Performance analysis of randomised space–time block codes for amplify-and-forward cooperative relaying

Francesco Verde
Department of Electrical Engineering and Information Technology, University Federico II, Naples I-80125, Italy
E-mail: f.verde@unina.it

Abstract: In recent studies (Verde et al.), capitalising on randomised space–time block coding (STBC), a decentralised coding method has been developed for amplify-and-forward (A&F) relays, which can closely achieve the performance of a centralised A&F linear dispersion coding (LDC) scheme and notably outperform the decentralised A&F LDC counterpart, by requiring a reduced amount of signalling and processing overhead. However, A&F relaying process introduces correlation among the noise samples at the destination, which, along with the non-Gaussian nature of the relaying channel and the fact that the relays are located in different positions, significantly complicates the performance analysis of the system. In this study, a theoretical performance analysis of the randomised STBC A&F rule is carried out in terms of average pairwise error probability to evaluate its attainable diversity order. Numerical results are provided to corroborate the analytical findings.

1 Introduction

Communication techniques exploiting temporal, spectral and/or spatial diversity allow to effectively counteract fading in wireless channels [1–6] and, thereby, meet the increasing demand for high-data rate wireless multimedia and Internet services. In cooperative networks [7–9], a given source and one or more relays cooperate among themselves in some way to reliably transmit information-bearing symbols to a certain destination, thus creating spatial diversity even when each terminal has only one antenna (so-called ‘cooperative diversity’). Potential application areas of cooperative diversity include ad-hoc/multihop local wireless networks [10–12] and modern cellular systems [13, 14].

Even though some works on cooperative networks have focused on a single-relay channel (e.g. [9]), to further increase the spatial degrees-of-freedom, a system with multiple relay nodes is considered hereinafter, which involve space–time block coding (STBC) [15–25]. In such cases, a two-phase protocol is involved: in Phase I, the source broadcasts the information to the relays and, in Phase II, the relays concurrently transmit a function of the source symbols to the destination such that the signal at the receiver appears as a space–time block code. From the information theory viewpoint, it is shown in [15] that, by assuming full coordination among the source, the relays and the destination, distributed STBC (DSTBC) rules can ensure higher bandwidth efficiency than repetition-based cooperative strategies [26–28], by achieving full diversity in the number of cooperative terminals. Moreover, contrary to beamforming techniques [8, 29], cooperative diversity can be harvested by means of DSTBC without requiring full channel state information (CSI) at both the relays and the destination (full CSI means that the channel coefficients from $S$ to all the relays and from all the relays to $D$ are known), and without involving considerable modifications to existing radio frequency front ends. Finally, unlike methods based on opportunistic selection of the best relay from a set of multiple potential relay nodes [30], cooperative diversity can be exploited through DSTBC without assuming that the forward and backward channels between each relay and destination are the same because of the reciprocity theorem, which is not true if transmissions occur on different frequency bands and/or different coherence intervals.

The works [16–19] have been devoted to the design of practical DSTBC methods for decode-and-forward (D&F) relays. In particular, similarly to [15], the presence of a centralised control terminal is required in [16] to preliminarily assign the codes to the relays; on the other hand, the scheme proposed in [18, 19, 31–33], referred to as randomised DSTBC (R-DSTBC), is decentralised in the sense that each relay chooses the code in a random and independent fashion, thus reducing the cooperation overhead. Moreover, no feedback is required from the relays to the destination in [18, 19], since decoding at the receiver can be accomplished by completely ignoring what the actual number of cooperating relays and the randomisation vectors are.

Amplify-and-forward (A&F) relaying is another popular cooperation protocol, where the relays simply scale the signal received from the source and then retransmit the resulting signal to the destination. Since this operation mode requires no decoding at relay nodes, it involves a less processing burden with respect to D&F relaying and, thus, it is well suited to systems with simple relay units that have stringent power constraints. Therefore we devote out attention to A&F relays in this paper. A DSTBC rule
for A&F cooperative networks has been proposed in [17], which is referred to as linear dispersion coding (LDC), where the signal sent by each relay in Phase II is a linear transformation of the received data block and/or its complex conjugate version by means of certain relaying matrices. Generalisations of [17] to the cases of multiple-antenna nodes and different orthogonal and quasi-orthogonal code designs can be found in [34, 35], respectively. The diversity properties of the LDC scheme have been recently studied in [20] for large values of the transmission block duration with reference to maximum likelihood (ML) and linear minimum-mean-squared error decoding. If the relaying matrices in [17, 34, 35] are not optimised, for example, they are chosen randomly at the relays, no control terminal is required to coordinate the relaying process. In this case, the relaying matrices can be absorbed into the channel matrix and the destination can directly estimate the equivalent channel for coherent decoding; however, a disadvantage of such an approach is that a large number of channel coefficients has to be estimated. An alternative way is that each relay transmits to the destination the random seed used to generate the transformation of the received data block and/or its decoding; however, a disadvantage of such an approach is that a large number of channel coefficients. In this paper, a theoretical diversity analysis is presented by showing that, when the transmit power is sufficiently large, the R-DSTBC A&F scheme achieves the same diversity order of its D&F counterpart [18]. A numerical performance analysis is also provided evidencing that the R-DSTBC A&F protocol achieves the same performance of a centralised LDC scheme, which involves a considerably higher cooperation cost in terms of coordination signalling, by outperforming the LDC method when the relaying matrices are randomly chosen at the relays.

The paper is organised as follows. Section 2 briefly reviews the R-DSTBC A&F scheme proposed in [1, 2] and we analyse in Section 3 the average PEP (APEP) of such a cooperative protocol. Section 4 provides numerical results in terms of average bit-error-rate (ABER), by corroborating the results of the theoretical analysis carried out and comparing the R-DSTBC A&F method with both centralised and decentralised versions of the LDC A&F scheme [17, 35]. Section 5 includes concluding remarks.

1.1 Notations and preliminaries

The fields of complex, real and integer numbers are denoted with $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{Z}$, respectively; matrices [vectors] are denoted with upper [lower] case boldface letters (e.g. $A$ or $a$); the field of $m \times n$ complex [real] matrices is denoted as $\mathbb{C}^{m \times n}$ [$\mathbb{R}^{m \times n}$], with $\mathbb{C}$ [$\mathbb{R}$] used as a shorthand for $\mathbb{C}^{1 \times 1}$ [$\mathbb{R}^{1 \times 1}$]; $|a|_2$, $\angle a$, $\Re\{a\}$ and $\Im\{a\}$ denote the magnitude, the phase, the real part and the imaginary part of $a \in \mathbb{C}$, respectively; the superscripts $*$, $T$, $H$ and $-$1 denote the conjugate, the transpose, the Hermitian (conjugate transpose) and the inverse of a matrix, respectively; the cardinality of a set $A$ is denoted by $|A|$; let $a = [a_1, a_2, \ldots, a_n]^T \in \mathbb{C}^n$ and $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$, the $i$th entry of the vector $a_i^{(\mathcal{I})} = \{a_i^*\}_i$, if $i \in \mathcal{I}$, otherwise $[a_i^T]_i = \{a_i^\dagger\}_i$; $0_m \in \mathbb{R}^m$, $O_{m \times n} \in \mathbb{R}^{m \times n}$ and $I_m \in \mathbb{R}^{m \times m}$ denote the null vector, the null matrix and the identity matrix, respectively; rank($A$) and det($A$) denote the rank and the determinant of the matrix $A$, respectively; $\|a\|$ is the Euclidean norm of $a \in \mathbb{C}^n$; the eigenvalues of $A \in \mathbb{C}^{m \times m}$ are denoted as $\mu_i(A)$, for $i \in \{1, 2, \ldots, m\}$, and, when they are real, they are ordered as $\mu_1(A) \geq \mu_2(A) \geq \cdots \geq \mu_m(A)$; $P(A) \geq \Psi$ denotes the probability that an event $A$ occurs and $P(A|B)$ is the conditional probability of $A$ given an event $B$; the operator $E[\cdot]$ denotes ensemble averaging and, specifically, $E[\cdot|\mathcal{I}]$ is the conditional mean given the random matrix $\mathcal{Y} \in \mathbb{C}^{n \times d}$; a circular symmetric complex Gaussian random vector $x \in \mathbb{C}^m$ with mean $m \in \mathbb{C}$ and covariance matrix $K \in \mathbb{C}^{m \times m}$ is denoted as $x \sim \mathcal{CN}(m, K)$; in the high signal-to-noise ratio (SNR) regime, the average symbol error probability (ASEP) for a digital communication system over a fading channel usually behaves as $\text{ASEP}(\gamma) \approx (G_\gamma \gamma)^{-\kappa}$ [6], where $\gamma$ denotes the average SNR without fading, $G_\gamma$ is the coding gain, and

$$G_d \triangleq \lim_{\gamma \to +\infty} -\log \frac{\text{ASEP}(\gamma)}{\log \gamma}$$

is the (asymptotic) diversity order.

2 R-DSTBC A&F cooperative diversity protocol

The considered wireless network (see Fig. 1) is composed of $N$ randomly and independently placed relay nodes $\left\{R_i^{(N)}\right\}_{i=1}^N$, one source station (S) and one destination terminal (D), each one employing a single transmit/receive antenna. For $i \in \{1, 2, \ldots, N\}$, let $f_i$ and $\eta_i$ denote the channel gain and the distance, respectively, between $S$ and $R_i$, whereas $g_i$ and $v_i$ denote the channel gain and the distance, respectively, between $R_i$ and $D$; the coherence time of $f_i$ is larger than $K$ symbol intervals, whereas the coherence time of $g_i$ is larger than $P$ symbol intervals.

