ADAPTIVE MIMO OFDM RECEIVERS: IMPLEMENTATION IMPAIRMENTS AND COMPLEXITY ISSUES *

Alireza Tarighat, Waleed M. Younis, and Ali H. Sayed

Department of Electrical Engineering
University of California
Los Angeles, CA 90095
Email: {tarighat,waleed,sayed}@ee.ucla.edu

Abstract: MIMO OFDM communications is a promising choice for future high data rate wireless systems. However, the commercial deployment of MIMO OFDM systems faces some challenges, including those arising from front-end analog implementation impairments and complex receiver structures. The impairments are caused by the analog processing of the received radio frequency (RF) signal and they cannot be efficiently nor entirely eliminated in the analog domain. This paper illustrates the effect of In-phase and Quadrature-phase (IQ) imbalances on system performance. Receiver algorithms are described that compensate for IQ imbalances. The structure of space-time block codes, along with the assumed structure of the distortion models, are exploited to design reduced-complexity adaptive receivers that are robust to IQ imbalances.

Keywords: Orthogonal frequency division multiplexing (OFDM), multi-input-multi-output (MIMO) systems, in-phase and quadrature-phase (IQ) imbalances, space-time coding, Alamouti scheme, equalization, adaptive receiver.

1. INTRODUCTION

High speed communications over broadband wireless channels has emerged as a key feature of future communications systems due in part to the explosive interest in information technology applications, including wireless networks, mobile computing, high-speed mobile internet, and video transmission over wireless channels. The demand for higher information capacity in these and other similar applications has motivated the use of broadband wireless channels in order to provide wider bandwidth and higher data rates. In addition, multi-user communication schemes are also being employed in order to allow users to share the same physical channel; thereby contributing to even higher data rates.

Broadband multi-antenna OFDM communications has emerged as a leading technology for such high speed wireless networks. OFDM-based physical layers have already been chosen for several systems such as IEEE 802.11a, IEEE 802.11g, IEEE 802.11n, IEEE 802.20, and IEEE 802.16. Still, there are challenges for the wide deployment of these systems. One such challenge is due to the implementation impairments that arise from analog component imperfections. Such impairments are difficult to eliminate using analog processing, and they are even more challenging at higher carrier frequencies and for higher bandwidths. Another challenge is due to the complexity of the receivers that are required for multi-antenna systems.

* This work was partially supported by NSF grants CCR-0208573 and ECS-0401188. ***Proc. IFAC workshop on Adaptation and Learning in Control and Signal Processing*, Yokohama, Japan, Aug. 2004.**

The structure of the underlying codes, as well as any prior knowledge about the distortion and imperfection models, should be exploited in order to reduce computational complexity.

1.1 Analog Impairments

With regards to the impairments in the analog components, they are mainly due to fabrication process variations which are not predictable nor controllable and tend to increase as fabrication technologies scale down (Pelgrom *et al.*, 1989). One major source of impairments is the imbalance that occurs between the In-phase (I) and Quadrature-phase (Q) branches; or equivalently, the real and imaginary parts of the complex signal (Razavi, 1998)-(Pengfei *et al.*, 2003). Usually, the received signal is down converted from the radio frequency (RF) to the baseband signal before it can be processed in the digital domain. A complex down-converter basically multiplies the RF signal by a complex waveform, $e^{-j2\pi f_{\mathrm{LO}}}$, and the spectrum of the received signal is shifted to the baseband by $2\pi f_{\mathrm{LO}}$. To perform the complex down conversion, both the *sine* and *cosine* oscillating waveforms are required at the receiver. The IQ imbalance is basically any mismatch between the I and Q branches from the ideal case, i.e., from the exact 90^o phase difference and equal amplitudes. The performance of OFDM receivers can be hindered by such IQ imbalances (see, e.g., (Baier, 1990)-(Tarighat and Sayed, 2004*b*)).

1.2 Receiver Complexity

Besides analog impairment issues, the complexity of OFDM receivers for multi-antenna (MIMO) systems operating in multi-user environments is also a concern as a result of the need to suppress inter-user interference. In addition, the receivers need to deal with distortions introduced by the MIMO channels and compensate for inter-symbol interference. Space-time codes are particularly useful in this regard since they simplify the structure of OFDM transceivers. For instance, they do not require channel state information at the transmitter. Moreover, as we are going to discuss further ahead, the rich structure of the codes should and can be exploited in order to simplify the overall computational complexity.

1.3 Objective

The purpose of this article is to overview some recent contributions in the area of adaptive MIMO OFDM receivers based on the developments in (Tarighat and Sayed, 2004*a*)-(Younis and Sayed, 2004*b*). In addition, some new results pertaining to the receiver structure are derived. Specifically, it will be shown that while IQ imbalances destroy some of the properties of space-time block codes, efficient receive algorithms can still be developed. The proposed receiver algorithms perform both distortion compensation and data decoding in a joint manner. Moreover, efficient adaptive receivers will be described with fast tracking/convergence abilities for joint distortion compensation and decoding for both single-user and multi-user environments with space-time block coded (STBC) transmissions. Frequency–selective channels will be assumed with two antennas per user for transmission and one antenna per user for reception. We shall indicate how the special structure of the space–time block code can be exploited to reduce complexity.

The paper is organized as follows. The next section reviews the model used for IQ imbalances in (Tarighat and Sayed, 2004*a*) and formulates the effect of IQ imbalances on SISO OFDM receivers. The receiver of (Tarighat and Sayed, 2004*b*) for Alamouti coded OFDM systems with compensation for IQ imbalances is then described in Sec. 3. In addition, an adaptive implementation of the receiver algorithm is presented in this section. The results are then extended to the multi-user scenario in Sec. 4. These results are variations of earlier adaptive solutions presented in (Younis and Sayed, 2004*a*)-(Younis and Sayed, 2004*b*). In comparison to the RLS implementations described in these references, the data in this article are first pre-processed by a Cholesky factor with the purpose of reducing the complexity of the RLS recursions to that of LMS-type recursions. Simulation results are presented in Sec. 5. Conclusions are given in Sec. 6.

2. IQ IMBALANCES IN SISO OFDM SYSTEMS

Following (Tarighat and Sayed, 2004*a*), let $b(t)$ represent the received complex signal before being distorted by the IQ imbalance caused by the analog signal processing. The distorted signal in the time domain can be written as (Baier, 1990),(Liu, 1998):

$$b'(t) = \mu b(t) + \nu b^*(t) \tag{1}$$

where the distortion parameters, μ and ν, are related to the amplitude (α) and phase (θ) imbalances between the I and Q branches. A simplified model for this distortion is given by (Liu, 1998):

$$\begin{aligned} \mu &= \cos(\theta/2) + j\alpha \sin(\theta/2) \\ \nu &= \alpha \cos(\theta/2) - j \sin(\theta/2) \end{aligned} \tag{2}$$

The values of θ and α are not known at the receiver since they are caused by manufacturing inaccuracies in the analog components.

Now, in OFDM systems, a block of data of size N (where N is a power of 2) is transmitted as an OFDM symbol, say

$$\mathbf{s} \triangleq \mathrm{col}\{\mathbf{s}(1), \mathbf{s}(2), \ldots, \mathbf{s}(N)\} \tag{3}$$

Each block is first passed through the IDFT operation:

$$\bar{\mathbf{s}} = \mathbf{F}^*\mathbf{s} \tag{4}$$

where $\mathbf{F}$ is the unitary discrete Fourier transform (DFT) matrix, and a cyclic prefix of length P is added to each transformed block prior to transmission through the channel–see Figure 1. An FIR model with $L+1$ taps is assumed for the channel, i.e.,

$$\mathbf{h} = \text{col}\{h_0, h_1, \ldots,\ h_L\} \tag{5}$$

with $L \leq P$. At the receiver, the incoming samples corresponding to the transmitted block $\bar{\mathbf{s}}$ are collected into a vector, after discarding the received cyclic prefix samples. The received block of data before being distorted by IQ imbalances can be written as (e.g., (Tarighat and Sayed, 2003)):

$$\bar{\mathbf{y}} = \mathbf{H}^c\bar{\mathbf{s}} + \bar{\mathbf{v}} \tag{6}$$

where

$$\mathbf{H}^c = \begin{bmatrix} h_0 & h_1 & \cdots & h_L & & & & \\ & h_0 & h_1 & \cdots & h_L & & & \\ & & \ddots & & & \ddots & & \\ & & & h_0 & h_1 & \cdots & h_L \\ \vdots & & & & \ddots & & \vdots \\ h_2 & \cdots h_L & & & & h_0 & h_1 \\ h_1 & \cdots & h_L & & & & h_0 \end{bmatrix} \tag{7}$$

is an $N \times N$ circulant matrix that can be diagonalized by the DFT matrix. Specifically, $\mathbf{H}^c = \mathbf{F}^*\mathbf{\Lambda}\mathbf{F}$ where

$$\mathbf{\Lambda} = \text{diag}\{\lambda\} \tag{8}$$

and the vector λ is related to $\mathbf{h}$ via

$$\lambda = \mathbf{F}^* \begin{bmatrix} \mathbf{h} \\ \mathbf{0}_{(N-(L+1))\times 1} \end{bmatrix} \tag{9}$$

Then (6) gives

$$\bar{\mathbf{y}} = \mathbf{F}^*\mathbf{\Lambda}\mathbf{F}\bar{\mathbf{s}} + \bar{\mathbf{v}} = \mathbf{F}^*\text{diag}\{\lambda\}\,\mathbf{F}\bar{\mathbf{s}} + \bar{\mathbf{v}} \tag{10}$$

The received block of data $\bar{\mathbf{y}}$ after being distorted by IQ imbalances will be transformed to (Tarighat and Sayed, 2004*a*; Tarighat and Sayed, 2004*b*):

$$\bar{\mathbf{z}} = \mu\bar{\mathbf{y}} + \nu\text{conj}(\bar{\mathbf{y}}) \tag{11}$$

where the notation $\text{conj}(\bar{\mathbf{y}})$ denotes a column vector whose entries are the complex conjugates of the entries of $\bar{\mathbf{y}}$. Now remember that the N-point DFT of the complex conjugate of a sequence is related to the DFT of the original sequence through a mirrored relation (assuming $1 \leq n \leq N$ and $1 \leq k \leq N$):

$$\begin{aligned} x(n) &\xrightarrow{\text{DFT}} X(k) \\ x^*(n) &\xrightarrow{\text{DFT}} X^*(N-k+2) \end{aligned} \tag{12}$$

For notational simplicity, we denote the operation which gives the DFT of the complex conjugate of a vector by the superscript $\#$, i.e., for a vector X of size N we write

$$X = \begin{bmatrix} X(1) \\ X(2) \\ \vdots \\ X(N/2) \\ X(N/2+1) \\ X(N/2+2) \\ \vdots \\ X(N) \end{bmatrix} \Longrightarrow X^{\#} = \begin{bmatrix} X^*(1) \\ X^*(N) \\ \vdots \\ X^*(N/2+2) \\ X^*(N/2+1) \\ X^*(N/2) \\ \vdots \\ X^*(2) \end{bmatrix} \tag{13}$$

so that if

$$X = \mathbf{F}x \quad \text{then} \quad X^{\#} = \mathbf{F}\text{conj}\,(x) \tag{14}$$

Now, proceedings as in (Tarighat and Sayed, 2004*a*; Tarighat and Sayed, 2004*b*), equation (6) gives

$$\text{conj}(\bar{\mathbf{y}}) = \text{conj}(\mathbf{H}^c)\text{conj}(\bar{\mathbf{s}}) + \text{conj}(\bar{\mathbf{v}}) \tag{15}$$

where $\text{conj}(\mathbf{H}^c)$ is a circulant matrix defined in terms of $\text{conj}(\mathbf{h})$ as in (7). In a manner similar to (8), (9), and (14), we have

$$\mathbf{F}^* \begin{bmatrix} \text{conj}(\mathbf{h}) \\ \mathbf{0}_{(N-(L+1))\times 1} \end{bmatrix} = \lambda^{\#} \tag{16}$$

and

$$\text{conj}(\mathbf{H}^c) = \mathbf{F}^*\text{diag}\left\{\lambda^{\#}\right\}\mathbf{F} \tag{17}$$

Substituting the above into (15) results in

$$\begin{aligned} \text{conj}(\bar{\mathbf{y}}) &= \mathbf{F}^*\text{diag}\left\{\lambda^{\#}\right\}\mathbf{F}\text{conj}(\bar{\mathbf{s}}) + \text{conj}(\bar{\mathbf{v}}) \\ &= \mathbf{F}^*\text{diag}\left\{\lambda^{\#}\right\}\mathbf{s}^{\#} + \text{conj}(\bar{\mathbf{v}}) \end{aligned} \tag{18}$$

where $\mathbf{F}\text{conj}(\bar{\mathbf{s}})$ is replaced by $\mathbf{s}^{\#}$ using (4) and the conjugate-mirrored notation defined by (12).

