-tracking capability and increased robustness against noise while incurring little extra computation. We should also note that throughout the development of our algorithm we assume that earlier detected symbols are correct. This suggests a possible susceptibility to error propagation, which we will address in a future paper.

5. Simulation. We present several simulation results comparing the proposed RQR scheme and the more traditional RLS adaptive equalizer as well as demonstrate the performance of our method for a variety of system parameters. Figure 4 shows symbol error rate results for the RLS and RQR adaptive equalizers for three different degrees of channel stationarity. Similar plots comparing the RQR approach to an MMSE equalizer with channel knowledge are given in Figure 5. In both, we specify channel stationarity by the Doppler rate, $f_d T$ (Doppler frequency multiplied by the symbol period)³. For comparison, we plot the performance of an MMSE equalizer for each of the channels with varying amounts of channel knowledge imprecision. We introduce channel imprecision by adding a Gaussian random variable to each channel coefficient whose standard deviation is the specified percentage of the corresponding coefficient's instantaneous amplitude, i.e. $h_{ij}^{(k)}(l) + \nu$ where $\nu \in \mathcal{N}(0, \|h_{ij}^{(k)}(l)\| * (\textit{imprecision ratio}))$. Each of the three plots corresponds to a (2,5) MIMO system with an ISI length of 4, and an equalizer order of 5. Each point represents an average over 600 simulations of 1150 symbols, each with randomly generated initial channels. Both the RLS and RQR algorithms are given the first 150 of these symbols for training purposes and perform unsupervised equalizer-tracking on the remaining 1000 symbols.

The proposed RQR scheme clearly outperforms the traditional RLS method for all degrees of channel stationarity, particularly in the low SNR regime. Furthermore, for the most stationary channel ($f_d T = .005$) the RQR equalizer performs better than the MMSE equalizer given channel knowledge with (somewhat high) imprecision at 20%. For less stationary channels, both the RLS and RQR methods are decisively inferior to the MMSE equalizer with exact channel knowledge, but the RQR scheme seems to offer comparable performance to that of the MMSE equalizer with slightly more imprecision ($\approx 20 - 30\%$). Also, note the performance improvement offered by interference cancellation. For all degrees of stationarity, interference cancellation in these two-user cases leads to improvements of approximately 1-2 dB. For the two most stationary channels (plots a.) and b.)), this interference cancellation approach rivals the MMSE equalizer with exact channel knowledge.

Figure 6 illustrates the performance of the RQR equalizer for various choices of equalizer order and system delay parameters. In both simulations, the ISI length, SNR, and channel stationarity were held constant. Figure 6a. shows the SER performance of the RQR equalizer for different equalizer orders. The theory presented in Section 3.3.1 predicts that postprocessing SNR can only increase with the filter order. These results appear to support this claim as the SER improves exponentially with an increase in equalizer order. We also predict that finding the optimal recovery filter

³For reference, a system with a 2.4 GHz carrier frequency, 20 kbps rate, and a mobile station traveling at 60 mph corresponds to a Doppler rate of approximately .01