The channel vectors

$$f \triangleq [f_1, f_2, \ldots, f_N]^T \in \mathbb{C}^N$$

and

$$g \triangleq [g_1, g_2, \ldots, g_N]^T \in \mathbb{C}^N$$

are statistically independent of each other and are modelled
as \( f \sim \mathcal{CN}(\mathbf{0}_N, \Sigma_f) \), with

\[
\Sigma_f \triangleq \text{diag}(\sigma_{f_1}^2, \sigma_{f_2}^2, \ldots, \sigma_{f_K}^2) \in \mathbb{R}^{N \times N}
\]

and \( g \sim \mathcal{CN}(\mathbf{0}_N, \Sigma_g) \), with

\[
\Sigma_g \triangleq \text{diag}(\sigma_{g_1}^2, \sigma_{g_2}^2, \ldots, \sigma_{g_N}^2) \in \mathbb{R}^{N \times N}
\]

where \( \sigma_{f_i}^2 \triangleq \eta_i^{-\rho} \) and \( \sigma_{g_i}^2 \triangleq \eta_i^{-\rho} \), with \( \rho \) being the path-loss exponent. It is assumed that \( f_i \) is known at the \( i \)-th relay. Indeed, the channel parameter \( f_i \) can be acquired at \( R_i \) by using well-known point-to-point training-based estimators \([15]\). It should be observed that no knowledge is required at \( R_i \) about the channel coefficient \( g_i \) over the \( R_i \rightarrow D \) link. Following the related literature, for example, \([15–17, 34, 35]\), perfect synchronisation is assumed at the symbol level among all the terminals. Such an assumption is made only for the sake of simplicity and can be relaxed. Indeed, if the signals transmitted by the relays do not arrive at \( D \) within a very small gap (compared to the symbol period), the cooperative link behaves as a frequency-selective channel \([19]\); in this case, one can use time-reverse space-time code. In the second phase, as done in standard STBC \([21–25]\), the vector \( z_i \) is mapped onto a STBC matrix \( \mathcal{C}(z_i) \in \mathbb{C}^{W \times K} \), where \( P \geq K \) is the block length and \( L \) denotes the number of virtual antennas in the underlying space–time code. In the case of a multiple-input system with colocated antennas, the parameter \( L \) is the effective number of transmitting antennas; in the distributed framework at hand, it is a design parameter that does not have a physical meaning, since there is no relationship between \( L \) and the number of cooperating relays \( N \). The considered space–time block codes satisfy the following three basic properties

\[ \mathcal{C}(\gamma_1 + \gamma_2) = \mathcal{C}(\gamma_1) + \mathcal{C}(\gamma_2), \forall \gamma_1, \gamma_2 \in \mathbb{C}^K \]

\[ \mathcal{C}(\mu \gamma) = \mu \mathcal{C}(\gamma), \forall \gamma \in \mathbb{C}^K \text{ and } \forall \mu \in \mathbb{R} \]

\[ \mathcal{E}(\mathcal{C}(\gamma) \mathcal{C}(\gamma')^\dagger) = \mathcal{E}[\|\gamma\|^2 J], \forall \gamma \in \mathbb{C}^K - \{0_K\}, \text{ with } J \in \mathbb{R}^{L \times L} \text{ being a non-singular matrix.} \]

By virtue of properties (p1) and (p2), and setting \(|\theta| = 1 \) and \( \lambda \theta = -\lambda \theta \), it comes from (2) that

\[ \mathcal{C}(z_i) = \lambda_i |f_i| \mathcal{C}(a) + \lambda_i \mathcal{C}(n_i) \]

where since the entries of \( n_i \) are circularly symmetric complex
Gaussian random variables, the vector \( \bar{n}_i = n_i e^{-jD_i} \in \mathbb{C}^K \) has the same probability distribution of \( n_i \), that is, \( \bar{n}_i \sim \mathcal{CN}(0_K, \sigma_n^2 I_K) \).

As a last step, let \( r_i \in \mathbb{C}^L \) be a random vector containing the linear combination coefficients for the \( i \)th node, the code \( x_i \in \mathbb{C}^p \) transmitted by the \( i \)th relay is given by \( x_i = \mathcal{C}(\xi_i, r_i) \). For the time being, we do not make any specific assumption on the statistical model of the randomisation matrix \( \mathcal{R} = [r_1; r_2; \ldots; r_N] \in \mathbb{C}^{L \times N} \), which collects the randomisation vectors used by the relays; we highlight only that, unlike [36–41], the R-DSTBC A&F coding rule is completely decentralised since the \( i \)th relay chooses \( r_i \) locally from a given distribution, which does not depend on \( i \).

The value of the parameter \( \lambda_i \) is determined by the power constraint at the \( i \)th relay, which is given by \( P_{r_i} = \mathbb{E}(\|x_i\|^2) = \text{Tr} \mathbb{E} \mathbb{E} \), where, by using property (p3) and scaling \( r_i \) such that \( \mathbb{E}(\|r_i\|^2) = 1 \), one has \( \mathbb{E}(\|x_i\|^2) = \mathbb{E}(\|z_i\|^2) = \lambda_i^2 K (\sigma^2 + \sigma_n^2) \). It results that

\[
\lambda_i = \sqrt{\frac{\sigma^2}{R_{\text{code}} (\sigma^2 + \sigma_n^2)}}
\]

which keeps the power constraint \( P_{r_i} = \sigma^2 r_i \) from the long term point of view, where the code rate \( R_{\text{code}} = K/P \leq 1 \) reflects the bandwidth efficiency of the coding scheme. It is worth noting that, let \( P_{\text{tot}} \) be the total power used in the network for transmitting a source symbol, it holds that

\[
P_{\text{tot}} = \sigma^2 + \frac{1}{R_{\text{code}}} \sum_{i=1}^{N} \sigma_{r_i}^2
\]

The baseband equivalent discrete-time signal received at \( D \) assumes the form

\[
y_d = \sum_{i=1}^{N} g_i \lambda_i f_i [\mathcal{C}(a) h] r_i + w_d = [\mathcal{C}(a) h] + w_d
\]

where \( h = \mathcal{R}^t \tilde{g} f \in \mathbb{C}^L \) is the overall \( S \) \( \rightarrow \) relays \( \rightarrow D \) channel, with \( \lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) \in \mathbb{R}^{N \times N} \) and \( \tilde{g} = \text{diag}(\tilde{g}) \in \mathbb{C}^{N \times N} \) and \( \tilde{g} = [g_1 e^{-jD_1}, g_2 e^{-jD_2}, \ldots, g_N e^{-jD_N}]^T \in \mathbb{C}^N \) and

\[
w_d = \sum_{i=1}^{N} g_i \lambda_i C(\tilde{n}_i) r_i + n_d
\]

denotes the effective noise at the receiver, with \( n_d \sim \mathcal{CN}(0_P, \sigma_n^2 I_p) \) denoting AWGN, which is statistically independent of \( a, f, g \) and \( \{n_i\}_{i=1}^{N} \). Since the entries of \( g \) are circularly symmetric complex Gaussian random variables, the vector \( \tilde{g} \) has the same probability distribution of \( g \), that is, \( \tilde{g} \sim \mathcal{CN}(0_N, \Sigma_g) \). Moreover, it is seen from (6) that, conditioned on \( \mathcal{R} \) and \( g \), \( w_d \sim \mathcal{CN}(0_P, K) \), where \( K = \mathbb{E}[w_d w_d^H] / [\mathcal{R}, g] \in \mathbb{C}^{p \times p} \) is given by

\[
K = \sum_{i=1}^{N} |g_i|^2 \lambda_i^2 \mathbb{E}[\mathcal{C}(\tilde{n}_i) r_i h_i^H \mathcal{C}(\tilde{n}_i)] + \sigma_n^2 I_p
\]

which shows that the noise correlation matrix \( K \) is not diagonal in general and, thus, the effective noise vector \( w_d \) might be additive correlated Gaussian noise (ACGN), which is an inherent shortcoming of many A&F relaying protocols.

The matrix \( K \) and the channel vector \( h \) can be directly estimated at \( D \) by allowing each data transmission in Phase II be preceded by a training period: within this time interval, all the relays send a symbol sequence \( a \) known to \( D \), which can be utilised to acquire \( h \); the noise-only correlation matrix \( K \) can then be estimated by subtracting the signal correlation matrix [depending on \( \mathcal{C}(a) \) and \( h \)] from the correlation matrix of the received vector \( y_d \) measured during the training period. The randomisation vectors \( \{r_i\}_{i=1}^{N} \) used during the training phase will be maintained in the subsequent data transmission. In the forthcoming analysis, both \( K \) and \( h \) are assumed to be perfectly known.

