

RESEARCH

Open Access



Low-complexity detection for noncoherent massive MIMO communications

Marc Vilà-Insa^{1*} and Jaume Riba¹

*Correspondence:
marc.vila.insa@upc.edu

¹ Signal Processing
and Communications
Group (SPCOM), Universitat
Politècnica de Catalunya (UPC),
08034 Barcelona, Spain

Abstract

This work studies a point-to-point MIMO uplink in which user equipment transmits data to a base station employing a massive array. Signal detection is noncoherent and fading is assumed to follow the Weichselberger model. By exploiting the spatial stationarity of fading at the base station, a cyclostationary structure emerges naturally in the space-time representation, which suggests formulating the statistical properties of the received signal in the Karhunen–Loève domain. This allows the derivation of a low-complexity receiver that approximates maximum likelihood detection even for a moderate array size. The spectral analysis of the problem provides valuable insights on the design of space-time codewords.

Keywords: Noncoherent detection, Massive MIMO communications, Cyclostationary processes, Spatially stationary channel, Weichselberger model, Karhunen–Loève representation

1 Introduction

The study of events that repeat periodically over time is ubiquitous within many areas of natural and social sciences [1]. From climatology [2] to econometrics [3], through radio astronomy [4] and genomics [5], the efforts to characterize random phenomena with periodic structures have shaped a vast and rich literature over the decades. Most of these processes emerge from the interaction between periodic and random circumstances. Therefore, although the resulting processes are not periodic themselves, their statistical characterizations are [6]. Of particular interest are those stochastic processes whose second-order moments (i.e. correlation structures) are periodic functions of time, which are known as *cyclostationary (CS)* or *periodically correlated* [1]. This is the kind of statistical periodicity which has been studied more extensively [1, 7, 8].

As one would expect, man-made signals (e.g the ones used in radar or sonar [9]) or those related to human activities (such as rotating machinery [10]) are perfectly captured by the CS framework. Wireless communications are no exception: periodicities occur due to the plethora of signal processing operations involved in modulation, sampling, multiplexing or coding, among many others [1]. The pervasiveness of CS processes in communication systems has motivated the design and development of techniques that exploit their distinct nature [11]. They rely on the CS structure that

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

arises on the oversampled (i.e. over the Nyquist rate) communications signals as a means to achieve improved estimation and detection performance. As a byproduct, they also offer improved robustness to perform these fundamental tasks in adverse scenarios. The advantages come either by avoiding the use of higher order statistics, which become very noisy at low signal-to-noise ratio (SNR) regimes, or by inducing improved richness on the signal space dimensionality. For example, in [12, 13] cyclostationarity is exploited for the estimation of fundamental synchronization parameters, such as timing, symbol rate and carrier frequency, and in [14] it is used to perform blind second-order channel estimation without losing phase information. In [15], cyclostationarity is proposed as an alternative to moment-based estimation of SNR, whereas works such as [16] harness it to perform spectrum sensing in opportunistic communications based on feature detection.

The work presented herein showcases a novel formulation of a well-known communication problem in which neither the transmitter nor the receiver is aware of instantaneous channel state information (CSI), referred to as *noncoherent*. Such schemes have recently rekindled an interest in the scientific community due to the introduction of potential applications that would benefit from circumventing the acquisition of instantaneous CSI [17]. The considered approach exploits periodicities in data for noncoherent detection by harnessing the statistical knowledge of fading (i.e. *statistical CSI*) through the *channel hardening* effected by a large array at the receiver [18]. This proposal differs from classical communications literature in the sense that cyclostationarity does not arise in the time domain. Instead, we present a channel model that naturally yields a repeating correlation function, by treating the spatial and temporal dimensions jointly. This fading profile, referred to as *spatially stationary Weichselberger model*, is the result of incorporating the spatial stationarity of the electromagnetic channel [19] into the well-known correlation model from [20]. This combination of ideas allows the mathematical machinery developed in [21] for SIMO systems to be extended to a MIMO scenario.