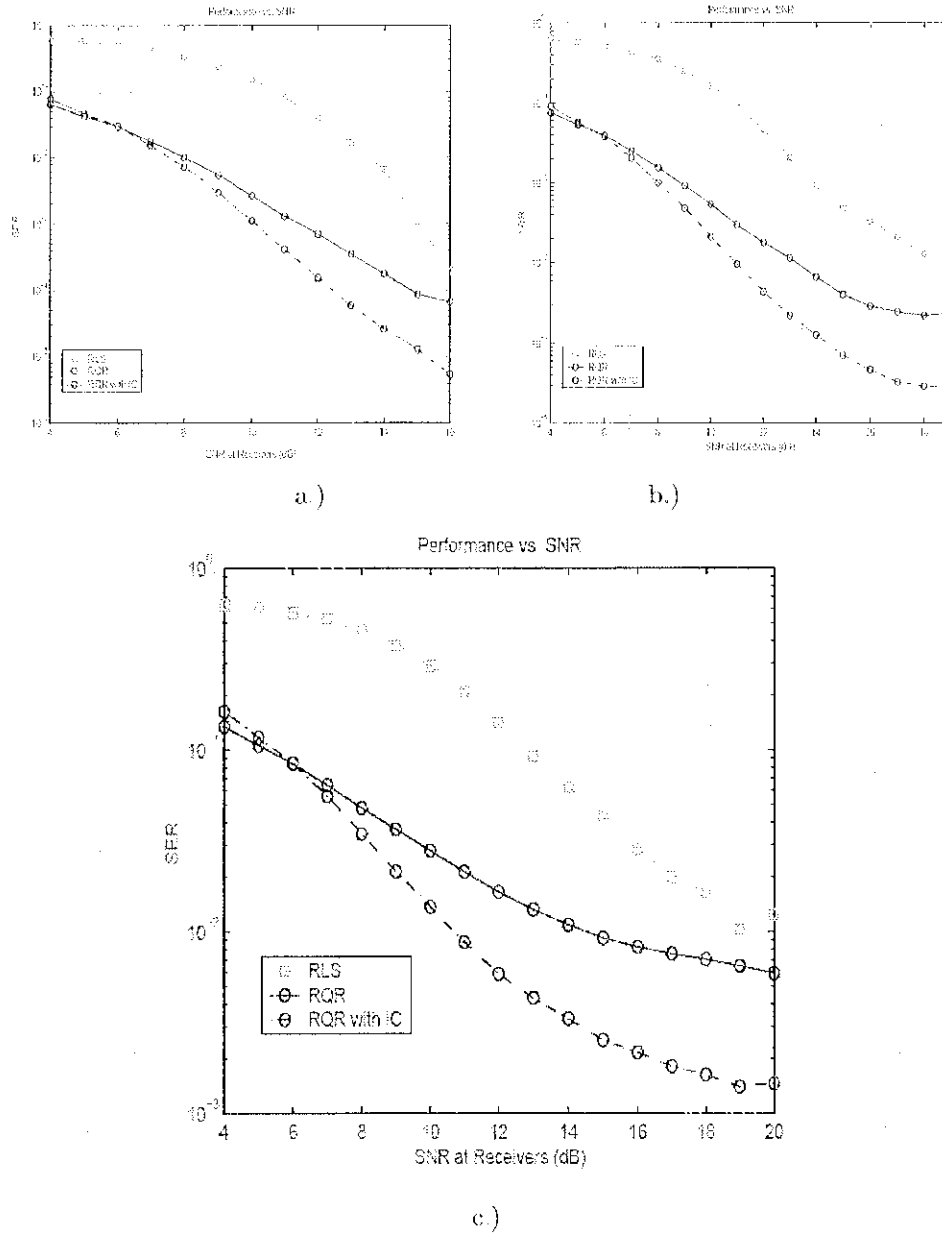
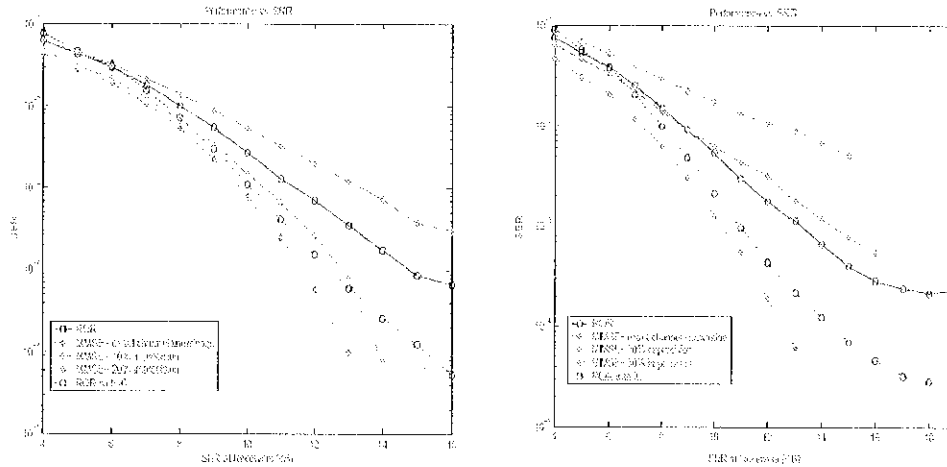