3 Complexity and performance analysis of the R-DSTBC A&F scheme

Optimum ML decoding from (6) corresponds [42] to the decision rule

\[
\hat{a}_{\text{opt}} = \arg \min_{a \in \mathbb{A}} [y_d - \mathcal{C}(a) h] \mathbb{H} [y_d - \mathcal{C}(a) h]
\]

\[
= \arg \min_{a \in \mathbb{A}} [\mathbb{H} \mathcal{C}(a) K^{-1} \mathcal{C}(a) h - 2 \Re \{\mathbb{H} \mathcal{C}(a) K^{-1} y_d\}]
\]

whose complexity is exponential in the number of possible constellation points. Alternatively, sphere decoding can be used with only polynomial complexity, for sufficiently high SNR [43]. It is stated in [2] (without a proof) that linear ML decoding complexity is possible at \( D \) if orthogonal full-rate STBC rules are used at the relays. We provide a formal proof of this property:

**Lemma 1:** If a complex orthogonal STBC rule \( \mathcal{C}(\cdot) \) is used at the relays, the R-DSTBC A&F rule is single-symbol ML decodable for \( R_{\text{code}} = 1 \). If the source symbol vector is real-valued, that is, \( a \in \mathbb{R}^K \) and, thus, a real orthogonal STBC design \( \mathcal{C}(\cdot) \) is used at the relays, the R-DSTBC A&F rule is single-symbol ML decodable for \( R_{\text{code}} = 1 \), provided that \( \mathcal{R} \) is real-valued, too.

**Proof:** See Appendix 1.

Although rate-one real orthogonal STBC matrices can be designed for any value of \( L \) [23], complex orthogonal STBC matrices with \( R_{\text{code}} = 1 \) exist only when \( L = 2 \), whereas, for \( L \geq 3 \), the noise correlation matrix is not a scaled identity matrix and, thus, the R-DSTBC A&F scheme is not single-symbol ML decodable. However, in this case, when decoding complexity is one of the major concerns, a suboptimal detector can be designed as if the noise correlation matrix were a scaled identity, that is, one
can resort to
\[
\tilde{a}_{\text{subopt}} = \arg\min_{a \in \mathcal{A}} \| y_d - C(a)h \|^2 \\
= \arg\min_{a \in \mathcal{A}} \left[ h^H C_{11}(a) C(a) h - 2R \left( h^H C_{11}(a) y_d \right) \right]
\] (10)

which is always single-symbol decodable, provided that \( C_{11}(a) C(a) = \|a\|^2 I_L \), and, according to Lemma 1, turns out to be optimum (i.e. \( \tilde{a}_{\text{subopt}} = a_{\text{opt}} \)) when \( L = 2 \). We will show that the suboptimum detector (10) achieves the same diversity order of the optimum one (9), by paying (see Section 4) only a negligible coding-gain penalty for complex orthogonal STBC rules with \( L \geq 3 \).

Another alternative is to use full-rate square (i.e. \( P = L = K \)) complex quasi-orthogonal codes [24, 25, 35], which are pair wise ML decodable (i.e. decoding pairs of symbols independently is possible). With reference to some known complex quasi-orthogonal codes, it is shown in Appendix 2 that the R-DSTBC A&F rule is pairwise ML decodable, provided that entries of \( \mathcal{R} \) are subject to a non-restrictive constraint depending on the particular structure of the quasi-orthogonal code.

Although at first sight the signal model (6) might appear similar to that reported in \([18, (3)]\), there are two key difference. First, even if a Rayleigh fading model were assumed for all the underlying channels, given the randomisation matrix \( \mathcal{R} \), the overall \( S \rightarrow relays \rightarrow D \) channel vector \( h \) is not Gaussian; on the contrary, under the assumption that the transmission from \( S \) to the relays is error-free, the relays \( \rightarrow D \) channel vector \( h \) in [18] has a Gaussian distribution for a given realisation of \( \mathcal{R} \). In D&K relaying, only the nodes that reliably decode the source message can retransmit the received data towards \( D \), which can be accomplished by using cyclic redundancy check coding at the relays [16] or by setting an SNR threshold at the relays that ensures negligible symbol error probability [44], and a node only forwards the source data if the received SNR is larger than the fixed threshold. Second, as a consequence of the randomisation of \( C(e) \) and the noise propagation from the relays to \( D \), the statistics of the effective noise \( w_2 \) in (6) depend on the relays \( \rightarrow D \) channel vector \( g \) and on the randomisation matrix \( \mathcal{R} \). This implies that, given both \( g \) and \( \mathcal{R} \), the vector \( w_2 \) is in general ACGN; instead, under the assumption that the transmission from \( S \) to the relays is error-free, the noise term at \( D \) is AWGN in [18]. Moreover, diversity analyses of A&F schemes are typically carried out (see, e.g. [17, 34, 35]) by assuming that the average path losses do not depend on \( i \), that is, \( \sigma_{\text{opt}}^2 = \sigma_r^2 \) and \( \sigma_{\text{opt}}^2 = \sigma_n^2 \), which is not realistic when the relays are located in different positions.

Without considering any specific space-time block code, we assume that, in addition to (p1)–(p3), the STBC matrix satisfies the rank criterion [22], which states that:

\[ (p4) \quad \text{for any pair } C_k \triangleq C(\alpha_k) \text{ and } C_{k,\ell} \triangleq C(\alpha_k, \alpha_\ell), \text{ where } \alpha_k, \alpha_\ell \in \mathcal{A} \text{ with } k \neq \ell \in \{1, 2, \ldots, Q\}, \text{ the matrix } C_{k,\ell} \triangleq C_k - C_\ell \text{ is full rank, that is, } \text{rank}(C_{k,\ell}) = \min(P, L). \]

For the sake of simplicity, we assume hereinafter that the noise variances at the relays and at \( D \) are all equal, that is, \( \sigma_{n_1}^2 = \sigma_{n_2}^2 = \cdots = \sigma_{n_K}^2 = \sigma_n^2 \triangleq \sigma_n^2 \). In the sequel, without loss of generality, we set \( \mathcal{P}_{\text{tot}} = 1 \) and, hence, the average SNR per symbol is defined as \( \gamma \triangleq 1/\sigma_n^2 \). Moreover, for the time being, we assume that \( \mathcal{R} \) is fixed with rank \( R \triangleq \text{rank}(\mathcal{R}) \leq \min(L, N) \).

In order to obtain the ASEP at the output of the ML detector (9), given \( \mathcal{R} \), for sufficiently high (but finite) SNR values, one can typically resort to the union bound [42] that depends on \( \text{APEP}_{\text{subopt}}(\mathcal{R}) \triangleq E \left[ P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h|\mathcal{R}) \right] \), with \( k \neq \ell \in \{1, 2, \ldots, Q\} \), where \( P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h) \) is the probability, conditioned on \( h \), that \( \alpha_k \) is detected at \( D \) when \( \alpha_k \) was transmitted, given by

\[
P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h) = Q \left( \sqrt{\frac{h^H C_{k,\ell} K^{-1} C_{k,\ell} h}{2}} \right)
\] (11)

Similarly, the ASEPs at the output of the suboptimum detector (10), given \( \mathcal{R} \), can be upper bounded by calculating \( \text{APEP}_{\text{subopt}}(\mathcal{R}) \triangleq E \left[ P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h|\mathcal{R}) \right] \), with \( k \neq \ell \in \{1, 2, \ldots, Q\} \), where the corresponding PEP, conditioned on \( h \), can be expressed as

\[
P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h) = Q \left( \frac{h^H C_{k,\ell} C_{\ell,\ell} h}{2 h^H C_{k,\ell} K C_{k,\ell} h} \right)
\] (12)

At this point, we provide the following lemma:

**Lemma 2**: The PEPs (11) and (12) admit the same upper bound that is given by

\[
P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h) \leq P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h) \leq P(\{\alpha_k \rightarrow \alpha_\ell\}_{\text{subopt}}|h)
\] (13)

**Proof**: See Appendix 3. \(\square\)

Based on the union bound and (13), we can assert that, if

\[
\lim_{\gamma \rightarrow +\infty} \frac{-\log \left[ \text{APEP}_{\text{det}}(\mathcal{R}) \right]}{\log (\gamma)} = G_d
\] (14)

for every \( k \neq \ell \in \{1, 2, \ldots, Q\} \) with \( \text{APEP}_{\text{det}}(\mathcal{R}) \triangleq E \left[ P(\{\alpha_k \rightarrow \alpha_\ell\}|h|\mathcal{R}) \right] \), then the ASEPs at the output of the detectors (9) and (10) exhibit in the high SNR regime the same slope (in dB/dB scale) against \( \gamma \) (this is true since the transmission rate is kept constant with respect to \( \gamma \)), that is, they ensure the same diversity order \( G_d \). Strictly speaking, Lemma 2 confirms that, compared to the ML detector (9), the performance penalty of its suboptimal counterpart (10) is only in terms of coding gain.