Now a standard OFDM receiver would apply the DFT operation on the received block of data $\bar{\mathbf{z}}$ directly. Thus applying the DFT matrix to (11), i.e., setting $\mathbf{z} = \mathbf{F}\bar{\mathbf{z}}$, and substituting (10) and (18) into (11), lead to

$$\mathbf{z} = \mu\text{diag}\{\lambda\}\,\mathbf{s} + \nu\text{diag}\left\{\lambda^{\#}\right\}\mathbf{s}^{\#} + \mathbf{v} \tag{19}$$

where $\mathbf{v}$ is a transformed version of the original noise vector $\bar{\mathbf{v}}$. As seen from (19), the vector $\mathbf{z}$ is not related to the transmitted block $\mathbf{s}$ through a diagonal matrix, as would have occurred in an OFDM system with ideal I and Q branches. The implication of this observation is the following. Discarding the samples corresponding to tones 1 and $N/2+1$, i.e., $\mathbf{z}(1)$ and $\mathbf{z}(N/2+1)$, and defining two new vectors (Tarighat and Sayed, 2004*a*):

$$\begin{aligned} \tilde{\mathbf{z}} &= \text{col}\{\mathbf{z}(2), \ldots, \mathbf{z}(N/2), \mathbf{z}^*(N/2+2), \ldots, \mathbf{z}^*(N)\} \\ \tilde{\mathbf{s}} &= \text{col}\{\mathbf{s}(2), \ldots, \mathbf{s}(N/2), \mathbf{s}^*(N/2+2), \ldots, \mathbf{s}^*(N)\} \end{aligned} \tag{20}$$

then equation (19) gives

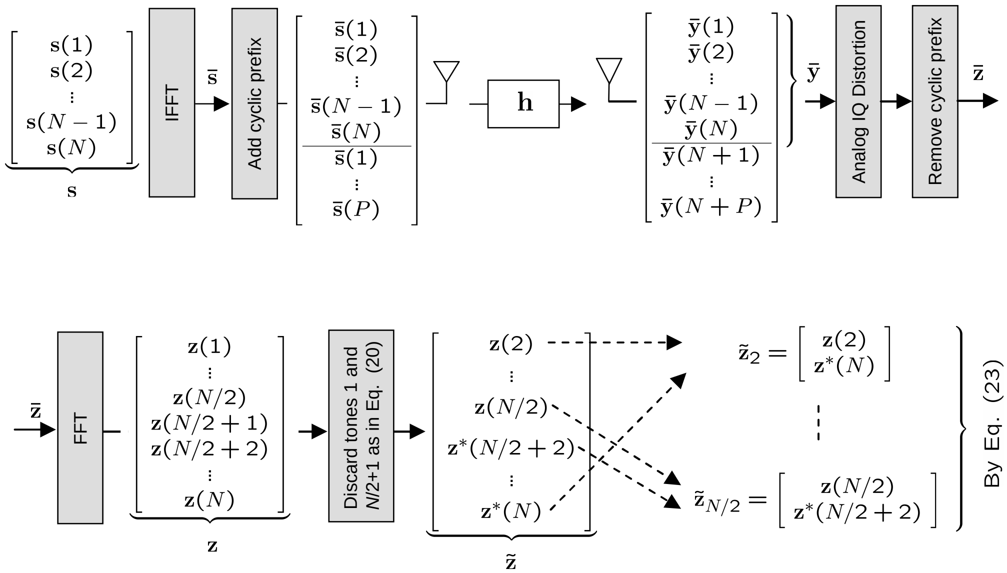

Fig. 1. An OFDM receiver with IQ imbalances and the notation used in the SISO derivations.

$$\tilde{\mathbf{z}} = \underbrace{\begin{bmatrix} \mu\lambda(2) & & & \nu\lambda^*(N) \\ & \ddots & \iddots & \\ & \mu\lambda(N/2) & \nu\lambda^*(N/2+2) & \\ & \nu^*\lambda(N/2) & \mu^*\lambda^*(N/2+2) & \\ & \iddots & \ddots & \\ \nu^*\lambda(2) & & & \mu^*\lambda^*(N) \end{bmatrix}}_{\tilde{\mathbf{\Lambda}}} \tilde{\mathbf{s}} + \tilde{\mathbf{v}} \tag{21}$$

where $\tilde{\mathbf{v}}$ is related to $\mathbf{v}$ in a manner similar to (20). Note that the matrix $\tilde{\mathbf{\Lambda}}$ in the above equation is not diagonal, as is the case for $\mathbf{\Lambda}$ in (10), although it collapses to a diagonal matrix by setting ν equal to zero. Still, equation (21) can be reduced to 2×2 decoupled sub-equations, for $k = \{2, \ldots, N/2\}$, each written as

$$\tilde{\mathbf{z}}_k = \tilde{\mathbf{\Gamma}}_k \tilde{\mathbf{s}}_k + \tilde{\mathbf{v}}_k \tag{22}$$

where

$$\tilde{\mathbf{z}}_k = \begin{bmatrix} \mathbf{z}(k) \\ \mathbf{z}^*(N-k+2) \end{bmatrix} \qquad \tilde{\mathbf{s}}_k = \begin{bmatrix} \mathbf{s}(k) \\ \mathbf{s}^*(N-k+2) \end{bmatrix} \tag{23}$$

$$\tilde{\mathbf{\Gamma}}_k = \begin{bmatrix} \mu\lambda(k) & \nu\lambda^*(N-k+2) \\ \nu^*\lambda(k) & \mu^*\lambda^*(N-k+2) \end{bmatrix} \tag{24}$$

The objective is to recover the data $\tilde{\mathbf{s}}_k$ from $\tilde{\mathbf{z}}_k$ in (22) for $k = \{2, \ldots, N/2\}$ or, equivalently, $\tilde{\mathbf{s}}$ from $\tilde{\mathbf{z}}$ in (21). Several algorithms, adaptive and otherwise, for both estimating the channel/disortion parameters and for recovering the $\tilde{\mathbf{s}}_k$ were proposed and studied in (Tarighat and Sayed, 2004*a*). In the next sections, we describe and derive similar algorithms for two broader scenarios, namely, i) for the case of a single user with 2-transmit 1-receive antenna and ii) for the case of multi-user transmissions with each user still using 2 transmit antennas and with one receive antenna per user.

3. ALAMOUTI SCHEME WITH IQ IMBALANCES

So let us now illustrate how IQ imbalances affect data recovery in a multi-antenna system (Tarighat and Sayed, 2004*b*). We consider first the case of a single user with a 2-transmit 1-receive antenna system employing the Alamouti scheme. This situation is illustrated in Figure 2, where it is indicated that two blocks of data

$$\mathbf{s}_1 \triangleq \begin{bmatrix} \mathbf{s}_1(1) \\ \vdots \\ \mathbf{s}_1(N) \end{bmatrix}, \quad \mathbf{s}_2 \triangleq \begin{bmatrix} \mathbf{s}_2(1) \\ \vdots \\ \mathbf{s}_2(N) \end{bmatrix}$$

are transmitted from both antennas before the IDFT operation. These blocks are then followed by the data

$$\begin{bmatrix} -\mathbf{s}_2^*(1) \\ \vdots \\ -\mathbf{s}_2^*(N) \end{bmatrix}, \quad \begin{bmatrix} \mathbf{s}_1^*(1) \\ \vdots \\ \mathbf{s}_1^*(N) \end{bmatrix}$$

At the receiver end, two blocks of data are received

$$\begin{bmatrix} \mathbf{z}_1(1) \\ \vdots \\ \mathbf{z}_1(N) \end{bmatrix}, \quad \begin{bmatrix} \mathbf{z}_2(1) \\ \vdots \\ \mathbf{z}_2(N) \end{bmatrix}$$

In this scenario, for instance, equation (6) would become

$$\bar{\mathbf{y}} = \mathbf{H}_1^c \bar{\mathbf{s}}_1 + \mathbf{H}_2^c \bar{\mathbf{s}}_2 + \bar{\mathbf{v}} \tag{26}$$

when $\bar{\mathbf{s}}_1 = \mathbf{F}^* \mathbf{s}_1$ and $\bar{\mathbf{s}}_2 = \mathbf{F}^* \mathbf{s}_2$ are the transmitted blocks from antennas 1 and 2, respectively, and $\mathbf{H}_1^c$ and $\mathbf{H}_2^c$ are the channel matrices from the transmit antennas 1 and 2 to the receiver. Applying the same arguments as before to the above equation instead of (6) results in the following sets of equations (which correspond to equations (22)-(24) in the single antenna case) (Tarighat and Sayed, 2004*b*):

$$\tilde{\mathbf{\Gamma}}_k^{(\mathrm{B})} = \begin{bmatrix} \mu\lambda_2(k) & \nu\lambda_2^*(N-k+2) & \mu\lambda_1(k) & \nu\lambda_1^*(N-k+2) \\ \nu^*\lambda_2(k) & \mu^*\lambda_2^*(N-k+2) & \nu^*\lambda_1(k) & \mu^*\lambda_1^*(N-k+2) \end{bmatrix} \tag{25}$$

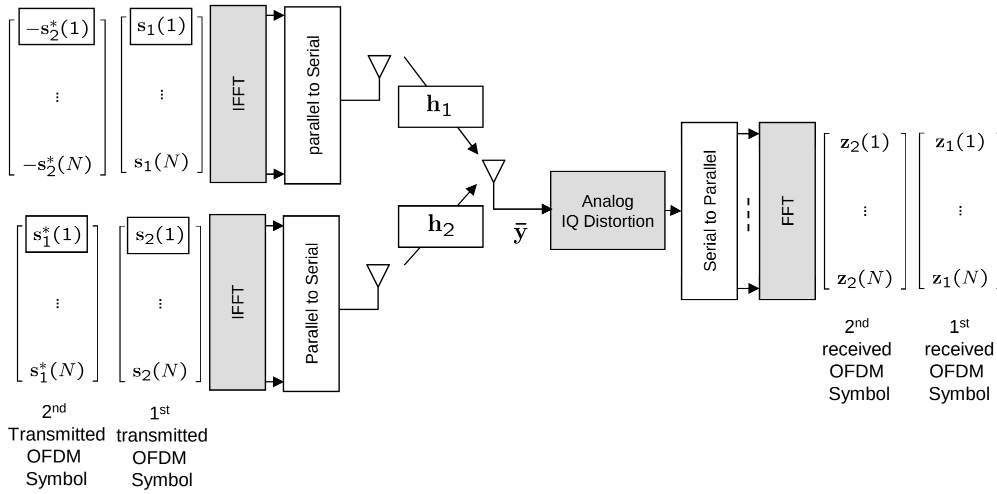

Fig. 2. Alamouti scheme applied to an OFDM system.