The main contributions in this work are the following. Beside unveiling the CS qualities of the fading process, the signal at the receiver is also identified as periodically correlated. Leveraging the massive number of antennas to obtain a large sample size, we apply an extension of the Szegő theorem from [22, 23] and derive the asymptotic *Cramér–Loève (CL)* and *Karhunen–Loève (KL)* expansions of the received signal. These representations broaden the frequency domain interpretations from classical signal processing of stationary processes and allow the materialization of clear spectral correlation structures. It is in these domains in which the fundamental entities in codeword detection (e.g. power spectra and signal subspaces) become conspicuous. In particular, we study the asymptotic expression of the maximum likelihood (ML) detector, from which we obtain insights onto the pairwise codeword detection problem and relate them to other works in the literature. As a byproduct of this theoretical analysis, we propose a low-complexity implementation of the ML receiver that approximates its asymptotic spectral form for a finite number of antenna elements. By taking advantage of well-known algorithms like the fast Fourier transform (FFT), the developed technique displays improved computational complexity while retaining near-optimal performance, even for moderately-sized systems.

The paper is structured as follows. Section 2 introduces the communication problem and signal detection framework. The statistics of fading and how cyclostationarity emerges from them are presented as well. Section 3 extends signal detection in the large array regime and exploits statistical periodicities as the number of receiving antennas grows unbounded. The limiting spectral expressions provide the foundation for a low-complexity detector with negligible performance penalties, even for moderately sized receivers. Finally, Sect. 4 provides a theoretical overview of this asymptotic formulation through the lens of detection theory. Relevant mathematical structures discernible in the spectral domain are linked to physical entities to unveil further insights on noncoherent codeword detection.

The notation used throughout the text is defined next. Vectors and matrices: boldface lowercase and uppercase. (r, c) -component of a matrix: $[\mathbf{A}]_{r,c}$. Transpose and conjugate transpose: $\cdot^T, \cdot^H$. Trace: $\text{Tr}[\cdot]$. Identity matrix of size N : $\mathbf{I}_N$. Column-wise vectorization: $\text{vec}(\cdot)$. Kronecker product: $\otimes$. Matrix determinant: $|\mathbf{A}|$. Euclidean norm: $\|\mathbf{a}\|$. Frobenius norm: $\|\mathbf{X}\|_F$. Diagonal matrix constructed from the elements of a set: $\text{Diag}(\cdot)$. Dirac delta distribution: $\delta(\cdot)$. Kronecker delta: δ_a . Random variables: sans serif font. Expectation: $E[\cdot]$. Circularly symmetric complex normal vector: $\mathbf{a} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$.

2 Problem formulation

2.1 Signal model

Consider a MIMO point-to-point uplink, in which a user equipment (UE) with N_T antennas communicates wirelessly with a base station (BS) employing a massive array of N_R antennas, such that $N_R \gg N_T$ (*asymmetric massive MIMO*). The channel displays frequency flat fading with a coherence time of $K \geq N_T$. It is represented by a matrix $\mathbf{H} \in \mathbb{C}^{N_T \times N_R}$ that remains constant during K channel uses (i.e. block flat fading). We assume a power normalization of $E[\|\mathbf{H}\|_F^2] \triangleq N_T N_R$. During a coherence block, the transmitter sends a codeword $\mathbf{X} \in \mathbb{C}^{K \times N_T}$ selected equiprobably from a finite alphabet $\mathcal{X}$ of size M .

The signal at the BS (i.e. the receiver) is represented by the following complex base-band time-space matrix:

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \mathbf{Z} \in \mathbb{C}^{K \times N_R}, \quad (1)$$

where $\mathbf{Z}$ is an additive white Gaussian noise component with i.i.d. entries, such that $[\mathbf{Z}]_{r,c} \sim \mathcal{CN}(0, \mathbf{P}_Z)$. Fading is assumed correlated Rayleigh with arbitrary covariance. To properly characterize its second-order statistics, it is convenient to vectorize $\mathbf{H}$ column-wise:

$$\tilde{\mathbf{h}} \triangleq \text{vec}(\mathbf{H}) \sim \mathcal{CN}(\mathbf{0}_{N_T N_R}, \mathbf{C}_{\tilde{\mathbf{h}}}), \quad \mathbf{C}_{\tilde{\mathbf{h}}} \triangleq E[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H]. \quad (2)$$

Similarly, noise is distributed as $\tilde{\mathbf{z}} \triangleq \text{vec}(\mathbf{Z}) \sim \mathcal{CN}(\mathbf{0}_{KN_R}, \mathbf{P}_Z \mathbf{I}_{KN_R})$. Unless stated otherwise, these distributions are known at the receiver but not at the transmitter, which is referred to as *statistical CSI at the receiver*.