At this point, we provide an upper bound on \( \text{APEP}_{\text{det}}(\mathcal{R}) \) in the following theorem:

**Theorem 1** *(deterministic \( \mathcal{R} \))*: Let us assume that \( P \geq L \) and the diagonal entries of \( \Sigma_f \hat{A}^{-1} \hat{G}^+ G^+ \hat{A}^{-1} \) are arranged in increasing order, that is, \( \sigma_{\ell,\ell}^2 \lambda_l^2 |g_l|^2 \leq \sigma_{\ell,\ell}^2 \lambda_{l+1}^2 |g_{l+1}|^2 \leq \cdots \leq \sigma_{\ell,\ell}^2 \lambda_N^2 |g_N|^2 \).

\[
\lim_{N \rightarrow +\infty} \sum_{i=1}^{N} \frac{\sigma_{\ell,\ell}^2 \lambda_i^4 |f_r|^2}{\ell^2} < +\infty
\] (15)
then, in the high-SNR regime and for sufficiently large finite values of $N$, it results that
\[
A_{\text{PEP}}_{k,l}(\mathcal{R}) \propto \gamma^{-k} [\ln (\gamma)]^{k} Y_{k,l}(\mathcal{R})
\] (16)
where
\[
Y_{k,l}(\mathcal{R}) \triangleq \frac{4^{R} \left( K \sum_{i=1}^{N} \sigma_{\phi_{i}}^{2} \lambda_{i} \lambda_{i}^{l} 1_{J_{r}} \right)^{R}} {2^{L} \prod_{i=1}^{L} \mu_{i}(\mathcal{R}^{H})} \left[ \prod_{i=L-R+1}^{L} \mu_{i}(C_{k,l}^{H} \mathcal{C}_{k,l}) \right]
\] (17)

\[\sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \triangleq \left( \frac{\sigma_{\phi_{i}}^{2}} {R_{\text{code}} \sigma_{\phi_{i}}^{2} \sigma_{\phi_{i}}^{2}} \right) \text{ for } i \in \{1, 2, \ldots, N\}
\]

**Proof:** See Appendix 4.

Before proceeding further, a comment on condition (15) is in order, which is a consequence of the Kolmogorov criterion of the strong law of large number [52]. A sufficient condition for fulfillment of (15) is that $\sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \phi_{i}^{l} J_{r}$ is bounded from above, that is, $\sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \phi_{i}^{l} J_{r} \leq Z < +\infty$; in this case, the sum in (15) is upper bounded by $(\pi Z)^{6} / 6$. It follows from (14) and (16) that, given $\mathcal{R}$, the diversity order $\gamma_{d}$ of the D-STBC A&F scheme is equal to $R$, that is, the rank of the matrix $\mathcal{R}$, whose maximum value is $\min(L, N)$. To additionally provide an upper bound on the average PEP performance of the diversity scheme at hand when the vectors $r_{1}, r_{2}, \ldots, r_{N}$ are randomly and independently chosen by the relays, we now evaluate the ensemble average of (16) with respect to $\mathcal{R}$. The result is summarised in the following lemma:

**Lemma 3 (random $\mathcal{R}$):** Let us assume that the matrix $\mathcal{R}$ has rank $R$ and satisfies (15) with probability one. If the random vectors $r_{1}, r_{2}, \ldots, r_{N}$ are statistically independent and
\[
\lim_{N \to +\infty} \sum_{i=1}^{N} \sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \left( E(\phi_{i}^{l} J_{r})^{2} \right) - 1 < +\infty
\] (18)
then, in the high-SNR regime and for sufficiently large finite values of $N$, it results that
\[
A_{\text{PEP}}_{k,l}(\mathcal{R}) \propto \gamma^{-k} [\ln (\gamma)]^{k} \Gamma_{k,l} E \left[ \prod_{r=1}^{R} \mu_{r}(\mathcal{R}^{H}) \right]
\] (19)
with
\[
\Gamma_{k,l} \triangleq \frac{4^{R} \left( K \sum_{i=1}^{N} \sigma_{\phi_{i}}^{2} \lambda_{i}^{l} + 1 \right)^{R}} {2^{L} \prod_{i=1}^{L} \sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \mu_{i}(\mathcal{C}_{k,l}^{H} \mathcal{C}_{k,l})}
\] (20)

\[\lambda_{i} \text{ has been already defined in Theorem 1.}
\]

**Proof:** See Appendix 5.

A sufficient condition for fulfillment of (18) is that $\sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \phi_{i}^{l} J_{r}^{2} \left( \frac{1}{\mu_{r}(\mathcal{R}^{H})} \right)^{R} - 1$ is bounded from above, that is,
\[
\sigma_{\phi_{i}}^{2} \lambda_{i}^{l} \phi_{i}^{l} J_{r}^{2} \left( \frac{1}{\mu_{r}(\mathcal{R}^{H})} \right)^{R} - 1 \leq Z < +\infty
\]
in this case, the sum in (18) is upper bounded by $(\pi Z)^{2} / 6$. The result of Lemma 3 can be further on explicated by assuming a specific randomisation rule, such as real/complex Gaussian, uniform phase, real/complex spherical distribution and antenna selection. We directly refer to [18] for details on how to explicitly calculate $E \left[ \prod_{r=1}^{R} \mu_{r}(\mathcal{R}^{H}) \right]$ in such cases.

(19) shows that the probability distribution of the matrix $\mathcal{R}$ impacts only on the coding gain. Thus, the diversity order of the R-DSTBC A&F scheme is equal to $\min(L, N)$ and full diversity in the number of cooperating nodes can be achieved when $L \geq N$. It is noteworthy that, for equal values of $L$ and $N$, the R-DSTBC A&F rule achieves the same asymptotic diversity order of its D&F counterpart [18], provided that correct decisions are made at the relays. However, in the A&F case at hand, because of the presence of the factor $[\ln(\gamma)]^{R}$ in (16), the convergence of the finite-SNR diversity order to its asymptotic value $R$ is slower than that of the D&F protocol [18]. Moreover, it should be observed that the finite-SNR diversity order of A&F LDC protocols is given by $\min(N, P) \left[ 1 - \log(\log(\gamma)) / \log(\gamma) \right]$ [34, 35] and, therefore full diversity is asymptotically (i.e. as $\gamma \to +\infty$) achieved when $P \geq N$. However, with respect to the R-DSTBC A&F method, an higher signaling and processing overhead is required for these schemes (see also Section 4).

### 4 Numerical examples

In this section, we present a Monte Carlo performance analysis of the R-DSTBC A&F protocol [referred to as ‘randomised A&F (MF)’], in terms of ABER for different STBC matrices $\mathcal{C}(\cdot)$ and in comparison with other DSTBC A&F schemes.

In all the examples, we adopt the following simulation setting. According to the assumptions made in Section 3, we set $\sigma_{\phi}^{2} = (1/2)$ and $\sigma_{\phi}^{2} = (R_{\text{code}} / (2N))$, for $i \in \{1, 2, \ldots, N\}$, which implies that $\gamma_{\text{tot}} = 1$; thus, all the ABER performances are reported as a function of the average SNR per symbol $\gamma$. QPSK signalling is used in Examples 4-A and 4-C, that is, $\tau_{k} = \{ \pm (1/\sqrt{2}) \pm (j/\sqrt{2}) \}$ for $k \in \{1, 2, \ldots, K\}$, whereas BPSK modulation is adopted in Example 4-B, that is, $\tau_{k} = \{ \pm 1 \}$ for $k \in \{1, 2, \ldots, K\}$. In the xOy Cartesian coordinate system, the $N$ relays are distributed randomly and independently in a circle with center at origin $O$ and radius $\delta$, where $S$ and $D$ are located at the positions $(\delta, 0)$ and $(\delta, 0)$, respectively. The radius $\delta$ is equal to $5$ m and the path-loss exponent is $\rho = 2$. As competitive alternatives, we consider: (i) the ZF variant [1] of the ‘randomised A&F (MF)’ scheme with $\theta_{z} = (1/3)$ in (2) [referred to as ‘randomised A&F (ZF)’]; (ii) both centralised and randomised versions of the A&F LDC rule [17, 35] (referred to as ‘centralised A&F LDC’ and ‘randomised A&F LDC’, respectively), for which the signal transmitted by the $n$th relay is given by
\[
x_{r} = \lambda_{n} (\theta_{r} A y_{r} + \theta_{r} B y_{r})
\] (21)
where $\lambda_i$ is given in (4), whereas $\theta_n$, which depends on the available CSI at the relays, and the relaying matrices $A_i \in \mathbb{C}^{P \times K}$ and $B_i \in \mathbb{C}^{P \times K}$, obeying the constraint $A_i^H A_i + B_i^H B_i = I_K$, for $i \in \{1, 2, \ldots, N\}$, will be specified example by example in both cases. Furthermore, we calculate the ABER performance of the $S \rightarrow D$ link (referred to as ‘w/o cooperation’) given by (see [6])

$$\text{ABER}_{x,d} = \frac{1}{2} \left(1 - \frac{\rho \sigma_{f,a}^2}{\rho \sigma_{f,a}^2 + \rho}\right)$$

where $\sigma_{f,a}^2 \triangleq (2\delta)^{-P}$ is the variance of the channel gain between $S$ and $D$, whereas $\delta = 1$ for BPSK signalling and $\delta = 2$ in the case of QPSK modulation. For the R-DSTBC protocols under comparison (‘randomised A&F (MF)‘ and ‘randomised A&F (ZF)‘), the matrix $\mathcal{R}$ is drawn from a Gaussian spherical distribution [18], which has complex-valued entries in Examples 4-A and 4-C, whereas it takes on real values in Example 4-B. For each of the $10^5$ Monte Carlo run carried out (wherein, besides the network configuration, channel coefficients, noise, data sequences and randomisation coefficients are randomly generated), an independent record of $10^3$ symbols is considered to evaluate the ABER of all the considered DSTBC schemes.