$$\begin{bmatrix} \mathbf{z}_1(k) \\ \mathbf{z}_1^*(N-k+2) \end{bmatrix} = \underbrace{\left[\tilde{\mathbf{\Gamma}}_{2,k} \;\; \tilde{\mathbf{\Gamma}}_{1,k}\right]}_{\tilde{\mathbf{\Gamma}}_k^{(\mathrm{B})}} \begin{bmatrix} \mathbf{s}_2(k) \\ \mathbf{s}_2^*(N-k+2) \\ \mathbf{s}_1(k) \\ \mathbf{s}_1^*(N-k+2) \end{bmatrix} + \tilde{\mathbf{v}}_{1,k} \tag{27}$$

and

$$\begin{bmatrix} \mathbf{z}_2(k) \\ \mathbf{z}_2^*(N-k+2) \end{bmatrix} = \left[\tilde{\mathbf{\Gamma}}_{2,k} \;\; \tilde{\mathbf{\Gamma}}_{1,k}\right] \begin{bmatrix} \mathbf{s}_1^*(k) \\ \mathbf{s}_1(N-k+2) \\ -\mathbf{s}_2^*(k) \\ -\mathbf{s}_2(N-k+2) \end{bmatrix} + \tilde{\mathbf{v}}_{2,k} \tag{28}$$

where

$$\begin{aligned} \tilde{\mathbf{\Gamma}}_{1,k} &= \begin{bmatrix} \mu\lambda_1(k) & \nu\lambda_1^*(N-k+2) \\ \nu^*\lambda_1(k) & \mu^*\lambda_1^*(N-k+2) \end{bmatrix} \\ \tilde{\mathbf{\Gamma}}_{2,k} &= \begin{bmatrix} \mu\lambda_2(k) & \nu\lambda_2^*(N-k+2) \\ \nu^*\lambda_2(k) & \mu^*\lambda_2^*(N-k+2) \end{bmatrix} \end{aligned} \tag{29}$$

The matrix $\tilde{\mathbf{\Gamma}}_k^{(\mathrm{B})}$ is now 2×4 and the $\lambda_i(k)$ in (29) denote the channel taps in the frequency domain corresponding to tone k from transmit antenna i to the receive antenna.

Compactly, we combine (27)–(28) as

$$\tilde{\mathbf{z}}_k^{(\mathrm{B})} = \tilde{\mathbf{\Gamma}}_k^{(\mathrm{B})} \tilde{\mathbf{s}}_k^{(\mathrm{B})} + \tilde{\mathbf{v}}_k^{(\mathrm{B})} \tag{30}$$

where the 2×4 matrix $\tilde{\mathbf{\Gamma}}_k^{(\mathrm{B})}$ is given by (25), and

$$\tilde{\mathbf{s}}_k^{(\mathrm{B})} = \begin{bmatrix} \mathbf{s}_1^*(k) & \mathbf{s}_2(k) \\ \mathbf{s}_1(N-k+2) & \mathbf{s}_2^*(N-k+2) \\ -\mathbf{s}_2^*(k) & \mathbf{s}_1(k) \\ -\mathbf{s}_2(N-k+2) & \mathbf{s}_1^*(N-k+2) \end{bmatrix} \tag{31}$$

$$\tilde{\mathbf{z}}_k^{(\mathrm{B})} = \begin{bmatrix} \mathbf{z}_2(k) & \mathbf{z}_1(k) \\ \mathbf{z}_2^*(N-k+2) & \mathbf{z}_1^*(N-k+2) \end{bmatrix} \tag{32}$$

In order to recover the data in $\tilde{\mathbf{s}}_k^{(\mathrm{B})}$ from $\tilde{\mathbf{z}}_k^{(\mathrm{B})}$ in (30), the space-time code structure can be exploited as follows. Let

$$\tilde{\mathbf{z}}_k = \begin{bmatrix} \mathbf{z}_1(k) \\ \mathbf{z}_2^*(k) \\ \mathbf{z}_1^*(N-k+2) \\ \mathbf{z}_2(N-k+2) \end{bmatrix}, \quad \tilde{\mathbf{s}}_k = \begin{bmatrix} \mathbf{s}_2(k) \\ \mathbf{s}_1(k) \\ \mathbf{s}_2^*(N-k+2) \\ \mathbf{s}_1^*(N-k+2) \end{bmatrix} \tag{33}$$

Then (30) gives

$$\tilde{\mathbf{z}}_k = \tilde{\mathbf{\Gamma}}_k \tilde{\mathbf{s}}_k + \tilde{\mathbf{v}}_k \tag{34}$$

with

$$\tilde{\mathbf{\Gamma}}_k = \begin{bmatrix} \mathcal{G}_{1,k} & \mathcal{G}_{2,k} \\ \mathcal{G}_{3,k} & \mathcal{G}_{4,k} \end{bmatrix} \quad (4 \times 4) \tag{35}$$

where, interestingly, all the sub-blocks have an Alamouti structure given by

$$\begin{aligned} \mathcal{G}_{1,k} &= \begin{bmatrix} \mu\lambda_2(k) & \mu\lambda_1(k) \\ -\mu^*\lambda_1^*(k) & \mu^*\lambda_2^*(k) \end{bmatrix} \\ \mathcal{G}_{2,k} &= \begin{bmatrix} \nu\lambda_2^*(N-k+2) & \nu\lambda_1^*(N-k+2) \\ -\nu^*\lambda_1(N-k+2) & \nu^*\lambda_2(N-k+2) \end{bmatrix} \\ \mathcal{G}_{3,k} &= \begin{bmatrix} \nu^*\lambda_2(k) & \nu^*\lambda_1(k) \\ -\nu\lambda_1^*(k) & \nu\lambda_2^*(k) \end{bmatrix} \\ \mathcal{G}_{4,k} &= \begin{bmatrix} \mu^*\lambda_2^*(N-k+2) & \mu^*\lambda_1^*(N-k+2) \\ -\mu\lambda_1(N-k+2) & \mu\lambda_2(N-k+2) \end{bmatrix} \end{aligned} \tag{36}$$

3.1 A Regularized Least-Squares Receiver

The matrix $\tilde{\mathbf{\Gamma}}_k$ in (35) would be block-diagonal if $\nu = 0$, i.e., if there were no IQ imbalances, in which case equation (34) would reduce to two 2×2 decoupled systems as in standard Alamouti decoding. The off-diagonal matrices in (35) are a result of the IQ imbalances. So we now need to deal with a 4×4 linear system of equations as opposed to two 2×2 systems of equations in Alamouti coded receivers with ideal IQ branches. Still, the Alamouti structure

$$\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k = \begin{bmatrix} (|\mu|^2 + |\nu|^2) \left(|\lambda_1(k)|^2 + |\lambda_2(k)|^2 \right) \mathbf{I}_{2\times 2} & \mathcal{G}_{1,k}^* \mathcal{G}_{2,k} + \mathcal{G}_{3,k}^* \mathcal{G}_{4,k} \\ \mathcal{G}_{2,k}^* \mathcal{G}_{1,k} + \mathcal{G}_{4,k}^* \mathcal{G}_{3,k} & (|\mu|^2 + |\nu|^2) \left(|\lambda_1(N-k+2)|^2 + |\lambda_2(N-k+2)|^2 \right) \mathbf{I}_{2\times 2} \end{bmatrix} \quad (37)$$

of the sub-matrices (36) allows a computationally efficient implementation of the regularized least-squares estimator (Sayed, 2003):

$$\hat{\tilde{\mathbf{s}}}_k = \left(\delta \mathbf{I}_{4\times 4} + \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k \right)^{-1} \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{z}}_k \quad (38)$$

where δ is a positive number. There are some advantages associated with the regularized solution compared to the standard least-squares solution, such as avoiding difficulties that may arise from data ill-conditioning. Now note that (Tarighat and Sayed, 2004*b*):

$$\mathcal{G}_{1,k}^* \mathcal{G}_{1,k} = |\mu|^2 \left(|\lambda_1(k)|^2 + |\lambda_2(k)|^2 \right) \mathbf{I}_{2\times 2}$$
$$\mathcal{G}_{2,k}^* \mathcal{G}_{2,k} = |\nu|^2 \left(|\lambda_1(N-k+2)|^2 + |\lambda_2(N-k+2)|^2 \right) \mathbf{I}_{2\times 2}$$
$$\mathcal{G}_{3,k}^* \mathcal{G}_{3,k} = |\nu|^2 \left(|\lambda_1(k)|^2 + |\lambda_2(k)|^2 \right) \mathbf{I}_{2\times 2}$$
$$\mathcal{G}_{4,k}^* \mathcal{G}_{4,k} = |\mu|^2 \left(|\lambda_1(N-k+2)|^2 + |\lambda_2(N-k+2)|^2 \right) \mathbf{I}_{2\times 2}$$

so that the product $\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k$ reduces to the result given by (37). This matrix would be diagonal if $\nu = 0$, i.e., if there were no IQ imbalances. When $\nu \neq 0$, the matrix $\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k$ is no longer diagonal, however it has a particular structure that is induced by the Alamouti code and the distortion model. Specifically, its 2×2 off-diagonal blocks are Alamouti, which means their inverses can be obtained by simple transposition. This is due to the fact that the sum and product of two Alamouti matrices is still an Alamouti matrix. Thus denote the 2×2 block entries of $\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k$ by

$$\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k = \begin{bmatrix} \mathbf{D}_1 & \mathbf{A}_1 \\ \mathbf{A}_1^* & \mathbf{D}_2 \end{bmatrix}$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal (actually scalar multiples of the identity due to the Alamouti structure), say $\mathbf{D}_1 = d_1\mathbf{I}$ and $\mathbf{D}_2 = d_2\mathbf{I}$, and $\mathbf{A}_1$ and $\mathbf{A}_2$ are also Alamouti. Then

$$\delta \mathbf{I}_{4\times 4} + \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k = \begin{bmatrix} (\delta + d_1)\mathbf{I} & \mathbf{A}_1 \\ \mathbf{A}_1^* & (\delta + d_2)\mathbf{I} \end{bmatrix} \quad (39)$$

Using the block inversion formula

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \Sigma^{-1} & -\Sigma^{-1} B D^{-1} \\ -D^{-1} C \Sigma^{-1} & D^{-1} + D^{-1} C \Sigma^{-1} B D \end{bmatrix}$$

where

$$\Sigma = A - BD^{-1}C$$

and applying it to (39), we get

$$\begin{aligned} \mathbf{\Sigma} &= (\delta + d_1)\mathbf{I} - \mathbf{A}_1((\delta + d_2)\mathbf{I})^{-1}\mathbf{A}_1^* \\ &= (\delta + d_1)\mathbf{I} - \frac{1}{\delta + d_2}\mathbf{A}_1\mathbf{A}_1^* \\ &= \left(\delta + d_1 - \frac{d_3}{\delta + d_2} \right) \mathbf{I} \end{aligned}$$

since $\mathbf{A}_1\mathbf{A}_1^* = d_3\mathbf{I}$ for some scalar d_3. Therefore, we see that all terms in the expression for $\left(\delta \mathbf{I}_{4\times 4} + \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k \right)^{-1}$ are trivial to compute.

Alternatively, the estimate of $\hat{\tilde{\mathbf{s}}}_k$ in (38) can be computed as follows (this alternative procedure is useful for the multi-user scenario. We explain it here for the single-user case for illustration purposes).