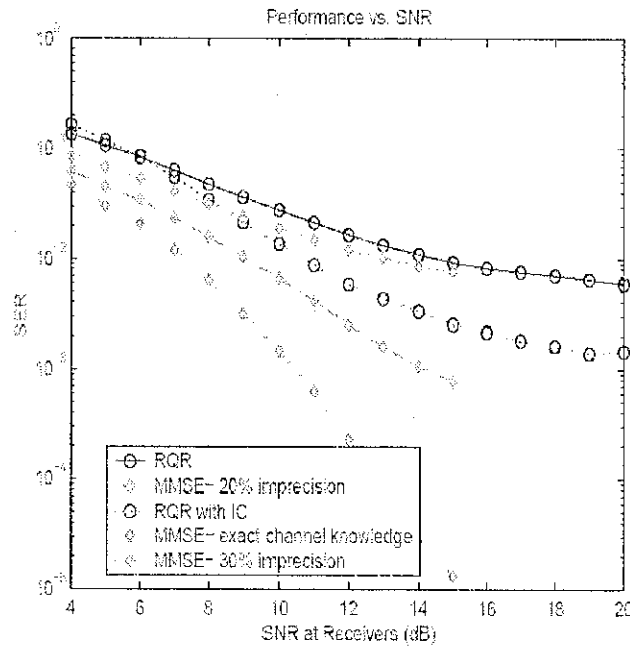
FIG. 4. RLS and RQR Comparison: Symbol Error Rate vs SNR (QPSK, 1150 symbols, 2 TX, 5 RX, ISI length: 4, Equalizer order: 5, Delay: 5, Δ : 1, Λ : 1, $K_{l,i,j}$: 10 dB $\forall l, i, j$) a.) $f_D T = .005$, b.) $f_D T = .01$, c.) $f_D T = .03$

becomes more difficult with larger equalizer order due to the increased search space size. Although we cannot see this trend here, caution must be exercised in selecting equalizer order. In fact, a blind version of the algorithm was presented in [22], where it was reported that excessive equalizer order can indeed yield poor performance



a.)

b.)



c.)

FIG. 5 RQR and MMSE Comparison: Symbol Error Rate vs SNR (QPSK, 1150 symbols, 2 TX, 5 RX, ISI length: 4, Equalizer order: 5, Delay: 5, Δ : 1, Λ : 1, $K_{l,i,j}$: 10 dB $\forall l, i, j$) a.) $f_D T = .005$, b.) $f_D T = .01$, c.) $f_D T = .03$.

Aside from this concern, it must be noted that the complexity of the RQR algorithm depends critically on the chosen filter order, and thus, we suggest this also be a major consideration in order selection

Figure 6b. shows the simulation results for a fixed equalizer order (5) and varying