As a final remark, it should be remembered that the ‘centralised A&F LDC‘ protocol is centralised since a control entity has to assign relays the codes; the ‘randomised A&F (MF)‘ and ‘randomised A&F (ZF)‘ schemes are completely decentralised and coherent detection at $D$ requires estimation of the $L$-dimensional vector $h$; on the other hand, although the relaying process for the ‘randomised A&F LDC‘ method is decentralised, when the relaying matrices $\{A_i\}_{i=1}^N$ and $\{B_i\}_{i=1}^N$ are unknown at $D$, estimation of the equivalent channel matrix $\Xi \triangleq \sum_{i=1}^N \{A_i^H g_i A_i, A_i^H g_i B_i, B_i^H B_i\} \in \mathbb{C}^{P \times (K \times K)}$ is required, which involves acquisition of $2KP \gg L$ parameters. This number of channel coefficients can be reduced to $N$ if each relay transmits to $D$ the random seed used to generate $\{A_i\}_{i=1}^N$ and $\{B_i\}_{i=1}^N$ through a finite-rate feedback channel. Moreover, in all the considered examples, both ‘centralised A&F LDC‘ and ‘randomised A&F LDC‘ methods are not single-symbol ML decodable.

4.1 Full-rate complex orthogonal STBC ($K = 2$, $P = 2$, $L = 2$, $N \in \{2, 4, 8\}$)

In this example, we use as complex STBC rule the full-rate (i.e. $R_{\text{code}} = 1$) Alamouti code of order $L = K = P = 2$ given by (see [4])

$$\mathbf{C}(a) = \begin{bmatrix} a_1 & a_2 \\ -a_2 & a_1 \end{bmatrix}$$

In such a case, according to Lemma 1, the ‘randomised A&F (MF)‘ rule is single-symbol ML decodable because of the orthogonal structure of $\mathbf{C}(a)$, that is, $\mathbf{C}(a)^H \mathbf{C}(a) = \|a\|^2 I_2$, and the fact that the noise correlation matrix $\mathbf{K}$ in (8) is a scaled identity matrix. The same property also holds for the ‘randomised A&F (ZF)‘ scheme [1]. We first consider a wireless network with $N = 2$ relays. In this case, for the ‘centralised A&F LDC‘ protocol, the relaying matrices are chosen as follows (see [35])

$$A_1 = I_2, \quad B_1 = A_2 = O_{2 \times 2}, \quad B_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

In the case of the ‘randomised A&F LDC‘ scheme, we set...
$B_1 = A_2 = O_{2 \times 2}$, whereas $A_1$ and $B_2$ are Haar-distributed unitary complex random matrices. It is easily verified that, for these LDC methods, the noise correlation matrix at $D$ is a scaled identity matrix. Since it is numerically shown in [35] that knowledge of the channel gains $\{f_i\}_i^N$ at the relays does not improve the performances of LDC methods in the case of the Alamouti design, we choose $\theta_i = 1$ in (21) for both the ‘centralised A&F LDC’ and ‘randomised A&F LDC’ methods.

It is seen from Fig. 2 that the direct transmission between $S$ and $D$ only achieves diversity order equal to 1, by providing much worse performance than the considered A&F schemes in the high-SNR region. On the other hand, all the considered A&F protocols achieve the maximal diversity order 2. In particular, the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ schemes exhibit substantially the same performance of the ‘centralised A&F LDC’ protocol, by significantly outperforming its decentralised version ‘randomised A&F LDC’. For the ‘centralised A&F LDC’ protocol, the relaying matrices are chosen as follows (see [35])

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B_1 = O_{2 \times 2}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_2 = O_{2 \times 2}$$

In the case of the ‘randomised A&F LDC’ scheme, we set $B_1 = B_2 = O_{2 \times 2}$, whereas $A_1$ and $A_2$ are Haar-distributed unitary complex random matrices. It results that the coding gain of all the considered A&F schemes improves as $N$ increases, and the performance advantage of the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ methods over the ‘randomised A&F LDC’ one is still 2 dB for $10^{-6} < \text{ABER} \leq 10^{-5}$.

### 4.2 Full-rate real orthogonal STBC ($K = 2, P = 2, L = 2, N = 2$)

In this example, we consider the case of $N = 2$ relays and, besides assuming that the modulation format and the randomisation matrix $\mathcal{R}$ are real-valued, we use as real STBC rule the full-rate (i.e. $R_{\text{code}} = 1$) orthogonal code of order $L = K = P = 2$ given by (see [23])

$$C(a) = \begin{bmatrix} a_1 & a_2 \\ -a_2 & a_1 \end{bmatrix}$$

(25)

It is noteworthy that, in this case, compensation of the phase of $f_i$ is unnecessary [2] and, hence, CSI is not exploited at the relays, thus leading to a blind relaying scheme, which is herein referred to as ‘randomised A&F’. For the ‘centralised A&F LDC’ protocol, the relaying matrices are chosen as follows (see [35])

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B_1 = O_{2 \times 2}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_2 = O_{2 \times 2}$$

(26)

In the case of the ‘randomised A&F LDC’ scheme, we set $B_1 = B_2 = O_{2 \times 2}$, whereas $A_1$ and $A_2$ are Haar-distributed unitary complex random matrices. It is easily verified that, for these LDC methods, the noise correlation matrix at $D$ is a scaled identity matrix. Since it is numerically shown in [35] that knowledge of the channel gains $\{f_i\}_i^N$ at the relays does not improve the performances of LDC methods in the case of the Alamouti design, we choose $\theta_i = 1$ in (21) for both the ‘centralised A&F LDC’ and ‘randomised A&F LDC’ methods.

It is seen from Fig. 2 that the direct transmission between $S$ and $D$ only achieves diversity order equal to 1, by providing much worse performance than the considered A&F schemes in the high-SNR region. On the other hand, all the considered A&F protocols achieve the maximal diversity order 2. In particular, the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ schemes exhibit substantially the same performance of the ‘centralised A&F LDC’ protocol, by significantly outperforming its decentralised version ‘randomised A&F LDC’. For the ‘centralised A&F LDC’ protocol, the relaying matrices are chosen as follows (see [35])

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B_1 = O_{2 \times 2}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_2 = O_{2 \times 2}$$

In the case of the ‘randomised A&F LDC’ scheme, we set $B_1 = B_2 = O_{2 \times 2}$, whereas $A_1$ and $A_2$ are Haar-distributed unitary complex random matrices. It results that the coding gain of all the considered A&F schemes improves as $N$ increases, and the performance advantage of the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ methods over the ‘randomised A&F LDC’ one is still 2 dB for $10^{-6} < \text{ABER} \leq 10^{-5}$.

### 4.2 Full-rate real orthogonal STBC ($K = 2, P = 2, L = 2, N = 2$)

In this example, we consider the case of $N = 2$ relays and, besides assuming that the modulation format and the randomisation matrix $\mathcal{R}$ are real-valued, we use as real STBC rule the full-rate (i.e. $R_{\text{code}} = 1$) orthogonal code of order $L = K = P = 2$ given by (see [23])

$$C(a) = \begin{bmatrix} a_1 & a_2 \\ -a_2 & a_1 \end{bmatrix}$$

(25)

It is noteworthy that, in this case, compensation of the phase of $f_i$ is unnecessary [2] and, hence, CSI is not exploited at the relays, thus leading to a blind relaying scheme, which is herein referred to as ‘randomised A&F’. For the ‘centralised A&F LDC’ protocol, the relaying matrices are chosen as follows (see [35])

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B_1 = O_{2 \times 2}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B_2 = O_{2 \times 2}$$

(26)

In the case of the ‘randomised A&F LDC’ scheme, we set $B_1 = B_2 = O_{2 \times 2}$, whereas $A_1$ and $A_2$ are Haar-distributed unitary complex random matrices. It results that the coding gain of all the considered A&F schemes improves as $N$ increases, and the performance advantage of the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ methods over the ‘randomised A&F LDC’ one is still 2 dB for $10^{-6} < \text{ABER} \leq 10^{-5}$.
unitary real random matrices. As in Example 4-A, for both the LDC methods, we choose $\theta_i = 1$ in (21) and the noise correlation matrix at $D$ is scaled identity matrix. Reading of results in Fig. 4 leads to the same conclusions drawn from Fig. 2. The only remarkable difference is the performance improvement exhibiting by all the considered schemes, which is basically due to the fact that only one bit per symbol is transmitted with BPSK modulation instead of two bits per symbol as in the QPSK case of Example 4-A.
4.3 Complex orthogonal STBC with code rate 3/4
\((K=3, P=4, L=3, N \in \{3, 6, 9\})\)

In this example, we use as complex STBC rule the orthogonal code of order \(L = K = 3\) given by (see [46])

\[
C(a) = \begin{bmatrix}
a_1 & a_2 & a_3 \\
-a_2^* & a_1 & 0 \\
-a_3^* & 0 & a_1^*
\end{bmatrix}
\]

(27)

whose code rate is \(R_{\text{code}} = 3/4\). We remember that, in this case, the noise correlation matrix \(K\) in (8) is not a scaled identity matrix and, thus, the ‘randomised A&F (MF)’ rule is not single-symbol ML decodable. The same conclusion applies to the ‘randomised A&F (ZF)’ scheme [1]. We first consider a wireless network with \(N = 3\) relays. In this case, for the ‘centralised A&F LDC’ protocol, the relaying matrices are chosen as follows

\[
A_1 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix},
B_1 = \begin{bmatrix}
0 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix},
\]

\[
A_2 = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix},
B_2 = \begin{bmatrix}
0 & 0 & 0 \\
1 & 0 & 0 \\
0 & 0 & 1 \\
\end{bmatrix},
\]

(28)

\[
A_3 = \begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix},
B_3 = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
-1 & 0 & 1 \\
\end{bmatrix}
\]

and, thus, the space–time code word formed at \(D\) has the form in (39). In the case of the ‘randomised A&F LDC’ scheme, \(\{A_i\}_{i=1}^3\) and \(\{B_i\}_{i=1}^3\) are Haar-distributed unitary complex random matrices. According to [35], we choose \(\theta = (f_i^*/|f_i|)\) in (21) such that, similarly to the R-DSTBC A&F method, both the LDC schemes can exploit CSI at the relays.