- Introduce the triangular (Cholesky) factorization of $\tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k$, namely,

$$\begin{aligned} \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k &\triangleq \mathcal{L}_k \mathcal{D}_k \mathcal{L}_k^* \\ &= \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{A}_1^*/d_1 & \mathbf{I} \end{bmatrix} \begin{bmatrix} d_1\mathbf{I} & \mathbf{0} \\ \mathbf{0} & (d_2 - d_3/d_1)\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}_1/d_1 \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \end{aligned}$$

  where

$$\mathcal{L}_k = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{A}_1^*/d_1 & \mathbf{I} \end{bmatrix}, \quad \mathcal{D}_k = \begin{bmatrix} d_1\mathbf{I} & \mathbf{0} \\ \mathbf{0} & (d_2 - d_3/d_1)\mathbf{I} \end{bmatrix} \quad (40)$$

- Let $\mathbf{s}_k' = \mathcal{L}_k^* \tilde{\mathbf{s}}_k$. Then we can rewrite (34) as

$$\begin{aligned} \tilde{\mathbf{z}}_k &= \left(\tilde{\mathbf{\Gamma}}_k \mathcal{L}_k^{-*} \right) (\mathcal{L}_k^* \tilde{\mathbf{s}}_k) + \tilde{\mathbf{v}}_k \\ &= \left(\tilde{\mathbf{\Gamma}}_k \mathcal{L}_k^{-*} \right) \mathbf{s}_k' + \tilde{\mathbf{v}}_k \end{aligned} \quad (41)$$

  where, from (40),

$$\mathcal{L}_k^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_1^*/d_1 & \mathbf{I} \end{bmatrix} \quad (42)$$

- Then, using (41), the least-squares estimator of $\mathbf{s}_k'$ is given by

$$\begin{aligned} \hat{\mathbf{s}}_k' &= \left(\delta'\mathbf{I} + \mathcal{L}_k^{-1} \tilde{\mathbf{\Gamma}}_k^* \tilde{\mathbf{\Gamma}}_k \mathcal{L}_k^{-*} \right)^{-1} \left(\mathcal{L}_k^{-1} \tilde{\mathbf{\Gamma}}_k^* \right) \tilde{\mathbf{z}}_k \\ &= (\delta'\mathbf{I} + \mathcal{D}_k)^{-1} \left(\mathcal{L}_k^{-1} \tilde{\mathbf{\Gamma}}_k^* \right) \tilde{\mathbf{z}}_k \end{aligned} \quad (43)$$

- Finally, the decisions for the entries of $\tilde{\mathbf{s}}_k$ are obtained from $\hat{\mathbf{s}}_k'$ via the transformation

$$\hat{\tilde{\mathbf{s}}}_k = \mathcal{L}_k^{-*} \hat{\mathbf{s}}_k'$$

3.2 An Adaptive Equalizer

The least-squares solution (38) requires knowledge of $\tilde{\mathbf{\Gamma}}_k$, which in turn requires knowledge of the channel parameters $\{\lambda_i(k)\}$ and the distortion model parameters $\{\mu, \nu\}$. By examining the structure of (38), we note that the mapping from $\tilde{\mathbf{z}}_k$ to $\hat{\tilde{\mathbf{s}}}_k$ can be written as follows

$$\hat{\tilde{\mathbf{s}}}_k = \mathbf{W}_k \tilde{\mathbf{z}}_k \quad (44)$$

where the 4×4 matrix $\mathbf{W}_k$ has the form

$$\mathbf{W}_k = \begin{bmatrix} \mathbf{W}_{1,k} & \mathbf{W}_{2,k} \\ \mathbf{W}_{3,k} & \mathbf{W}_{4,k} \end{bmatrix} \quad (45)$$

and each $\mathbf{W}_{i,k}$, $i = 1, \ldots, 4$ is a 2×2 Alamouti matrix. That is,

$$\mathbf{W}_{i,k} = \begin{bmatrix} \mathbf{W}_{i,k}(1) & \mathbf{W}_{i,k}(2) \\ -\mathbf{W}_{i,k}^*(2) & \mathbf{W}_{i,k}^*(1) \end{bmatrix} \quad (46)$$

This result follows from the following properties of Alamouti matrices:

- The sum or difference of two Alamouti matrices is an Alamouti matrix.
- The inverse of an Alamouti matrix is another Alamouti matrix.
- The inverse of a block matrix with Alamouti subblocks is a block matrix with Alamouti subblocks.

The matrix $\mathbf{W}_k$ in (44) can be interpreted as the coefficient matrix of an equalizer that operates on the received data $\tilde{\mathbf{z}}_k$ and provides the signal estimates $\hat{\tilde{\mathbf{s}}}_\mathbf{k}$. One way to avoid the need for explicit channel and distortion model information at the receiver, as well as to enable the receiver to track variations in this model, is to determine the entries of $\mathbf{W}_k$ adaptively.

By expanding (44), we get

$$\left[\begin{array}{c}\hat{\mathbf{s}}_2(k)\\ \hat{\mathbf{s}}_1(k)\\ \hline \hat{\mathbf{s}}_2^*(N-k+2)\\ \hat{\mathbf{s}}_1^*(N-k+2)\end{array}\right]=\left[\begin{array}{c|c}\mathbf{W}_{1,k}&\mathbf{W}_{2,k}\\ \hline \mathbf{W}_{3,k}&\mathbf{W}_{4,k}\end{array}\right]\left[\begin{array}{c}\mathbf{z}_1(k)\\ \mathbf{z}_2^*(k)\\ \hline \mathbf{z}_1^*(N-k+2)\\ \mathbf{z}_2(N-k+2)\end{array}\right] \tag{47}$$

Then (47) can be rewritten as

$$\left[\begin{array}{c}\hat{\mathbf{s}}_2(k)\\ \hat{\mathbf{s}}_1(k)\end{array}\right]=\left[\begin{array}{c|c}\mathbf{W}_{1,k}&\mathbf{W}_{2,k}\end{array}\right]\left[\begin{array}{c}\mathbf{z}_1(k)\\ \mathbf{z}_2^*(k)\\ \hline \mathbf{z}_1^*(N-k+2)\\ \mathbf{z}_2(N-k+2)\end{array}\right] \tag{48}$$

and

$$\left[\begin{array}{c}\hat{\mathbf{s}}_2^*(N-k+2)\\ \hat{\mathbf{s}}_1^*(N-k+2)\end{array}\right]=\left[\begin{array}{c|c}\mathbf{W}_{3,k}&\mathbf{W}_{4,k}\end{array}\right]\left[\begin{array}{c}\mathbf{z}_1(k)\\ \mathbf{z}_2^*(k)\\ \hline \mathbf{z}_1^*(N-k+2)\\ \mathbf{z}_2(N-k+2)\end{array}\right] \tag{49}$$

Alternatively, (48) and (49) can be written as (50) and (51), respectively. Now let i denote a block iteration index and define

$$\tilde{\mathbf{Z}}_i=\left[\begin{array}{cc|cc}\mathbf{z}_1(k)&\mathbf{z}_2^*(k)&\mathbf{z}_1^*(N-k+2)&\mathbf{z}_2(N-k+2)\\ -\mathbf{z}_2(k)&\mathbf{z}_1^*(k)&-\mathbf{z}_2^*(N-k+2)&\mathbf{z}_1(N-k+2)\end{array}\right]$$

where the entries of $\tilde{\mathbf{Z}}_i$ are obtained from the received OFDM symbol at block time i, as given by (47). Let further

$$\hat{\mathbf{S}}_{1,i}=\left[\begin{array}{c}\hat{\mathbf{s}}_2(k)\\ -\hat{\mathbf{s}}_1^*(k)\end{array}\right],\ \hat{\mathbf{S}}_{2,i}=\left[\begin{array}{c}\hat{\mathbf{s}}_2^*(N-k+2)\\ -\hat{\mathbf{s}}_1(N-k+2)\end{array}\right],$$

$$\mathcal{W}_{1,i}=\left[\begin{array}{c}\mathbf{W}_{1,k}(1)\\ \mathbf{W}_{1,k}(2)\\ \mathbf{W}_{2,k}(1)\\ \mathbf{W}_{2,k}(2)\end{array}\right],\ \mathcal{W}_{2,i}=\left[\begin{array}{c}\mathbf{W}_{3,k}(1)\\ \mathbf{W}_{3,k}(2)\\ \mathbf{W}_{4,k}(1)\\ \mathbf{W}_{4,k}(2)\end{array}\right],$$

and assume we collect the received blocks corresponding to 2 consecutive OFDM symbols into a 4×4 matrix $\mathcal{Z}_i$ as follows

$$\mathcal{Z}_i=\left[\begin{array}{c}\tilde{\mathbf{Z}}_i\\ \tilde{\mathbf{Z}}_{i+1}\end{array}\right]$$

Assuming that the channel is fixed over the 2 OFDM symbols, then we can concatenate the vectors in (48) and (49) from 2 OFDM symbols and write

$$\hat{\mathcal{S}}_{1,i}=\left[\begin{array}{c}\hat{\mathbf{S}}_{1,i}\\ \hat{\mathbf{S}}_{1,i+1}\end{array}\right]=\mathcal{Z}_i\mathcal{W}_{1,i} \tag{52}$$

$$\hat{\mathcal{S}}_{2,i}=\left[\begin{array}{c}\hat{\mathbf{S}}_{2,i}\\ \hat{\mathbf{S}}_{2,i+1}\end{array}\right]=\mathcal{Z}_i\mathcal{W}_{2,i} \tag{53}$$

where $\mathcal{Z}_i$ has Alamouti subblocks. Moreover, $\mathcal{W}_{1,i}$ and $\mathcal{W}_{2,i}$ are 4×1 vectors containing the entries of the matrix $\mathbf{W}$ in (44). Equations (52) and (53) reveal the special structure of the STBC problem. Rather than estimate the 16 entries of $\mathbf{W}_k$, we only need to estimate the 8 entries of $\mathcal{W}_{1,i}$ and $\mathcal{W}_{2,i}$. Moreover, the matrix $\mathcal{Z}_i$ itself has Alamouti subblocks. Now we show how $\mathcal{W}_{1,i}$ and $\mathcal{W}_{2,i}$ can be estimated adaptively by using, for example, a block version of the RLS algorithm. In addition, the structure of the STBC, as revealed by $\mathcal{Z}_i$, will be exploited to reduce the complexity of the RLS implementation.