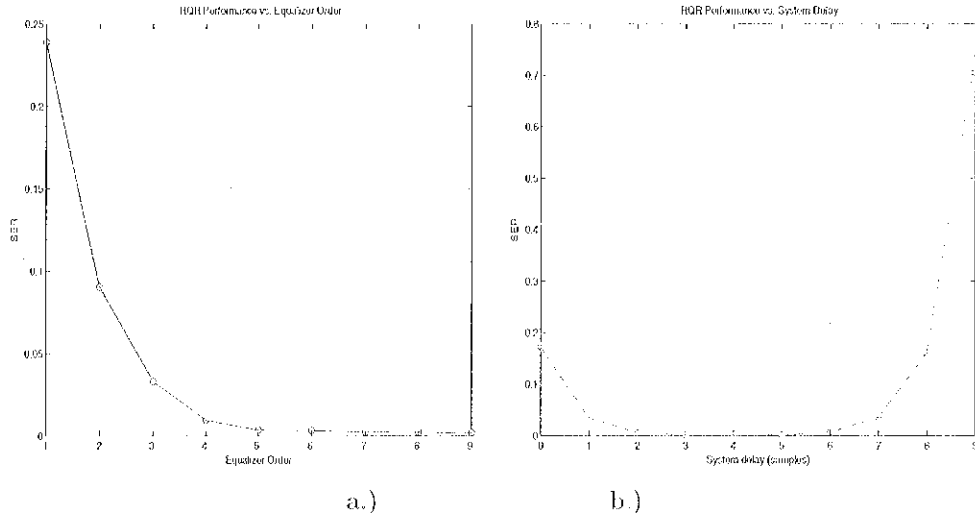


FIG. 6. a) Symbol Error Rate vs. Equalizer order (QPSK, 1150 symbols, 2 TX, 5 RX, ISI length: 4, Delay: 5, SNR: 12 dB, Δ : 1, $f_D T$: 0.1, $K_{l,i,j}$: 10 dB $\forall l, i, j$) b) Symbol Error Rate vs. System Delay (QPSK, 1150 symbols, 2 TX, 5 RX, ISI length: 4, Equalizer order: 5, $f_D T$: 0.1, $K_{l,i,j}$: 10 dB $\forall l, i, j$)

system delays. The theory presented in Section 3.3.1 also suggests that a minimum system delay must be allowed for correct signal recovery. These results support this claim as we only see acceptable SER's for delays of 3,4,5 with a minimum occurring at 4. Not surprisingly, too large of a delay also hinders signal recovery. Through several simulations, we empirically find that the optimum delay is approximately $\lfloor (L-1-\rho)/2 \rfloor$.

Finally, Figure 7 demonstrates the capability of the proposed RQR algorithm for different transmitter and receiver scenarios. The plotted data reflects an average over 100 simulations for which the ISI length, equalizer order, delay, SNR, and channel stationarity were held constant, while the number of transmitters and receivers was adjusted. As expected, for each transmitter scenario, the added diversity offered by extra receivers improves the equalizer's performance. It is also not surprising that as the number of inputs to be recovered increases, the SER performance suffers.

An important question to be addressed is the computational practicality of the equalization approach presented. Clearly, the component of the algorithm incurring the most complexity is the Cholesky factorization required for equalizer updating. Traditional Cholesky factorization is an $O(n^3)$ algorithm, but we can save computation by noting the special structure of the matrix we wish to factor. Consider the case when $\Delta = 1$ (equalizer is updated upon arrival of each new observation). Then, Q_Δ of (41) is simply a column vector ($Q_\Delta Q_\Delta^*$ has rank 1) and the Cholesky factorization of $[Q_\Delta Q_\Delta^* + \lambda^{2\Delta} I_{r(\rho+1)}]$ can be obtained by a simple update of the trivial factorization of $\lambda^{2\Delta} I_{r(\rho+1)}$. This simpler updating has a complexity of only $O(n^2)$, which is