Fig. 5 reports the ABER performances of all the considered techniques with ML detection at \(D\). We underline that, in addition to estimation of the relevant channel coefficients, the ML detector of all the A&F schemes under comparison requires knowledge of the noise correlation matrix at \(D\). It results from Fig. 5 that the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ schemes perform comparably to the ‘centralised A&F LDC’ protocol, by achieving the maximal diversity order 3. Even in this example, with respect to the ‘decentralised A&F LDC’ method (whose diversity order is 3, too), the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ techniques enable a noticeable saving in terms of average SNR, which is a little more than 3 dB at ABER of \(10^{-5}\).

It is interesting to evaluate the impact of the knowledge at \(D\) of the noise correlation matrix. To this aim, Fig. 6 depicts the performances of the A&F schemes under comparison when the correlation properties of the noise are ignored in the detection process. In this case, estimation of the noise correlation matrix is unnecessary for all the considered A&F methods; additionally, because of the orthogonal property of \(C(a)\) in (39), the resulting suboptimal detection rule (10) is single-symbol decodable for the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ schemes. It is apparent from Figs. 5 and 6 that the performance loss with respect to the optimum ML detector is imperceptible.

In Fig. 7, we only consider the decentralised schemes with ML decoding at \(D\), by simulating a network with \(N = 6\) and \(N = 9\) relays. With reference to the ‘randomised A&F LDC’
approach, the relaying matrices \( [A_i]_{i=1}^{N} \) and \( [B_j]_{j=1}^{N} \) are Haar-distributed unitary complex random matrices. Results show that performance gap between the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ methods and the ‘randomised A&F LDC’ one diminishes as the number of relays increases, by anyway being greater than 1 dB for ABER values ranging from \( 10^{-3} \) to \( 10^{-6} \). Moreover, while the ‘randomised A&F (MF)’ and ‘randomised A&F (ZF)’ schemes substantially exhibit the same ABER performance in the previous reported numerical results, the ‘randomised A&F (MF)’ rule slightly outperforms in Fig. 7 its ZF counterpart, which leads to noise amplification at the relays.

For blocks of \( P = 4 \) symbols in Phase II, if bandwidth efficiency is a stringent system requirement, one can use a full-rate (i.e. \( R_{\text{code}} = 1 \)) complex quasi-orthogonal code of order \( L = K = 4 \) (see Appendix 2) with a QPSK modulation or, alternatively, a full-rate real orthogonal design of the same order [23] together with a 4-ASK constellation. In the former case, the price to pay is a slighty increase in decoding complexity, whereas in the latter one the disadvantage is a lower noise immunity.

5 Conclusions

In this paper, we studied the R-DSTBC A&F rule, which is based on the idea that the cooperating nodes transmit an independent random linear combination of the \( L \) columns of a given STBC matrix associated with a scaled version of the received data block. Such a relaying scheme is decentralised, that is, a preliminary code allocation procedure is not required, and does not involve knowledge of the random seeds used to generate the relaying matrices. A future research issue consists of analytically evaluating the performance of the considered schemes in the presence of imperfect CSI.

6 References

7 Appendix

7.1 Appendix 1: Proof of Lemma 1

Complex orthogonal STBC rules exhibit the property $C(\gamma|\gamma) = \| \gamma \|_2^4, \forall \gamma \in C^K \setminus \{0\}$. A nice property of such codes is that $C(\cdot)$ can be designed so as to exhibit a kind of 'commutative' property, which can be stated as follows: there exists a subset $I \subseteq \{1, 2, \ldots, P\}$ such that

$$[C(\gamma|\beta)_{I}]_{\gamma} = \mathcal{D}(\beta|\gamma),$$

for any $\gamma \in C^K$ and $\beta \in C^L$ (29)

where $\mathcal{D}(\beta|\gamma) \in C^P \times K$ is a complex orthogonal transformation of $\beta$, that is, $\mathcal{D}^H(\beta|\gamma) \mathcal{D}(\beta|\gamma) = \| \beta \|^2 I_K$, for any $\beta \in C^L \setminus \{0\}$, whose structure is uniquely determined from $C(\cdot)$. It can be shown [45] that (29) holds if and only if the $p$th row of $C(\cdot)$, $\forall p \in I$, has all its non-zero entries conjugated (‘conjugate row’), whereas the $p$th row of $C(\cdot)$, $\forall p \in \{1, 2, \ldots, P\} \setminus I$, has all its non-zero entries non-conjugated (non-conjugate row). Complex orthogonal STBC designs exhibiting such a desired property can be found in [46]. For instance, the well-known $2 \times 2$ complex orthogonal STBC design proposed by Alamouti [4], for which $P = K = L = 2$, fulfills condition (29) with $I = \{2\}$. 

References


© The Institution of Engineering and Technology 2013

IET Commun., pp. 1–16
Accounting for (6), one has
\[ y_d^{[1]} = \left[ C(a)h \right]^{[1]} + \hat{w}_d^{[1]} = \mathcal{D}(h)u + \hat{w}_d^{[1]} \] (30)
which can be used instead of \( y_d \) for ML decoding since, by the theorem of reversibility [42], no information is lost in such a transformation. By virtue of (7) and (29), the correlation matrix of \( \hat{w}_d^{[1]} \) assumes the form
\[
\tilde{K} \triangleq E\left[ \hat{w}_d^{[1]} \left( \hat{w}_d^{[1]} \right) \right] = \sum_{i=1}^{N} \left| g_i \right|^2 \lambda_i^2 \sigma_n^2 \mathcal{D}(r_i) e^{j\angle g_i} \mathcal{D}^H(r_i) e^{-j\angle g_i} + \sigma_n^2 \mathbf{I}_P
\] (31)
If the STBC matrix has rate \( R_{\text{code}} = 1 \), that is, \( P = K \), one also has \( \mathcal{D}(r_i) e^{j\angle g_i} \mathcal{D}^H(r_i) e^{-j\angle g_i} = \|r_i\|^2 \mathbf{I}_K \), which implies that \( \tilde{K} = \left( \sum_{i=1}^{N} \left| g_i \right|^2 \lambda_i^2 \sigma_n^2 \|r_i\|^2 + \sigma_n^2 \right) \mathbf{I}_K \), that is, the noise correlation matrix is a scaled identity matrix. In this case, since \( \mathcal{D}^H(h) \mathcal{D}(h) = \|h\|^2 \mathbf{I}_K \), the rule
\[
\tilde{a}_{\text{opt}} = \arg \min_{a \in \mathbb{A}} \left[ \tilde{y}_d^{[1]} - \mathcal{D}(h)u \right]^H \tilde{K}^{-1} \left[ \tilde{y}_d^{[1]} - \mathcal{D}(h)u \right]
\] (32)
is equivalent to \( K \) independent single-symbol ML decision rules.

Let \( \gamma \triangleq \left\{ \gamma_1, \gamma_2, \ldots, \gamma_K \right\} \in \mathbb{C}^K - \{0_K\} \), real orthogonal STBC designs exhibit the characteristic property \( C^T(\gamma)C(\gamma) = \left( \sum_{i=1}^{K} \gamma_i^2 \right) \mathbf{I}_K \). In this case, similarly to (29), the following property holds
\[ C(\gamma)\beta = \mathcal{D}(\beta)\gamma \] for any \( \gamma \in \mathbb{C}^K \) and \( \beta \in \mathbb{C}^L \) (33)
where \( \mathcal{D}(\beta) \in \mathbb{C}^{P \times K} \) is a real orthogonal transformation of \( \beta \triangleq \left\{ \beta_1, \beta_2, \ldots, \beta_K \right\} \in \mathbb{C}^L - \{0_L\} \), that is, \( \mathcal{D}^H(\beta)\mathcal{D}(\beta) = \sum_{i=1}^{K} \beta_i^2 \mathbf{I}_K \), whose structure is uniquely determined from \( C(\cdot) \). By virtue of (7) and (33), the noise correlation matrix (8) becomes
\[
K = \sum_{i=1}^{N} \left| g_i \right|^2 \lambda_i^2 \sigma_n^2 \mathcal{D}(r_i) \mathcal{D}^H(r_i) + \sigma_n^2 \mathbf{I}_P
\] (34)
If the STBC matrix \( C(\cdot) \) has rate \( R_{\text{code}} = 1 \), that is, \( P = K \), and all the entries of the randomisation vectors \( r_1, r_2, \ldots, r_N \) are real-valued, one also has \( \mathcal{D}(r_i) \mathcal{D}^H(r_i) = \mathcal{D}(r_i)^2 \), \( \mathcal{D}(r_i) = \|r_i\|^2 \mathbf{I}_K \), which implies that \( K = \left( \sum_{i=1}^{N} \left| g_i \right|^2 \lambda_i^2 \sigma_n^2 \|r_i\|^2 + \sigma_n^2 \right) \mathbf{I}_K \), that is, the noise correlation matrix is a scaled identity matrix. In this case, since \( C^T(a)C(a) = \sum_{i=1}^{K} (\lambda_i^2 \sigma_n^2 \|r_i\|^2 \mathbf{I}_K) \), the ML decision rule (6) is equivalent to \( K \) independent single-symbol ML decision rules if all the entries of the symbol vector \( a \) are real-valued, too.