The previous vectorization approach can be applied onto the received signal $\mathbf{Y}$ as well. Moreover, it will prove useful when defining the maximum likelihood detector in the sequel. Therefore,

$$\tilde{\mathbf{y}} \triangleq \text{vec}(\mathbf{Y}) = \tilde{\mathbf{X}}\tilde{\mathbf{h}} + \tilde{\mathbf{z}} \in \mathbb{C}^{KN_R}, \quad (3)$$

where $\tilde{\mathbf{X}} \triangleq \mathbf{I}_{N_R} \otimes \mathbf{X}$.

2.2 Unconditional maximum likelihood detection

Given an equiprobable alphabet, the receiver that minimizes the error probability of codeword detection is known to be the maximum likelihood (ML) detector [24, Thm. 21.3.3]. It is derived from the distribution of the received signal conditioned to a transmitted codeword $\mathbf{X}_i \in \mathcal{X}$, and a channel realization $\mathbf{H}$:

$$f_{\mathbf{Y}|\mathbf{X}_i, \mathbf{H}}(\tilde{\mathbf{y}}) = \frac{1}{(\pi P_{\mathbf{Z}})^{KN_R}} \exp\left(-\frac{1}{P_{\mathbf{Z}}} \|\tilde{\mathbf{y}} - \tilde{\mathbf{X}}_i \tilde{\mathbf{h}}\|^2\right). \quad (4)$$

In a noncoherent communication setting, the channel realization is unknown at both ends of the link and treated as random (i.e. *unconditional model* [25]). It is removed from conditioning by marginalizing (4). The resulting distribution of $\mathbf{Y}|\mathbf{X}_i$ becomes

$$f_{\mathbf{Y}|\mathbf{X}_i}(\tilde{\mathbf{y}}) = E_{\mathbf{H}}[f_{\mathbf{Y}|\mathbf{X}_i, \mathbf{H}}(\tilde{\mathbf{y}})] = \frac{1}{\pi^{KN_R} |\mathbf{C}_i|} \exp(-\tilde{\mathbf{y}}^H \mathbf{C}_i^{-1} \tilde{\mathbf{y}}), \quad (5)$$

where

$$\mathbf{C}_i \triangleq \tilde{\mathbf{X}}_i \mathbf{C}_{\tilde{\mathbf{h}}} \tilde{\mathbf{X}}_i^H + P_{\mathbf{Z}} \mathbf{I}_{KN_R} \quad (6)$$

is the covariance matrix of the received signal. With the likelihood function (5), we finally define the normalized ML detector as

$$\begin{aligned} \hat{\mathbf{X}}_{\text{ML}} &\triangleq \arg \max_{\mathbf{X}_j \in \mathcal{X}} f_{\mathbf{Y}|\mathbf{X}_j}(\tilde{\mathbf{y}}) = \arg \min_{\mathbf{X}_j \in \mathcal{X}} -\frac{1}{KN_R} \ln f_{\mathbf{Y}|\mathbf{X}_j}(\tilde{\mathbf{y}}) \triangleq \arg \min_{\mathbf{X}_j \in \mathcal{X}} \mathcal{L}_j(\tilde{\mathbf{y}}) \\ &= \arg \min_{\mathbf{X}_j \in \mathcal{X}} \frac{1}{KN_R} \tilde{\mathbf{y}}^H \mathbf{C}_j^{-1} \tilde{\mathbf{y}} + \frac{1}{KN_R} \ln |\mathbf{C}_j|. \end{aligned} \quad (7)$$

It is composed of a data-dependent quadratic term and a logarithmic term:

$$Q_j(\tilde{\mathbf{y}}) \triangleq \frac{1}{KN_R} \tilde{\mathbf{y}}^H \mathbf{C}_j^{-1} \tilde{\mathbf{y}} \quad , \quad \ell_j \triangleq \frac{1}{KN_R} \ln |\mathbf{C}_j|, \quad (8)$$

such that $\mathcal{L}_j(\tilde{\mathbf{y}}) = Q_j(\tilde{\mathbf{y}}) + \ell_j$.