To begin with, we note that the 4×4 matrix $\mathcal{Z}_i^*\mathcal{Z}_i$ has a block structure similar to that of $\tilde{\mathbf{\Gamma}}_k^*\tilde{\mathbf{\Gamma}}_k$ in (39). That is, $\mathcal{Z}_i^*\mathcal{Z}_i$ is Hermitian with diagonal blocks being multiples of the identity matrix $\mathbf{I}$ and with the off diagonal blocks being 2×2 Alamouti, i.e.,

$$\mathcal{Z}_i^*\mathcal{Z}_i=\left[\begin{array}{cc}\alpha_1\mathbf{I}&\mathbf{Z}_1\\ \mathbf{Z}_1^*&\alpha_2\mathbf{I}\end{array}\right] \tag{54}$$

for some scalars α_1 and α_2. To derive an adaptive algorithm for estimating $\{\mathcal{W}_1,\mathcal{W}_2\}$, we first introduce an iteration index i. At iteration i, the data matrix $\mathcal{Z}$ is denoted by $\mathcal{Z}_i$. Let

$$\begin{aligned}\mathcal{Z}_i^*\mathcal{Z}_i &\stackrel{\triangle}{=}\mathcal{L}_i\mathcal{D}_i\mathcal{L}_i^*\\ &=\left[\begin{array}{cc}\mathbf{I}&\mathbf{0}\\ \mathbf{Z}_1^*/\alpha_1&\mathbf{I}\end{array}\right]\left[\begin{array}{cc}\alpha_1\mathbf{I}&\mathbf{0}\\ \mathbf{0}&(\alpha_2-\alpha_3/\alpha_1)\,\mathbf{I}\end{array}\right]\left[\begin{array}{cc}\mathbf{I}&\mathbf{Z}_1/\alpha_1\\ \mathbf{0}&\mathbf{I}\end{array}\right]\end{aligned}$$

where

$$\mathcal{L}_i=\left[\begin{array}{cc}\mathbf{I}&\mathbf{0}\\ \mathbf{Z}_1^*/\alpha_1&\mathbf{I}\end{array}\right],\quad \mathcal{D}_i=\left[\begin{array}{cc}\alpha_1\mathbf{I}&\mathbf{0}\\ \mathbf{0}&(\alpha_2-\alpha_3/\alpha_1)\,\mathbf{I}\end{array}\right] \tag{55}$$

and $\mathbf{Z}_1\mathbf{Z}_1^*=\alpha_3\mathbf{I}$ for some scalar α_3. Then, we can rewrite (52) and (53) at block time i as

$$\hat{\mathcal{S}}_{1,i}=\left(\mathcal{Z}_i\mathcal{L}_i^{-*}\right)\left(\mathcal{L}_i^*\mathcal{W}_{1,i}\right)\stackrel{\triangle}{=}\tilde{\mathcal{Z}}_i\tilde{\mathcal{W}}_{1,i} \tag{56}$$

$$\hat{\mathcal{S}}_{i,2}=\left(\mathcal{Z}_i\mathcal{L}_i^{-*}\right)\left(\mathcal{L}_i^*\mathcal{W}_{2,i}\right)\stackrel{\triangle}{=}\tilde{\mathcal{Z}}_i\tilde{\mathcal{W}}_{2,i} \tag{57}$$

in terms of transformed variables $\{\tilde{\mathcal{Z}}_i,\tilde{\mathcal{W}}_{1,i},\tilde{\mathcal{W}}_{2,i}\}$. By working with the transformed variables, the complexity of the RLS recursions would simplify to that of an LMS implementation; the additional cost would be that of evaluating the transformed variable $\tilde{\mathcal{Z}}_i$. Using (56) and (57), the estimates of $\mathcal{W}_1$ and $\mathcal{W}_2$ are updated at every 2 blocks according to the following RLS recursions:

$$\begin{aligned}\tilde{\mathcal{W}}_{1,i+2}&=\tilde{\mathcal{W}}_{1,i}+\lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\mathbf{\Pi}_{i+2}\left[\mathcal{D}_{1,i+2}-\tilde{\mathcal{Z}}_{i+2}\tilde{\mathcal{W}}_{1,i}\right]\\ \tilde{\mathcal{W}}_{2,i+2}&=\tilde{\mathcal{W}}_{2,i}+\lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\mathbf{\Pi}_{i+2}\left[\mathcal{D}_{2,i+2}-\tilde{\mathcal{Z}}_{i+2}\tilde{\mathcal{W}}_{2,i}\right]\end{aligned} \tag{58}$$

where

$$\mathcal{P}_{i+2}=\lambda^{-1}\left[\mathcal{P}_i-\lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\mathbf{\Pi}_{i+2}\tilde{\mathcal{Z}}_{i+2}\mathcal{P}_i^*\right] \tag{59}$$

$$\begin{bmatrix} \hat{\mathbf{s}}_2(k) \\ -\hat{\mathbf{s}}_1^*(k) \end{bmatrix} = \left[\begin{array}{cc|cc} \mathbf{z}_1(k) & \mathbf{z}_2^*(k) & \mathbf{z}_1^*(N-k+2) & \mathbf{z}_2(N-k+2) \\ -\mathbf{z}_2(k) & \mathbf{z}_1^*(k) & -\mathbf{z}_2^*(N-k+2) & \mathbf{z}_1(N-k+2) \end{array}\right] \begin{bmatrix} \mathbf{W}_{1,k}(1) \\ \mathbf{W}_{1,k}(2) \\ \mathbf{W}_{2,k}(1) \\ \mathbf{W}_{2,k}(2) \end{bmatrix} \tag{50}$$

and

$$\begin{bmatrix} \hat{\mathbf{s}}_2^*(N-k+2) \\ -\hat{\mathbf{s}}_1(N-k+2) \end{bmatrix} = \left[\begin{array}{cc|cc} \mathbf{z}_1(k) & \mathbf{z}_2^*(k) & \mathbf{z}_1^*(N-k+2) & \mathbf{z}_2(N-k+2) \\ -\mathbf{z}_2(k) & \mathbf{z}_1^*(k) & -\mathbf{z}_2^*(N-k+2) & \mathbf{z}_1(N-k+2) \end{array}\right] \begin{bmatrix} \mathbf{W}_{3,k}(1) \\ \mathbf{W}_{3,k}(2) \\ \mathbf{W}_{4,k}(1) \\ \mathbf{W}_{4,k}(2) \end{bmatrix} \tag{51}$$

and

$$\mathbf{\Pi}_{i+2} = \left(\mathbf{I} + \lambda^{-1}\tilde{\mathcal{Z}}_{i+2}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\right)^{-1} \tag{60}$$

and λ is a forgetting factor that is usually close to 1. The initial conditions are $\mathcal{W}_{1,0} = 0$ and $\mathcal{W}_{2,0} = 0$ and $\mathcal{P}_0 = \delta\mathbf{I}$, δ is a large number. Moreover, $\mathcal{D}_{1,i+2}$ and $\mathcal{D}_{2,i+2}$ are the desired response vectors given by

$$\mathcal{D}_{1,i+2} = \begin{cases} \mathcal{S}_{1,i+2} & \text{for training} \\ \check{\mathcal{S}}_{1,i+2} & \text{for tracking} \end{cases}$$

and

$$\mathcal{D}_{2,i+2} = \begin{cases} \mathcal{S}_{2,i+2} & \text{for training} \\ \check{\mathcal{S}}_{2,i+2} & \text{for tracking} \end{cases}$$

where the notation $\check{\mathbf{s}}$ denotes the output of a decision device during a decision-directed mode of operation. By using the matrix inversion Lemma (Sayed, 2003), we can rewrite (60) as

$$\begin{aligned}\mathbf{\Pi}_{i+2} &= \mathbf{I} - \tilde{\mathcal{Z}}_{i+2}\left(\lambda\mathcal{P}_i^{-1} + \tilde{\mathcal{Z}}_{i+2}^*\tilde{\mathcal{Z}}_{i+2}\right)^{-1}\tilde{\mathcal{Z}}_{i+2}^* \\ &= \mathbf{I} - \tilde{\mathcal{Z}}_{i+2}\left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2}\right)^{-1}\tilde{\mathcal{Z}}_{i+2}^* \end{aligned} \tag{61}$$

Substituting (61) into (59), we get

$$\begin{aligned}\mathcal{P}_{i+2} =& \lambda^{-1}\Big[\mathcal{P}_i - \lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\Big(\mathbf{I} - \tilde{\mathcal{Z}}_{i+2} \\ &\left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2}\right)^{-1}\tilde{\mathcal{Z}}_{i+2}^*\Big)\tilde{\mathcal{Z}}_{i+2}\mathcal{P}_i^*\Big] \\ =& \lambda^{-1}\mathcal{P}_i - \lambda^{-2}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2}^*\tilde{\mathcal{Z}}_{i+2} \\ &\left[\mathbf{I} - \left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2}\right)^{-1}\tilde{\mathcal{Z}}_{i+2}^*\tilde{\mathcal{Z}}_{i+2}\right]\mathcal{P}_i^* \\ =& \lambda^{-1}\mathcal{P}_i - \lambda^{-2}\mathcal{P}_i\mathcal{D}_{i+2} \\ &\left[\mathbf{I} - \left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2}\right)^{-1}\mathcal{D}_{i+2}\right]\mathcal{P}_i^* \end{aligned} \tag{62}$$

where (58) and (62) are used to update the adaptive equalizer coefficients. Due to the special structure of the space–time block–code, we now verify that the complexity of the RLS algorithm can be reduced to that of an LMS implementation. The reasoning is as follows. It follows by induction that $\mathcal{P}_{i+2}$ has a diagonal structure of the form

$$\mathcal{P}_{i+2} = \begin{pmatrix} p_1(i+2)\mathbf{I} & \mathbf{0} \\ \mathbf{0} & p_2(i+2)\mathbf{I} \end{pmatrix} \tag{63}$$

for some scalars $\{p_1(i+2), p_2(i+2)\}$. This statement holds at time $i+2=0$ since, by assumption, $\mathcal{P}_0 = \delta\mathbf{I}$ (so that $p_1 = p_2 = \delta$). Now assume the statement holds at time i. Then it is easy to see that $p_1(i+2)$ and $p_2(i+2)$ are given by

$$\begin{aligned} p_1(i+2) &= \lambda^{-1}p_1(i) - \lambda^{-2}p_1^2(i)\alpha_1\left[1 - \frac{\alpha_1}{\lambda p_1^{-1}(i) + \alpha_1}\right] \\ &= \frac{1}{\lambda p_1^{-1}(i) + \alpha_1} \end{aligned} \tag{64}$$

$$p_2(i+2) = \frac{1}{\lambda p_2^{-1}(i) + \alpha_2 - \alpha_3/\alpha_1} \tag{65}$$

which has the desired diagonal structure. Substituting (61) and (63) into (58) we get the following RLS update equation:

$$\begin{aligned} \tilde{\mathcal{W}}_{1,i+2} &= \tilde{\mathcal{W}}_{1,i} + \mathcal{P}_{i+2}\tilde{\mathcal{Z}}_{i+2}^*\left[\mathcal{D}_{1,i+2} - \tilde{\mathcal{Z}}_{i+2}\tilde{\mathcal{W}}_{1,i}\right] \\ \tilde{\mathcal{W}}_{2,i+2} &= \tilde{\mathcal{W}}_{2,i} + \mathcal{P}_{i+2}\tilde{\mathcal{Z}}_{i+2}^*\left[\mathcal{D}_{2,i+2} - \tilde{\mathcal{Z}}_{i+2}\tilde{\mathcal{W}}_{2,i}\right] \end{aligned} \tag{66}$$

Finally, equations (63)–(66) are used to update the adaptive equalizer coefficients. The block diagram of the adaptive receiver is shown in Figure 3.