### 7.2 Appendix 2: Pairwise decoding of the R-DSTBC A&F rule

Full-rank square (i.e. \( L = P = K \)) complex quasi-orthogonal codes fulfil property (29) exhibited by complex orthogonal STBC rules, with the difference [24] that, in this case, \( \mathcal{D}(\beta) \in \mathbb{C}^{K \times K} \) is a complex quasi-orthogonal transformation of \( \beta \), that is, \( \mathcal{D}^H(\beta)\mathcal{D}(\beta) = \mathcal{D}(\beta)\mathcal{D}^H(\beta) = \|\beta\|^2 \mathbf{I}_K \), for any \( \beta \in \mathbb{C}^K - \{0_K\} \), where \( \mathbf{I}_K \) is a sparse matrix with ones on its main diagonal. For instance, for the ABBA code proposed in [47]
\[
C(a) = \begin{bmatrix}
  a_1 & a_2 & a_3 & a_4 \\
  a_2^* & -a_1 & a_4 & -a_3 \\
  a_3 & a_4 & -a_1 & -a_2 \\
  a_4^* & a_3^* & a_2^* & -a_1^*
\end{bmatrix}
\] (35)
it results that \( K = 4, I = \{2, 4\} \) and
\[
\Delta(\beta) = \begin{bmatrix}
  1 & 0 & X_{\text{ABBA}}(\beta) & 0 \\
  0 & 1 & 0 & X_{\text{ABBA}}(\beta) \\
  0 & X_{\text{ABBA}}(\beta) & 0 & 1 \\
  0 & 0 & X_{\text{ABBA}}(\beta) & 0
\end{bmatrix}
\] (36)
where \( X_{\text{ABBA}}(\beta) \triangleq 2 \Re \{ \beta_1 \beta_2^* + \beta_2 \beta_3^* \}/\|\beta\|^2 \). For the extended Alamouti (EA) code proposed in [48]
\[
C(a) = \begin{bmatrix}
  a_1 & a_2 & a_3 & a_4 \\
  a_2^* & -a_1 & a_4 & -a_3 \\
  a_3 & a_4 & -a_1 & -a_2 \\
  a_4^* & a_3^* & a_2^* & -a_1^*
\end{bmatrix}
\] (37)
it results that \( K = 4, I = \{2, 3\} \) and
\[
\Delta(\beta) = \begin{bmatrix}
  1 & 0 & 0 & X_{\text{EA}}(\beta) \\
  0 & 1 & 0 & -X_{\text{EA}}(\beta) \\
  0 & -X_{\text{EA}}(\beta) & 1 & 0 \\
  0 & 0 & X_{\text{EA}}(\beta) & 1
\end{bmatrix}
\] (38)
where \( X_{\text{EA}}(\beta) \triangleq 2 \Re \{ \beta_1 \beta_2^* + \beta_2 \beta_3^* \}/\|\beta\|^2 \). For such codes, the expression (31) of the correlation matrix of \( \hat{w}_d^{[1]} \) is still valid and assumes the particular form
\[
\tilde{K} = \sum_{i=1}^{N} \left| g_i \right|^2 \lambda_i^2 \sigma_n^2 \|r_i\|^2 \Delta(\beta) + \sigma_n^2 \mathbf{I}_K
\] (34)
where \( \Delta(\beta) \) can boil down to the identity matrix by judiciously choosing the entries of the vector \( r_i = [r_{i,1}, r_{i,2}, r_{i,3}, r_{i,4}]^T \). Indeed, in the case of the ABBA code, if \( R\{r_{i,1}r_{i,3}^* + r_{i,2}r_{i,4}^*\} = 0 \), for example, \( r_{i,1} \) and \( r_{i,2} \) are real-valued.
7.3 Appendix 3: Proof of Lemma 2

The fact that the detector (10) is suboptimum can also be shown analytically by observing that

\[
\begin{align*}
    h^H C_{k,j}^c h &= (h^H C_{k,j}^c K^{-1/2})(K^{1/2} C_{k,j} h) \\
    &\leq \|K^{-1/2} C_{k,j} h\| \|K^{1/2} C_{k,j} h\| 
\end{align*}
\]

(42)

where the inequality comes from the Cauchy-Schwarz inequality [50], which, accounting for (11) and (12), leads to

\[
P\left(\left\{ \alpha_k \rightarrow \alpha_\ell \right\}_{\text{subopt}} \right) \leq P\left(\left\{ \alpha_k \rightarrow \alpha_\ell \right\}_{\text{subopt}} \right) \leq \frac{1}{2} \left| \left| h^H C_{k,j}^c h \right| \right|
\]

(43)

By invoking the Rayleigh-Ritz inequality [50], one has that

\[
\mu_1 (K) h^H C_{k,j}^c h \leq h^H C_{k,j} h
\]

Thus, the PEP of the ML detector given by (11) can be upper bounded as

\[
P\left(\left\{ \alpha_k \rightarrow \alpha_\ell \right\}_{\text{subopt}} \right) \leq \frac{1}{2} \left| \left| h^H C_{k,j}^c h \right| \right|
\]

(43)

7.4 Appendix 4: Proof of Theorem 1

Accounting for (13) and remembering the structure of \( h \) in (6), one obtains

\[
P\left(\left\{ \alpha_k \rightarrow \alpha_\ell \right\}\left| h \right| \right) \leq \frac{1}{2} \exp \left[ -\frac{\sum \left| g_i \right|^2 \lambda_i^2 \left| \Delta \right|_2 \left| \Delta \right|_2 2 \right] \frac{\| h^H C_{k,j}^c h \|^2}{4 \mu_1 (K)}
\]

(45)

where the inequality comes from the Chernoff bound [42]. Using the result [51] that

\[
E \left[ \exp \left( -\sum \left| g_i \right|^2 \lambda_i^2 \left| \Delta \right|_2 \left| \Delta \right|_2 2 \right) \right] \leq \frac{1}{\left( \sum \left| g_i \right|^2 \lambda_i^2 \left| \Delta \right|_2 \left| \Delta \right|_2 2 \right)}
\]

(46)

where \( \Phi_{k,\ell} \triangleq R^H C_{k,\ell}^c R \in \mathbb{C}^{N \times N} \) is a given matrix. At this point, let us focus on the largest eigenvalue \( \mu_1 (K) \) of the noise correlation matrix \( K \) given by (8). Using the Weyl’s theorem [50] and invoking property (p3) of the STBC matrix, one obtains

\[
\mu_1 (K) \leq \sum_{i=1}^{N} |g_i|^2 \lambda_i R_i + \sigma_n^2
\]

(47)

where \( \sigma_n^2 = \sigma_n^2 = \cdots = \sigma_n^2 = \sigma_n^2 \). Consequently, (46) can be further upper bounded as

\[
P\left(\left\{ \alpha_k \rightarrow \alpha_\ell \right\}\left| h \right| \right) \leq \frac{1}{2} \exp \left[ -\frac{\sum \left| g_i \right|^2 \lambda_i^2 \left| \Delta \right|_2 \left| \Delta \right|_2 2 \right] \frac{1}{4 \mu_1 (K)}
\]

(48)

with \( \xi(g) \triangleq 4 \sigma_n^2 (K \sum_{i=1}^{N} |g_i|^2 \lambda_i R_i + 1) \). Since \( \Sigma_{f}, \Lambda \) and \( \widetilde{G} \) are nonsingular and the rank of a matrix is unchanged upon left or right multiplication by a nonsingular matrix [50], it follows that

\[
\text{rank} (\Phi_{k,\ell}) = \text{rank} (C_{k,\ell}^c R)
\]

Moreover, since \( \text{rank} (C_{k,\ell}) = \min (P, L) \) by property (p4) and \( \text{rank} (R) = R \leq L \leq \min (L, N) \), it follows that

\[
\text{rank} (\Phi_{k,\ell}) = \text{rank} (C_{k,\ell}^c R) = \min (P, L) + R - L \leq \text{rank} (C_{k,\ell} R) \leq \min (P, L, R).
\]

In the case of \( P \geq L \), one has \( \text{rank} (C_{k,\ell} R) = R \). Therefore since \( \Phi_{k,\ell} \) is a positive semidefinite Hermitian matrix, that is, its eigenvalues are non-negative real numbers, and the eigenvalues of \( \Sigma_{f} \Lambda \widetilde{G} \) are simply equal to its diagonal...
entries \( \{ \sigma_{g_i}^2 \lambda_i^2 | g_i |^2 \} \) (which are positive and are assumed to be arranged in increasing order, that is, \( \sigma_{g_1}^2 \lambda_1^2 | g_1 |^2 \leq \sigma_{g_2}^2 \lambda_2^2 | g_2 |^2 \leq \cdots \leq \sigma_{g_N}^2 \lambda_N^2 | g_N |^2 \)), one obtains
\[
\det \left[ I_N + \mathbf{\Phi}_{k,i} \Sigma_i \Lambda^2 \mathbf{G} \mathbf{G}^* \right] = 1 + \frac{\mu_r (\mathbf{\Phi}_{k,i} \Sigma_i \Lambda^2 \mathbf{G} \mathbf{G}^*)}{\xi (g)} \approx 1 + \left( \prod_{r=1}^R \mu_r (\mathbf{\Phi}_{k,i}) \right) \left( \prod_{r=1}^R \sigma_{g_r}^2 \lambda_r^2 | g_r |^2 \right) \xi (g) ^{-1} \tag{49} \]

where the approximation \( \approx \) holds in the high-SNR regime, that is, for \( \sigma^2 \to 0 \). Starting from (49), the next step consists of evaluating the ensemble average with respect to \( \mathbf{g} \) of the right-hand side of (48).