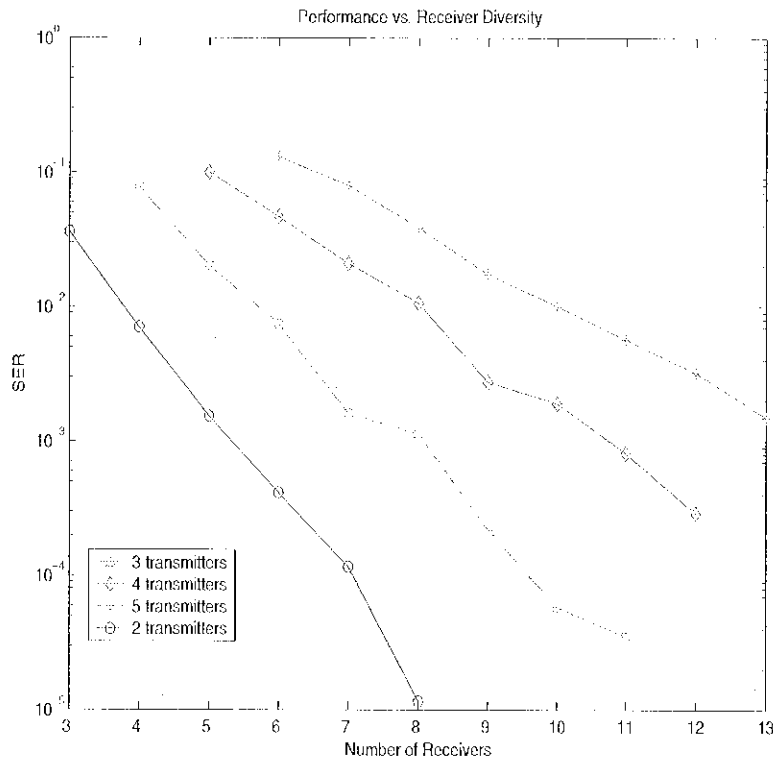


FIG. 7 Symbol Error Rate vs. Number of Receivers (QPSK, 1150 symbols, ISI length: 4, Delay: 5, SNR: 12 dB, Δ : 1, $f_D T$: 0.1, $K_{L,i,j}$: 10 dB $\forall i, j$)

comparable to the computational complexity of the RLS adaptive equalizer

6. Conclusion. This paper presents a recursive QR approach to semi-blind equalization of time-varying ISI MIMO channels. The theoretical foundations of the proposed approach are rooted in signal recovery results derived from the generalized Bézout identity and the finite alphabet property inherent in digital communication schemes. Concerning the behavior of the algorithm, three issues of main concern are addressed: *existence* and *exclusiveness* of fixed points, and *robustness*. It is recognized that it is necessary to impose a proper equalizer order and system delay for correct and robust results. Theoretical and practical bounds for such parameters are provided. Under the theoretical framework established in the paper, we develop a computationally efficient recursive QR scheme for adaptive equalization of time-varying MIMO systems and present simulation results confirming its performance.