4. MULTI–USER SCENARIO

Consider now an M-user scenario where each user is equipped with two transmit antennas and with the receiver having one receive antenna per user. Let $\tilde{\mathbf{\Gamma}}_{k,i}^{(j)}$, $i,j = \{1,\ldots,M\}$, denote the 4×4 channel matrix between user j and receive antenna i, defined in a manner similar to (35). Furthermore, let $\tilde{\mathbf{s}}_k^{(j)}$ represent the transmitted data by user j and let $\tilde{\mathbf{z}}_{k,i}$ represent the received data at receive antenna i, both defined as in (33). It can be verified that the system of equations (34) generalizes to

$$\begin{bmatrix} \tilde{\mathbf{z}}_{k,1} \\ \tilde{\mathbf{z}}_{k,2} \\ \vdots \\ \tilde{\mathbf{z}}_{k,M} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{\Gamma}}_{k,1}^{(1)} & \tilde{\mathbf{\Gamma}}_{k,1}^{(2)} & \cdots & \tilde{\mathbf{\Gamma}}_{k,1}^{(M)} \\ \tilde{\mathbf{\Gamma}}_{k,2}^{(1)} & \tilde{\mathbf{\Gamma}}_{k,2}^{(2)} & \cdots & \tilde{\mathbf{\Gamma}}_{k,2}^{(M)} \\ \vdots & \ddots & \ddots & \vdots \\ \tilde{\mathbf{\Gamma}}_{k,M}^{(1)} & \tilde{\mathbf{\Gamma}}_{k,M}^{(2)} & \cdots & \tilde{\mathbf{\Gamma}}_{k,M}^{(M)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{s}}_k^{(1)} \\ \tilde{\mathbf{s}}_k^{(2)} \\ \vdots \\ \tilde{\mathbf{s}}_k^{(M)} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{v}}_{k,1} \\ \tilde{\mathbf{v}}_{k,2} \\ \vdots \\ \tilde{\mathbf{v}}_{k,M} \end{bmatrix} \tag{67}$$

The transmission scheme is shown in Figure 4. The above equation can be represented more compactly as

$$\mathbf{Z} = \mathbf{\Gamma}\mathbf{S} + \mathbf{V} \tag{68}$$

so that the regularized least squares estimation of $\mathbf{S}$ is

$$\hat{\mathbf{S}} = \left(\delta\mathbf{I} + \mathbf{\Gamma}^*\mathbf{\Gamma}\right)^{-1}\mathbf{\Gamma}^*\mathbf{Z} \tag{69}$$

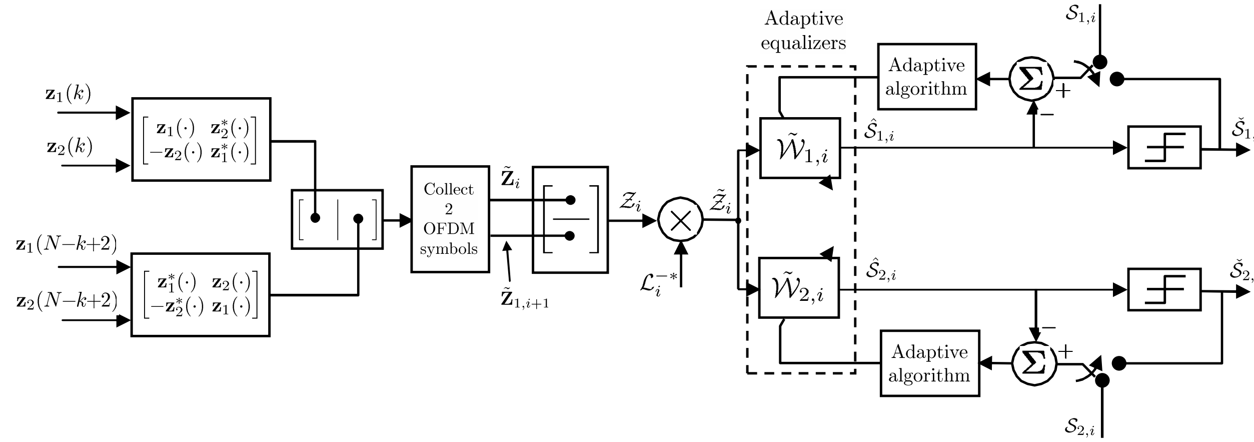

Fig. 3. An OFDM adaptive receiver for single user transmissions with IQ imbalances using 2-transmit and 1-receive antennas.

By inspecting the structure of (69), we observe that $\mathbf{\Gamma}^*\mathbf{\Gamma}$ is a Hermitian block matrix with 2×2 diagonal subblocks that are multiples of $\mathbf{I}$ and with off diagonal blocks that are 2×2 Alamouti.

4.1 A Regularized Least-Squares Receiver

As in Sec. 3.1, we may proceed as follows to evaluate $\hat{\mathbf{S}}$. Let

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathcal{L}\mathcal{D}\mathcal{L}^* \tag{70}$$

and define $\mathbf{S}' = \mathcal{L}^*\mathbf{S}$. Then we can rewrite (68) as

$$\begin{aligned}\mathbf{Z} &= \left(\mathbf{\Gamma}\mathcal{L}^{-*}\right)\left(\mathcal{L}^*\mathbf{S}\right) + \mathbf{V}\\ &= \left(\mathbf{\Gamma}\mathcal{L}^{-*}\right)\mathbf{S}' + \mathbf{V}\end{aligned} \tag{71}$$

so that the least-squares estimate of $\mathbf{S}'$ is given by

$$\begin{aligned}\hat{\mathbf{S}}' &= \left(\delta\mathbf{I} + \mathcal{L}^{-1}\mathbf{\Gamma}^*\mathbf{\Gamma}\mathcal{L}^{-*}\right)^{-1}\left(\mathcal{L}^{-1}\mathbf{\Gamma}^*\right)\mathbf{Z}\\ &= \left(\delta\mathbf{I} + \mathcal{D}\right)^{-1}\left(\mathcal{L}^{-1}\mathbf{\Gamma}^*\right)\mathbf{Z}\end{aligned} \tag{72}$$

and, subsequently,

$$\hat{\mathbf{S}} = \mathcal{L}^{-*}\hat{\mathbf{S}}'$$

It remains to show how to evaluate $\mathcal{L}$ and $\mathcal{D}$. To do so we diagonalize $\mathbf{\Gamma}^*\mathbf{\Gamma}$ via successive Schur complementations as follows. We first express $\mathbf{\Gamma}^*\mathbf{\Gamma}$ as

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \left(\begin{array}{c|c}\gamma_0\mathbf{I}_2 & \mathbf{B}_0\\ \hline \mathbf{B}_0^* & \mathbf{D}_0\end{array}\right) \tag{73}$$

with $\gamma_0\mathbf{I}_2$, $\mathbf{B}_0$, $\mathbf{D}_0$ denoting the 2×2 upper–left, $2\times 2(2M-1)$ upper–right, and $2(2M-1)\times 2(2M-1)$ lower–right matrices, respectively. Now consider the block triangular factorization

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \begin{bmatrix}\mathbf{I}_2 & \mathbf{0}\\ \mathbf{B}_0^*/\gamma_0 & \mathbf{I}_{2(2M-1)}\end{bmatrix}\begin{bmatrix}\gamma_0\mathbf{I}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_0\end{bmatrix}\begin{bmatrix}\mathbf{I}_2 & \mathbf{B}_0/\gamma_0\\ \mathbf{0} & \mathbf{I}_{2(2M-1)}\end{bmatrix} \tag{74}$$

where $\mathbf{\Delta}_0 = \mathbf{D}_0 - \mathbf{B}_0^*\mathbf{B}_0/\gamma_0$. Equation (74) can be expressed more compactly as

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathcal{L}_0\begin{bmatrix}\gamma_0\mathbf{I}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_0\end{bmatrix}\mathcal{L}_0^* \tag{75}$$

Using the properties of Alamouti matrices, $\mathbf{\Delta}_0$ has a similar structure to $\mathbf{D}_0$ with diagonal subblocks that are scaled multiples of $\mathbf{I}_2$ and with off diagonal blocks that are 2×2 Alamouti. Thus we can similarly decompose $\mathbf{\Delta}_0$ as

$$\mathbf{\Delta}_0 = \left[\begin{array}{c|c}\gamma_1\mathbf{I}_2 & \mathbf{B}_1\\ \hline \mathbf{B}_1^* & \mathbf{D}_1\end{array}\right] \tag{76}$$

and factor it as

$$\mathbf{\Delta}_0 = \mathbf{L}_1\begin{bmatrix}\gamma_1\mathbf{I}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_1\end{bmatrix}\mathbf{L}_1^* \tag{77}$$

where

$$\mathbf{L}_1 = \begin{bmatrix}\mathbf{I}_2 & \mathbf{0}\\ \mathbf{B}_1^*/\gamma_1 & \mathbf{I}_{2(2M-2)}\end{bmatrix} \tag{78}$$

Then

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathcal{L}_0\left(\begin{array}{c|c}\gamma_0\mathbf{I}_2 & \mathbf{0}\\ \hline \mathbf{0} & \mathbf{L}_1\begin{bmatrix}\gamma_1\mathbf{I}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{\Delta}_1\end{bmatrix}\mathbf{L}_1^*\end{array}\right)\mathcal{L}_0^* \tag{79}$$

i.e.,

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathcal{L}_0\begin{bmatrix}\mathbf{I}_2 & \\ & \mathbf{L}_1\end{bmatrix}\begin{bmatrix}\gamma_0\mathbf{I}_2 & & \\ & \gamma_1\mathbf{I}_2 & \\ & & \mathbf{\Delta}_1\end{bmatrix}\begin{bmatrix}\mathbf{I}_2 & \\ & \mathbf{L}_1^*\end{bmatrix}\mathcal{L}_0^* \tag{80}$$

where again $\mathbf{\Delta}_1$ has a block structure with diagonal subblocks that are scaled multiples of $\mathbf{I}_2$ and with off diagonal blocks that are 2×2 Alamouti. Let

$$\mathcal{L}_1 = \mathcal{L}_0\begin{bmatrix}\mathbf{I}_2 & \\ & \mathbf{L}_1\end{bmatrix} \tag{81}$$

If we continue the decomposition in the same fashion, we end up with

$$\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathcal{L}\mathcal{D}\mathcal{L}^* \tag{82}$$

where $\mathcal{L} = \mathcal{L}_{M-1}$ is a $4M\times 4M$ lower triangular block matrix with identity diagonal subblocks and with subblocks below the diagonal that are 2×2 Alamouti matrices. Moreover, $\mathcal{D}$ is a diagonal matrix with entries

$$\mathcal{D} = \begin{bmatrix}\gamma_0\mathbf{I}_2 & & \\ & \ddots & \\ & & \gamma_{2M-1}\mathbf{I}_2\end{bmatrix} \tag{83}$$

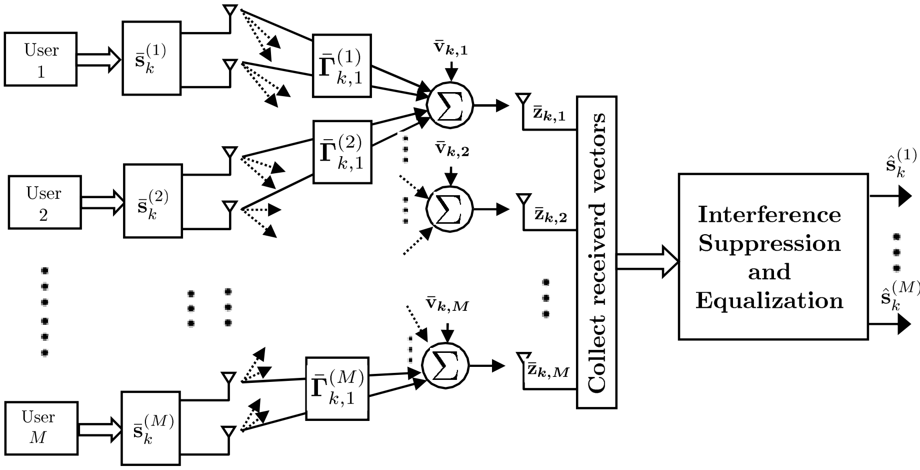

Fig. 4. OFDM multi–user transmissions with 2-transmit 1-receive antennas per user.