Let us now observe that \( \xi (g) \) can be equivalently expressed as \( \xi (g) \triangleq 4 \sigma^2 (K S_N + 1) \), where we have defined the random variable \( S_N \triangleq \sum_{i=1}^N | g_i |^2 \lambda_i^2 r_i^H \mathbf{J}_r \). It is easily seen that \( | g_i |^2 \lambda_i^2 r_i^H \mathbf{J}_r \), for \( i \in \{ 1, 2, \ldots, N \} \), are mutually independent exponential random variables with means \( \sigma_{g_r}^2 \lambda_r^2 r_i^H \mathbf{J}_r \) and, hence, variances \( \left( \sigma_{g_r}^2 \lambda_r^2 r_i^H \mathbf{J}_r \right)^2 \); thus, such random variables are not identically distributed. At this point, we can rely on the following result:

**Lemma 4:** Under the assumption that
\[
\lim_{N \to +\infty} \sum_{r=1}^N \frac{\left( \sigma_{g_r}^2 \lambda_r^2 r_i^H \mathbf{J}_r \right)^2}{\sigma_{g_r}^2} = +\infty \tag{50} \]
the sequence \( \{ | g_i |^2 \lambda_i^2 r_i^H \mathbf{J}_r \} \) obeys the strong law of large numbers, that is, to every pair of real numbers \( \epsilon > 0 \) and \( \Delta > 0 \), there corresponds an \( M > 0 \) such that, for every \( H > 0 \), the probability of the simultaneous fulfillment of the \( H + 1 \) inequalities
\[
| S_N - E(S_N | \mathcal{R}) | < \epsilon, \tag{51} \]
for \( N \in \{ M, M + 1, \ldots, M + H \} \)
is greater than or equal to \( 1 - \Delta \).

**Proof:** This is a direct consequence of the Kolmogorov criterion [52]. \qed

**Lemma 5:** states that \( S_N \) converges almost surely to \( E(S_N | \mathcal{R}) \), provided that (50) holds. On the basis of this result, we resort to the approximation \( S_N \approx \sum_{r=1}^R \sigma_{g_r}^2 \lambda_r^2 r_i^H \mathbf{J}_r \), especially for

\[
E \left[ E\left[ P \left( \{ \mathbf{a}_k \to \mathbf{a}_i \} | h | \mathbf{g}, \mathcal{R} \right) \right] \right] \rightarrow \frac{1}{2} E \left[ \prod_{r=1}^R \left( 1 + \chi_{k,r} | g_r |^2 \right)^{-1} | \mathcal{R} \right] \tag{52} \]

where
\[
\chi_{k,r} \triangleq \frac{\mu_r (\mathbf{\Phi}_{k,i}) \sigma_{g_r}^2 \lambda_r^2}{4 \sigma^2 (K \sum_{i=1}^N \sigma_{g_i}^2 \lambda_i^2 r_i^H \mathbf{J}_r + 1)} \]
and \( | g_r |^2 \), for \( r \in \{ 1, 2, \ldots, R \} \), are mutually independent exponential random variables with means \( \sigma_{g_r}^2 \). After some calculations, it is verified that
\[
E \left[ \prod_{r=1}^R \left( 1 + \chi_{k,r} | g_r |^2 \right)^{-1} | \mathcal{R} \right] \rightarrow \prod_{r=1}^R \frac{1}{1 + \chi_{k,r} x \sigma_{g_r}^2} \exp \left( -\frac{x}{\sigma_{g_r}^2} \right) dx \tag{53} \]

and
\[
\prod_{r=1}^R \frac{1}{1 + \chi_{k,r} \sigma_{g_r}^2} \exp \left( -\frac{1}{\sigma_{g_r}^2 \chi_{k,r}} \right) E \left( \frac{1}{1 + \sigma_{g_r}^2 \chi_{k,r}} \right) \tag{54} \]

We observe that, for high SNR values, the parameter \( \lambda_i \) in (4) (with \( | \theta | = 1 \)) is well approximated as
\[
\lambda_i \approx \frac{\sigma_{g_i}^2}{\sqrt{R \sigma_{g_i}^2 \sigma_{\theta_i}^2}} \tag{55} \]

Hence, recalling the definition of \( \chi_{k,i,r} \), in the high-SNR regime, we obtain
\[
E \left[ P \left( \{ \mathbf{a}_k \to \mathbf{a}_i \} | h | \mathbf{g}, \mathcal{R} \right) \right] \rightarrow \ln \left( 1 + \sigma_{g_i}^2 \chi_{k,i,r} \right) \sim \ln \left( \left( \frac{1}{\sigma^2} \right) \right) \tag{54} \]

\[
= \ln \left( \frac{1}{\sigma^2} \right) + \ln \left( \frac{\mu_r (\mathbf{\Phi}_{k,i}) \sigma_{g_r}^2 \sigma_{g_i}^2 \lambda_r^2}{4 \left( K \sum_{i=1}^N \sigma_{g_i}^2 \lambda_i^2 r_i^H \mathbf{J}_r + 1 \right)} \right) \approx \ln \left( \frac{1}{\sigma^2} \right) \tag{55} \]
Moreover, it can be shown (see, e.g. [54]) that
\[
\prod_{r=1}^{R} \mu_r(\Phi_{k,i}) \geq \left[ \prod_{r=1}^{R} \mu_r(\mathcal{R}\mathcal{R}^H) \right] \times \left[ \prod_{r=L-R+1}^{L} \mu_r(\mathcal{C}_{k,i}^H) \right] \geq \prod_{r=1}^{R} \mu_r(\mathcal{R}\mathcal{R}^H) \times \prod_{r=L-R+1}^{L} \mu_r(\mathcal{C}_{k,i}^H)
\] (56)

Equation (16) follows by substituting \(x_{k,i}^r\) in (54), using (55), (56), and remembering that \(\gamma = 1/\sigma^2\).

7.5 Appendix 5: Proof of Lemma 4

Since the matrix \(\mathcal{R}\) has rank \(R\) and satisfies (15) with probability one, averaging (16) with respect to \(\mathcal{R}\) ends up to evaluate \(E[\tilde{Y}_{k,i}(\mathcal{R})]\). To this end, we observe that, since \(r_1, r_2, \ldots, r_N\) are statistically independent vectors, the random variables \(\sigma_{k,i}^2 \lambda_i^2 r_i^H j r_i\), for \(i \in \{1, 2, \ldots, N\}\), are mutually independent with means \(\sigma_{k,i}^2 \lambda_i^2\), where we have used the fact that \(E(r_i^H j r_i) = 1\) by assumption; the corresponding variances are given by

\[
\sigma_{k,i}^4 \lambda_i^4 \left| E \left( (r_i^H j r_i)^2 \right) - 1 \right|
\]

Lemma 6: Let \(\widetilde{S}_N = \sum_{i=1}^{N} \sigma_{k,i}^2 \lambda_i^2 r_i^H j r_i\), under the assumption that

\[
\lim_{N \to +\infty} \sum_{i=1}^{N} \frac{\sigma_{k,i}^2 \lambda_i^4}{\sigma_{k,i}^2 \lambda_i^4 + 1} R \left[ \prod_{r=1}^{R} \sigma_{k,i}^2 \sigma_{k,i}^2 \mu_r(\mathcal{R}\mathcal{R}^H) \right] \left[ \prod_{r=L-R+1}^{L} \mu_r(\mathcal{C}_{k,i}^H) \right] \]

is greater than or equal to \(1 - \Delta\).

Proof: This is a direct consequence of the Kolmogorov criterion [52].

As a consequence of Lemma 6, for all \(N \geq M\), the random variable \(\widetilde{S}_N\) converges almost surely to \(E(\widetilde{S}_N) = \sum_{i=1}^{N} \sigma_{k,i}^2 \lambda_i^2\). By virtue of this result, for sufficiently large finite values of \(N\), a reasonable approximation for (17) is

\[
\tilde{Y}_{k,i}(\mathcal{R}) \approx \frac{4^R \left( K \sum_{i=1}^{N} \sigma_{k,i}^2 \lambda_i^2 + 1 \right)^R}{2 \left[ \prod_{r=1}^{R} \sigma_{k,i}^2 \sigma_{k,i}^2 \mu_r(\mathcal{R}\mathcal{R}^H) \right] \left[ \prod_{r=L-R+1}^{L} \mu_r(\mathcal{C}_{k,i}^H) \right]}
\] (59)

Equation (19) immediately follows from (59).