From a theoretical perspective, (7) entails some analysis issues. If $\mathbf{C}_{\tilde{\mathbf{h}}}$ is an unstructured, arbitrary covariance matrix, the set of ML metrics $\{\mathcal{L}_j(\tilde{\mathbf{y}})\}$ does not provide any recognizable insights onto the detection problem nor onto the design of $\mathcal{X}$ based on $\{Q_j(\tilde{\mathbf{y}})\}$ and $\{\ell_j\}$. Therefore, the next section is devoted to defining a more specific correlation model for $\mathbf{H}$. Its statistical properties, paired with the use of a large array at the BS, will allow us to represent the received signal through its Karhunen–Loève and Cramér–Loève expansions in Sect. 3. It is in such spectral domains in which the detection-relevant structures emerge. They are properly discussed in Sect. 4, where we relate them to the idea of *singular detection* [26]. As a byproduct of this exploration, a low-complexity implementation of (7) will arise and be assessed in Sect. 3.2.

2.3 Channel model and spatial stationarity

The model that will be assumed for the channel throughout the rest of this work is known as *Weichselberger model* [20], which is a particularization of correlated Rayleigh fading:

$$\mathbf{H} \triangleq \mathbf{U}_T \mathbf{H}^\circ \mathbf{U}_R^H, \quad (9)$$

where $\mathbf{U}_T \in \mathbb{C}^{N_T \times N_T}$ and $\mathbf{U}_R \in \mathbb{C}^{N_R \times N_R}$ are fixed unitary bases that represent the UE and BS array geometries, respectively. The only stochastic parameters in this model are contained in $\mathbf{H}^\circ \in \mathbb{C}^{N_T \times N_R}$, which are independent and distributed as $[\mathbf{H}^\circ]_{r,c} \sim \mathcal{CN}(0, \gamma_{r,c})$. It generalizes various classical *nongeometrical stochastic models* of fading [27], such as i.i.d. Rayleigh, keyhole and virtual channel (*beam-domain*) [28, Sec. 3.6.1]. By relaxing the structural constraints of the *Kronecker model* and allowing correlations between transmitter and receiver, it solves some of its shortcomings without fully sacrificing signal structure.

To simplify subsequent analyses, we will make further assumptions on the statistical properties of $\mathbf{H}$. It is common in the literature to consider channel fading behaves as a *spatially stationary* process, i.e. its second-order moments are invariant to spatial shifts [29]. This idea can be effectively harnessed when using a massive array (such as in the BS of the studied system), in the same manner as how it is done in the time domain when a large data set is available [30], thus giving rise to the classical frequency analysis toolkit grounded on the core property that Fourier coefficients at different frequencies become uncorrelated [31]. Therefore, we only need stationarity to manifest at the BS side of the wireless link, which translates into the rows of $\mathbf{H}$ being jointly WSS:

$$\mathbb{E}[\mathbf{H}]_{r,c} [\mathbf{H}]_{r',c'}^* \equiv \mathbb{E}[\mathbf{H}]_{r,c+l} [\mathbf{H}]_{r',c'+l}^*, \quad \forall l \in \mathbb{Z}. \quad (10)$$