REFERENCES

- [1] S. K. C. MYERS AND X. ZHANG, *A recursive QR approach to semi-blind equalization of time-varying MIMO channels*, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP), Montreal, Canada, May 2004.
- [2] G. CHANNAKIS, Y. HUA, P. STOICA, AND L. TONG, *Signal Processing Advances in Wireless and Mobile Communications*, ser. Trends in Channel Estimation and Equalization. Upper Saddle River, NJ: Prentice Hall, 2001, No. 1.
- [3] C. KOMNINAKIS, C. FRAGOUI, A. SAYED, AND R. WESEI, *Adaptive multi-input multi-output fading channel equalization using Kalman Estimation*, in: IEEE International Conference on Communications (ICC 2000), Vol. 3, New Orleans, Louisiana, USA, June 2000, pp. 1655-1659.
- [4] N. AL-DHAHR AND A. SAYED, *The finite-length multi-input multi-output MMSE-DFE*, IEEE Transactions on Signal Processing, 48:10(2000), pp. 2921-2936.
- [5] A. TEHRANI, *Space-time equalization in wireless systems with multiple transmit/receive array antennas*, Ph.D. dissertation, Stanford University, Stanford, CA, 2000.
- [6] E. MOULINES, P. DUHAMEL, J. F. CARDOSO, AND S. MAYRARGUE, *Subspace methods for the blind identification of multichannel FIR filters*, IEEE Transactions on Signal Processing, 43:2(1995), pp. 516-525.
- [7] L. TONG, G. XU, AND T. KAILATH, *A new approach to blind identification and equalization of multipath channels*, in: Proceedings Of the 25th Asilomar Conference on Signals, Systems and Computers, Vol. 2, Pacific Grove, CA, USA, November 1991, pp. 856-860.
- [8] G. H. XU AND H. LIU AND L. TONG AND T. KAILATH, *A least-squares approach to blind channel identification*, IEEE Transactions on Signal Processing, 43:5(1995), pp. 2982-2993.
- [9] K. ABED-MERAIM, J. F. CARDOSO, A. Y. GOROKHOV, P. LOUBATON, AND E. MOULINES, *On subspace methods for blind identification of single-input multiple-output fir systems*, IEEE Transactions on Signal Processing, 45:1(1997), pp. 42-55.
- [10] L. TONG AND S. PERREAU, *Multichannel blind identification: From subspace to maximum likelihood methods*, Proceedings of IEEE, 86:10(1998), pp. 1951-1968.
- [11] V. U. REDDY, C. B. PAPADIAS, AND A. J. PAULRAJ, *Blind identifiability of certain classes of multipath channels from second-order statistics using antenna arrays*, IEEE Signal Processing Letters, 4:5(1997), pp. 138-141.
- [12] H. ZENG AND L. TANG, *Blind channel estimation using the second-order statistics: Algorithms*, IEEE Transactions on Signal Processing, 45:8(1997), pp. 1919-1930.
- [13] R. W. LUCKY, *Techniques for adaptive equalization of digital communication systems*, Bell Lab Technical Journal, 45(1966), pp. 255-286.
- [14] C. R. JOHNSON, P. SCHULTER JR., I. FIDALKOW, L. TONG, J. D. BEHM, M. G. LARIMORE, D. R. BROWN, R. A. CASAS, T. J. FENDRES, S. LAMBOHARAN, H. H. ZENG, A. TOUZNI, M. GREEN, AND J. R. TREICHLER, *The Core of FSI-CMA Behavior Theory*, 1st ed., ser. Unsupervised Adaptive Filtering II: Blind Deconvolution. New York, NY: John Wiley & Sons, 2000.
- [15] A. LESHEM, N. PETROCHLOS, AND A. J. VAN DER VEEN, *Finite sample identifiability of multiple constant modulus sources*, IEEE Trans. on Information Theory, 49:9(2003), pp. 2314-2319.
- [16] S. TALWAR, M. VIBERG, AND A. J. PAULRAJ, *Blind separation of synchronous co-channel digital signals using an antenna array. part i. Algorithms*, IEEE Trans. Signal Processing, 44:5(1996), pp. 1184-1197.
- [17] A. J. VAN DER VEEN, S. TALWAR, AND A. J. PAULRAJ, *A subspace approach to blind space-time signal processing for wireless communication systems*, IEEE Trans. Signal Processing, 45:1(1997), pp. 173-190.

- [18] G. A. JONES AND J. M. JONES, *Elementary Number Theory*, Berlin, Germany: Springer-Verlag, 1998.
- [19] S. Y. KUNG, *Multivariable and multidimensional systems: Analysis and design*, Ph.D. dissertation, Stanford University, Stanford, CA, 1977.
- [20] T. KAILATH, *Linear Systems*, Englewood Cliffs, NJ: Prentice Hall, 1980.
- [21] S. Y. KUNG, T. KAILATH, AND M. MÖRF, *A generalized resultant matrix for matrix polynomials*, in: Proceedings of the Conference on Decisions and Control, Clearwater Beach, Florida, USA, December 1976, pp. 892-895.
- [22] S. Y. KUNG AND X. ZHANG, *An associative memory approach to blind signal recovery for SIMO/MIMO systems*, in: IEEE Workshop on Neural Network for Signal Processing (NNSP), Falmouth, Massachusetts, USA, September 2001, pp. 1073-1076.
- [23] S. Y. KUNG, Y. WU, AND X. ZHANG, *Bezout space-time precoding and equalization for MIMO channels*, IEEE Transactions on Signal Processing, 50:10(2002), pp. 2499-2514.
- [24] A. SAYED AND T. KAILATH, *A state-space approach to adaptive RLS filtering*, IEEE Signal Processing Magazine, 11:3(1994), pp. 18-60.
- [25] W. ZHUANG, *RLS algorithm with variable forgetting factor for decision feedback equalizer over time-variant fading channels*, Wireless Personal Communications, 8:1(1998), pp. 15-29.

Q
1
2

Y
3
4
5