4.2 An Adaptive Equalizer

As was done in Sec. 3.2, by examining the structure of (69), we note that the mapping from $\mathbf{Z}$ to $\hat{\mathbf{S}}$ can be written as

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{Z} \tag{84}$$

where the equalizer coefficient matrix

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1^{(1)} & \mathbf{W}_1^{(2)} & \cdots & \mathbf{W}_1^{(M)} \\ \mathbf{W}_2^{(1)} & \mathbf{W}_2^{(2)} & \cdots & \mathbf{W}_2^{(M)} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{W}_M^{(1)} & \mathbf{W}_M^{(2)} & \cdots & \mathbf{W}_M^{(M)} \end{bmatrix} \tag{85}$$

is such that each of its entries $\mathbf{W}_j^{(l)}$ has a block structure similar to $\mathbf{W}_k$ in (45) with 2×2 Alamouti subblocks. That is,

$$\mathbf{W}_j^{(l)} = \begin{bmatrix} \mathbf{W}_{1,j}^{(l)} & \mathbf{W}_{2,j}^{(l)} \\ \mathbf{W}_{3,j}^{(l)} & \mathbf{W}_{4,j}^{(l)} \end{bmatrix} \tag{86}$$

where each $\mathbf{W}_{m,j}^{(l)}$, $m = 1,\ldots,4$ is a 2×2 Alamouti matrix with entries

$$\mathbf{W}_{m,j}^{(l)} = \begin{bmatrix} \mathbf{W}_{m,j}^{(l)}(1) & \mathbf{W}_{m,j}^{(l)}(2) \\ -\mathbf{W}_{m,j}^{*(l)}(2) & \mathbf{W}_{m,j}^{*(l)}(1) \end{bmatrix} \tag{87}$$

Thus, the estimates of the data samples from the l–th user, $\tilde{\mathbf{s}}_k^{(l)}$, can be expressed as

$$\tilde{\mathbf{s}}_k^{(l)} = \sum_{j=1}^{M} \mathbf{W}_j^{(l)} \tilde{\mathbf{z}}_{k,j} \tag{88}$$

or, equivalently,

$$\begin{bmatrix} \hat{\mathbf{s}}_2^{(l)}(k) \\ \hat{\mathbf{s}}_1^{(l)}(k) \\ \hline \hat{\mathbf{s}}_2^{*(l)}(N-k+2) \\ \hat{\mathbf{s}}_1^{*(l)}(N-k+2) \end{bmatrix} = \sum_{j=1}^{M} \left[\begin{array}{c|c} \mathbf{W}_{1,j}^{(l)} & \mathbf{W}_{2,j}^{(l)} \\ \hline \mathbf{W}_{3,j}^{(l)} & \mathbf{W}_{4,j}^{(l)} \end{array}\right] \begin{bmatrix} \mathbf{z}_{1,j}(k) \\ \mathbf{z}_{2,j}^*(k) \\ \hline \mathbf{z}_{1,j}^*(N-k+2) \\ \mathbf{z}_{2,j}(N-k+2) \end{bmatrix} \tag{89}$$

Now let i denote a block iteration index and define

$$\tilde{\mathbf{Z}}_i = \left[\, \mathbf{Z}_{1,i} \ \cdots \ \mathbf{Z}_{M,i} \,\right]$$

where each $\mathbf{Z}_{i,j}$ is given by

$$\mathbf{Z}_{j,i} = \left[\begin{array}{cc|cc} \mathbf{z}_{1,j}(k) & \mathbf{z}_{2,j}^*(k) & \mathbf{z}_{1,j}^*(N-k+2) & \mathbf{z}_{2,j}(N-k+2) \\ -\mathbf{z}_{2,j}(k) & \mathbf{z}_{1,j}^*(k) & -\mathbf{z}_{2,j}^*(N-k+2) & \mathbf{z}_{1,j}(N-k+2) \end{array}\right]$$

In other words, the entries of each $\mathbf{Z}_{j,i}$ are obtained from the received OFDM symbol at the j–th antenna at block time i, as given by (89). Let further

$$\hat{\mathbf{S}}_{1,i}^{(l)} = \begin{bmatrix} \hat{\mathbf{s}}_2^{(l)}(k) \\ -\hat{\mathbf{s}}_1^{*(l)}(k) \end{bmatrix}, \ \hat{\mathbf{S}}_{2,i}^{(l)} = \begin{bmatrix} \hat{\mathbf{s}}_2^{*(l)}(N-k+2) \\ -\hat{\mathbf{s}}_1^{(l)}(N-k+2) \end{bmatrix},$$

$$\mathcal{W}_{1,i}^{(l)} = \begin{bmatrix} \mathbf{W}_{1,1}^{(l)}(1) \\ \mathbf{W}_{1,1}^{(l)}(2) \\ \hline \vdots \\ \hline \mathbf{W}_{2,M}^{(l)}(1) \\ \mathbf{W}_{2,M}^{(l)}(2) \end{bmatrix}, \ \mathcal{W}_{2,i}^{(l)} = \begin{bmatrix} \mathbf{W}_{3,1}^{(l)}(1) \\ \mathbf{W}_{3,1}^{(l)}(2) \\ \hline \vdots \\ \hline \mathbf{W}_{4,M}^{(l)}(1) \\ \mathbf{W}_{4,M}^{(l)}(2) \end{bmatrix},$$

and assume we collect the received blocks corresponding to $2M$ consecutive OFDM symbols into a $4M \times 4M$ matrix $\mathcal{Z}_i$ as follows

$$\mathcal{Z}_i = \begin{bmatrix} \tilde{\mathbf{Z}}_i \\ \vdots \\ \tilde{\mathbf{Z}}_{i+2M-1} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{1,i} & \cdots & \mathbf{Z}_{M,i} \\ \vdots & \ddots & \vdots \\ \mathbf{Z}_{1,i+2M-1} & \cdots & \mathbf{Z}_{M,i+2M-1} \end{bmatrix}$$

Assuming that the channel is fixed over the $2M$ OFDM symbols, then we can concatenate the vectors in (89) from $2M$ OFDM symbols and write

$$\hat{\mathcal{S}}_{1,i}^{(l)} = \begin{bmatrix} \hat{\mathbf{S}}_{1,i}^{(l)} \\ \vdots \\ \hat{\mathbf{S}}_{1,i+2M-1}^{(l)} \end{bmatrix} = \mathcal{Z}_i \mathcal{W}_{i,1}^{(l)} \tag{90}$$

$$\hat{\mathcal{S}}_{2,i}^{(l)} = \begin{bmatrix} \hat{\mathbf{S}}_{2,i}^{(l)} \\ \vdots \\ \hat{\mathbf{S}}_{2,i+2M-1}^{(l)} \end{bmatrix} = \mathcal{Z}_i \mathcal{W}_{2,i}^{(l)} \tag{91}$$

where $\mathcal{Z}_i$ has Alamouti subblocks. Moreover, $\mathcal{W}_{1,i}^{(l)}$ and $\mathcal{W}_{2,i}^{(l)}$ are $4M \times 1$ vectors containing the entries of the matrix $\mathbf{W}$ in (84). Equations (90) and (91) suggest that $\mathcal{W}_{1,i}^{(l)}$ and $\mathcal{W}_{2,i}^{(l)}$ can be computed adaptively. We proceed in a manner similar to Section 3.2 to show that the STBC structure can be exploited to reduce the complexity of the RLS algorithm. We first note that, $\mathcal{Z}^*\mathcal{Z}$ is Hermitian with diagonal blocks that are multiples of the identity matrix $\mathbf{I}$ and with off diagonal blocks that are 2×2 Alamouti. At block time i, we introduce the Cholesky factorization

$$\mathcal{Z}_i^*\mathcal{Z}_i \triangleq \mathcal{L}_i\mathcal{D}_i\mathcal{L}_i^*$$

and rewrite (90) and (91) as

$$\hat{\mathcal{S}}_{1,i}^{(l)} = \left(\mathcal{Z}_i\mathcal{L}_i^{-*}\right)\left(\mathcal{L}_i^*\mathcal{W}_{1,i}^{(l)}\right) \triangleq \tilde{\mathcal{Z}}_i\tilde{\mathcal{W}}_{1,i}^{(l)} \quad (92)$$

$$\hat{\mathcal{S}}_{i,2}^{(l)} = \left(\mathcal{Z}_i\mathcal{L}_i^{-*}\right)\left(\mathcal{L}_i^*\mathcal{W}_{2,i}^{(l)}\right) \triangleq \tilde{\mathcal{Z}}_i\tilde{\mathcal{W}}_{2,i}^{(l)} \quad (93)$$

in terms of transformed variables. Now, using (92) and (93), the entries of $\mathcal{W}_1$ and $\mathcal{W}_2$ are updated at every $2M$ blocks according to the following RLS recursions:

$$\begin{aligned}\tilde{\mathcal{W}}_{1,i+2M}^{(l)} &= \tilde{\mathcal{W}}_{1,i}^{(l)} + \lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\mathbf{\Pi}_{i+2M} \\ &\quad\times\left[\mathcal{D}_{1,i+2M}^{(l)} - \tilde{\mathcal{Z}}_{i+2M}\tilde{\mathcal{W}}_{1,i}^{(l)}\right] \\ \tilde{\mathcal{W}}_{2,i+2M}^{(l)} &= \tilde{\mathcal{W}}_{2,i}^{(l)} + \lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\mathbf{\Pi}_{i+2M} \\ &\quad\times\left[\mathcal{D}_{2,i+2M}^{(l)} - \tilde{\mathcal{Z}}_{i+2M}\tilde{\mathcal{W}}_{2,i}^{(l)}\right]\end{aligned} \quad (94)$$

where

$$\mathcal{P}_{i+2M} = \lambda^{-1}\left[\mathcal{P}_i - \lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\mathbf{\Pi}_{i+1}\tilde{\mathcal{Z}}_{i+2M}\mathcal{P}_i^*\right] \quad (95)$$

and

$$\mathbf{\Pi}_{i+2M} = \left(\mathbf{I} + \lambda^{-1}\tilde{\mathcal{Z}}_{i+2M}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\right)^{-1} \quad (96)$$

where $\mathcal{D}_{1,i+2M}^{(l)}$ and $\mathcal{D}_{2,i+2M}^{(l)}$ are the desired response vectors given by

$$\mathcal{D}_{1,i+2M}^{(l)} = \begin{cases}\mathcal{S}_{1,i+2M}^{(l)} & \text{for training}\\ \check{\mathcal{S}}_{1,i+2M}^{(l)} & \text{for tracking}\end{cases}$$

and

$$\mathcal{D}_{2,i+2M}^{(l)} = \begin{cases}\mathcal{S}_{2,i+2M}^{(l)} & \text{for training}\\ \check{\mathcal{S}}_{2,i+2M}^{(l)} & \text{for tracking}\end{cases}$$

By using the matrix inversion Lemma (Sayed, 2003), we can rewrite (96) as

$$\begin{aligned}\mathbf{\Pi}_{i+2M} &= \mathbf{I} - \tilde{\mathcal{Z}}_{i+2M}\left(\lambda\mathcal{P}_i^{-1} + \tilde{\mathcal{Z}}_{i+2M}^*\tilde{\mathcal{Z}}_{i+2M}\right)^{-1}\tilde{\mathcal{Z}}_{i+2M}^* \\ &= \mathbf{I} - \tilde{\mathcal{Z}}_{i+2M}\left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2M}\right)^{-1}\tilde{\mathcal{Z}}_{i+2M}^*\end{aligned} \quad (97)$$

Substituting (97) into (95), we get

$$\begin{aligned}\mathcal{P}_{i+2M} &= \lambda^{-1}\Big[\mathcal{P}_i - \lambda^{-1}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\Big(\mathbf{I} - \tilde{\mathcal{Z}}_{i+2M} \\ &\qquad \left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2M}\right)^{-1}\tilde{\mathcal{Z}}_{i+2M}^*\Big)\tilde{\mathcal{Z}}_{i+2M}\mathcal{P}_i^*\Big] \\ &= \lambda^{-1}\mathcal{P}_i - \lambda^{-2}\mathcal{P}_i\tilde{\mathcal{Z}}_{i+2M}^*\tilde{\mathcal{Z}}_{i+2M} \\ &\qquad \left[\mathbf{I} - \left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2M}\right)^{-1}\tilde{\mathcal{Z}}_{i+2M}^*\tilde{\mathcal{Z}}_{i+2M}\right]\mathcal{P}_i^* \\ &= \lambda^{-1}\mathcal{P}_i - \lambda^{-2}\mathcal{P}_i\mathcal{D}_{i+2M} \\ &\qquad \left[\mathbf{I} - \left(\lambda\mathcal{P}_i^{-1} + \mathcal{D}_{i+2M}\right)^{-1}\mathcal{D}_{i+2M}\right]\mathcal{P}_i^*\end{aligned} \quad (98)$$

where (94) and (98) are used to update the adaptive equalizer coefficients. In a manner similar to Section 3.2, we can show that $\mathcal{P}_i$ has the following diagonal structure:

$$\mathcal{P}_i = \left[\begin{array}{cc|c|cc} p_1^{(1)}\mathbf{I} & & & & \\ & p_2^{(1)}\mathbf{I} & & & \\ \hline & & \ddots & & \\ \hline & & & p_1^{(M)}\mathbf{I} & \\ & & & & p_2^{(M)}\mathbf{I}\end{array}\right]$$

where the $p_j^{(i)}$ are scalars. The block diagram of the multi–user adaptive receiver is shown in Figure 5.

5. SIMULATION RESULTS

A typical OFDM system with space-time coding is simulated to evaluate the performance of the proposed schemes in comparison to an OFDM receiver with ideal IQ branches and to a receiver with no IQ compensation. In the simulations, the Alamouti scheme is applied to a (2×1)-OFDM system, as was described in Sec. 3. The parameters used in the simulation are as follows. OFDM symbol length of $N = 64$, cyclic prefix of $P = 16$, and channel length of $(L+1) = 4$. The channel taps corresponding to two transmit antennas are chosen independently with complex Gaussian distribution. The BER versus SNR for the proposed scheme are simulated and shown in Figures 6-8. In all figures, 'Ideal IQ' legend refers to a receiver with perfect IQ branches and perfect channel knowledge. The 'IQ Imbalance/No Compensation' refers to a receiver with IQ imbalances without any compensation scheme. Figure 6 shows the simulation results for the proposed least-squares solution. Figure 7 depicts the results when the adaptive solution is applied. The results are depicted for different constellation sizes (4QAM, 16QAM, and 64QAM) and different phase and amplitude IQ imbalances. The simulation results presented in this section are performed for a single-user case.

6. CONCLUDING REMARKS

The effect of IQ imbalances on Alamouti coded OFDM systems was studied and a framework for deriving receivers with compensation for IQ imbalances was presented. The simulation results showed that the achievable BER vs. SNR in a system assuming ideal IQ branches can be severely limited by implementation impairments such as the IQ imbalances. Combating the IQ imbalances in the digital domain has many advantages to the analog domain compensation in terms of overall cost and complexity. Compensating the IQ imbalances in the digital domain requires advanced signal processing techniques that efficiently exploit both the structure of space-times codes as well as the distortion models. Efficient receiver designs are needed to address the complexity issue.

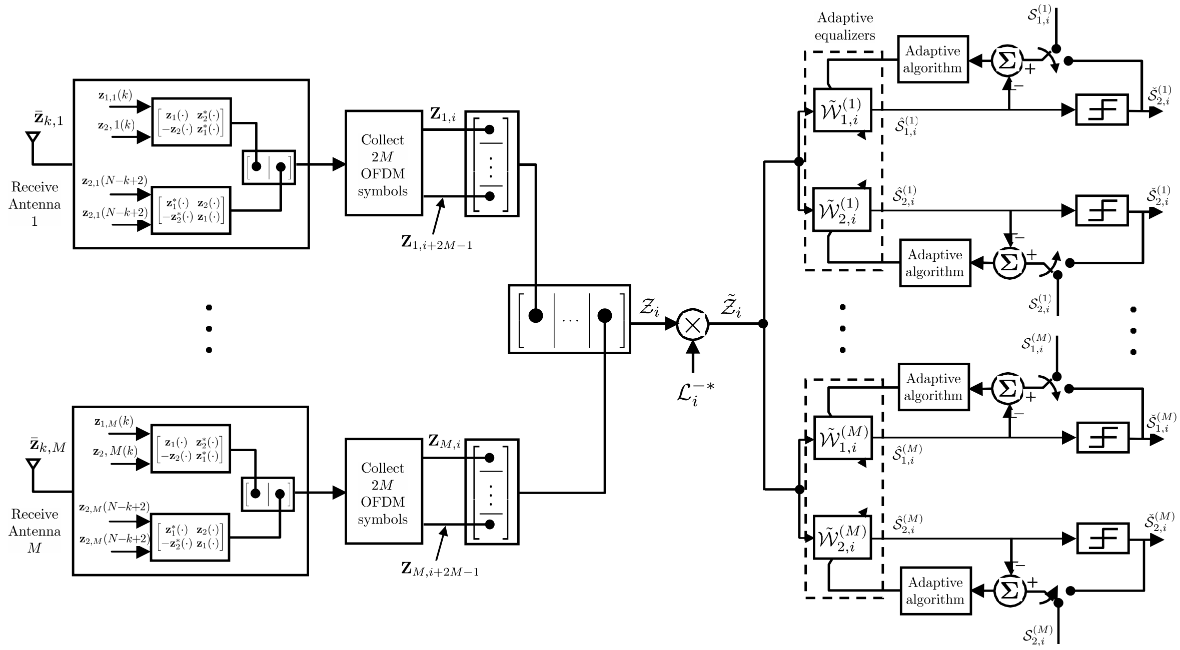

Fig. 5. An OFDM adaptive receiver for multi–user transmissions with IQ imbalances using 2-transmit 1-receive antennas per user.

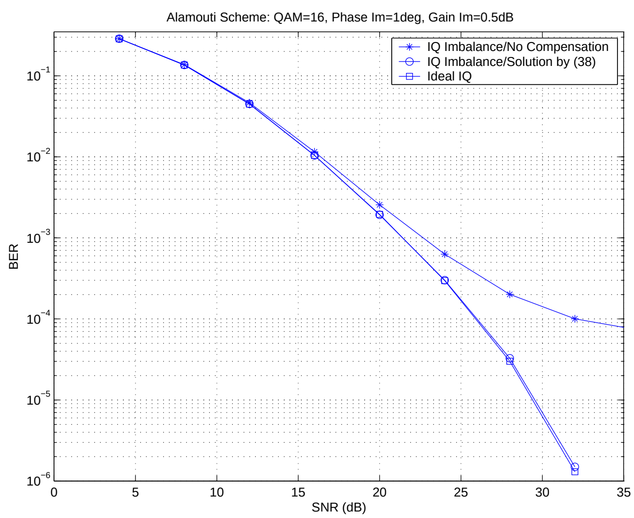

Fig. 6. BER vs. SNR for an OFDM system with Alamouti space-time coding. The simulation parameters are: phase mismatch of $\theta = 1^o$, amplitude mismatch of $\alpha = 0.5$dB–see (2), and 16QAM constellation. The proposed *least-squares* solution is simulated in this plot.

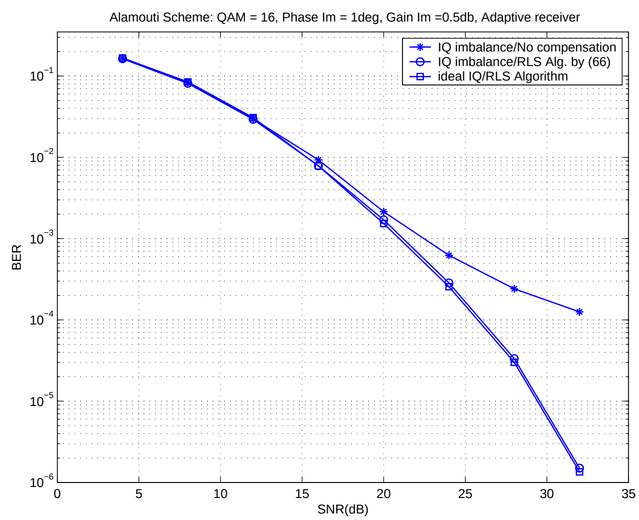

Fig. 7. BER vs. SNR for an OFDM system with Alamouti space-time coding. The simulation parameters are: phase mismatch of $\theta = 1^o$, amplitude mismatch of $\alpha = 0.5$dB–see (2), and 16QAM constellation. The proposed *adaptive (RLS)* solution is simulated in this plot.

REFERENCES

Baier, A. (1990). Quadrature mixer imbalances in digital TDMA mobile radio receivers. In: *Proc. International Zurich Seminar on Digital Communications, Electronic Circuits and Systems for Communications*. pp. 147–162.

Liu, C. L. (1998). Impacts of I/Q imbalance on QPSK-OFDM-QAM detection. **44**, 984–989.

Pelgrom, M. J. M., A. C. J. Duinmaijer and A. P. G. Welbers (1989). Matching properties of MOS transistors. **24**, 1433–1439.

Pengfei, Z., N. Thai, C. Lam, D. Gambetta, C. Soorapanth, C. Baohong, S. Hart, I. Sever, T. Bourdi, A. Tham and B. Razavi (2003). A direct conversion CMOS transceiver for IEEE 802.11a WLANs. In: *IEEE International Solid-State Circuits Conference Digest of Technical Papers*. Vol. 1. pp. 354–498.

Razavi, B. (1998). *RF Microelectronics*. Prentice Hall. NJ.

Sayed, A. H. (2003). *Fundamentals of Adaptive Filtering*. Wiley. NY.

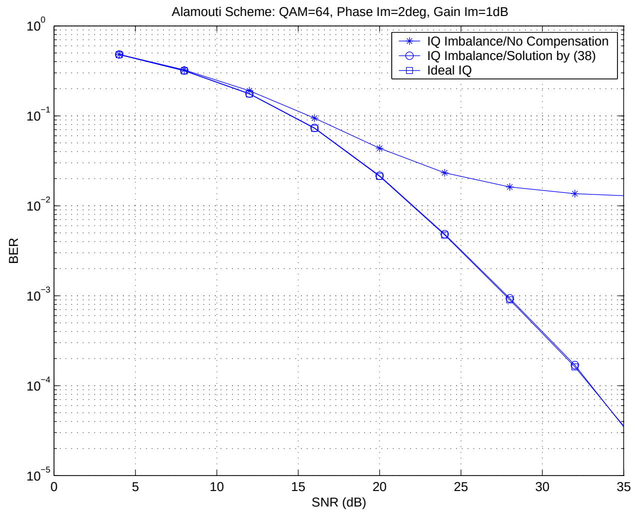

Fig. 8. BER vs. SNR for an OFDM system with Alamouti space-time coding. The simulation parameters are: phase mismatch of $\theta = 2^o$, amplitude mismatch of $\alpha = 1$dB–see (2), and 64QAM constellation.

Tarighat, A. and A. H. Sayed (2003). An optimum OFDM receiver exploiting cyclic prefix for improved data estimation. In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 4. Hong Kong. pp. 217–220.

Tarighat, A. and A. H. Sayed (2004*a*). On the baseband compensation of IQ imbalances in OFDM systems. In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 4. Montreal, Canada. pp. 1021–1024.

Tarighat, A. and A. H. Sayed (2004*b*). Space-time coding in MISO-OFDM systems with implementation impairments. In: *Proc. Third IEEE Sensor Array and Multichannel (SAM) Signal Processing Workshop*. Barcelona, Spain.

Younis, W. M. and A. H. Sayed (2004*a*). Adaptive channel estimation for MIMO space-time coded communications. In: *Proc. Third IEEE Sensor Array and Multichannel (SAM) Signal Processing Workshop*. Barcelona, Spain.

Younis, W. M. and A. H. Sayed (2004*b*). Space-time block coded receivers over frequency selective fading channels. In: *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*. Lisbon, Portugal.