To expand on how the previous notion can be exploited, we define the correlation matrix

$$\mathbf{R}_{BS} \triangleq \mathbb{E}[\mathbf{H}^H \mathbf{H}] = \mathbf{U}_R \mathbb{E}[\mathbf{H}^\circ \mathbf{H}^\circ] \mathbf{U}_R^H = \mathbf{U}_R \text{Diag}(\{\sum_{r=0}^{N_T-1} \gamma_{r,c}\}_c) \mathbf{U}_R^H \in \mathbb{C}^{N_R \times N_R}, \quad (11)$$

in which we have used the Weichselberger model from (9). The channel being spatially stationary at the BS implies that $[\mathbf{R}_{BS}]_{r,c} \equiv [\mathbf{R}_{BS}]_{r+l,c+l}$ for any l , by (10). Therefore, $\mathbf{R}_{BS}$ is constant along its diagonals which is known as a *Toeplitz matrix* (see Fig. 2a). This type of matrices have the special property of asymptotically diagonalizing onto the Fourier basis $[\mathbf{F}_{N_R}]_{r,c} \triangleq e^{j2\pi rc/N_R} / \sqrt{N_R}$ as $N_R \rightarrow \infty$ [29]. The interest behind this structure is that it can be exploited with well-known precoding implementations, such as the phase-shifter architectures used for beamspace techniques [32].

Under the premise of using a massive array in the BS, and to simplify upcoming mathematical derivations, we will assume that $\mathbf{U}_R \triangleq \mathbf{F}_{N_R}^H$, even for finite N_R . As the number of receiving antennas grows without bound, the previously described Fourier basis will asymptotically become the true eigenbasis of $\mathbf{H}$. With this, (9) becomes the *spatially stationary Weichselberger model (SSW-model)*:

$$\mathbf{H} \triangleq \mathbf{U}_T \mathbf{H}^\circ \mathbf{F}_{N_R}^H. \quad (12)$$

Returning to the formulation from Sect. 2.1, we may want to express it in vector form:

$$\tilde{\mathbf{h}} = (\mathbf{F}_{N_R} \otimes \mathbf{U}_T) \text{vec}(\mathring{\mathbf{H}}). \quad (13)$$

Its covariance matrix is

$$\mathbf{C}_{\tilde{\mathbf{h}}} = (\mathbf{F}_{N_R} \otimes \mathbf{U}_T) \mathbf{\Gamma} (\mathbf{F}_{N_R}^H \otimes \mathbf{U}_T^H), \quad (14)$$

where $\mathbf{\Gamma} \triangleq \mathbb{E}[\text{vec}(\mathring{\mathbf{H}})\text{vec}(\mathring{\mathbf{H}})^H] \in \mathbb{R}^{N_T N_R \times N_T N_R}$ is a diagonal matrix which contains its eigenvalues. By (9), we know they are nonnegative and $[\mathbf{\Gamma}]_{r+cN_T, r+cN_T} = \gamma_{r,c}$.

The following result sheds light onto the statistical structure of $\tilde{\mathbf{h}}$. A graphical intuition behind it can be found in Fig. 1.

1 Proposition 1

Under the SSW-model, fading behaves as a CS process of period N_T .

1 Proof

Recall (10). Expressing this relationship in terms of the vectorized form $\tilde{\mathbf{h}}$ yields

$$\mathbb{E}[\tilde{\mathbf{h}}_{r+cN_T} \tilde{\mathbf{h}}_{r'+cN_T}^*] \equiv \mathbb{E}[\tilde{\mathbf{h}}_{r+(c+l)N_T} \tilde{\mathbf{h}}_{r'+(c'+l)N_T}^*] \quad (15a)$$

$$[\mathbf{C}_{\tilde{\mathbf{h}}}]_{r+cN_T, r'+cN_T} = [\mathbf{C}_{\tilde{\mathbf{h}}}]_{(r+cN_T)+lN_T, (r'+cN_T)+lN_T}. \quad (15b)$$

By setting $p \triangleq r + cN_T$ and $q \triangleq r' + c'N_T$, we have that

$$[\mathbf{C}_{\tilde{\mathbf{h}}}]_{p,q} = [\mathbf{C}_{\tilde{\mathbf{h}}}]_{p+lN_T, q+lN_T}, \quad (16)$$

which is a defining property of N_T -periodic CS processes. $\square$

CS processes have been widely studied in the literature [1], since they display second order statistics that are invariant to shifts of length equal to their period. For $\tilde{\mathbf{h}}$, this

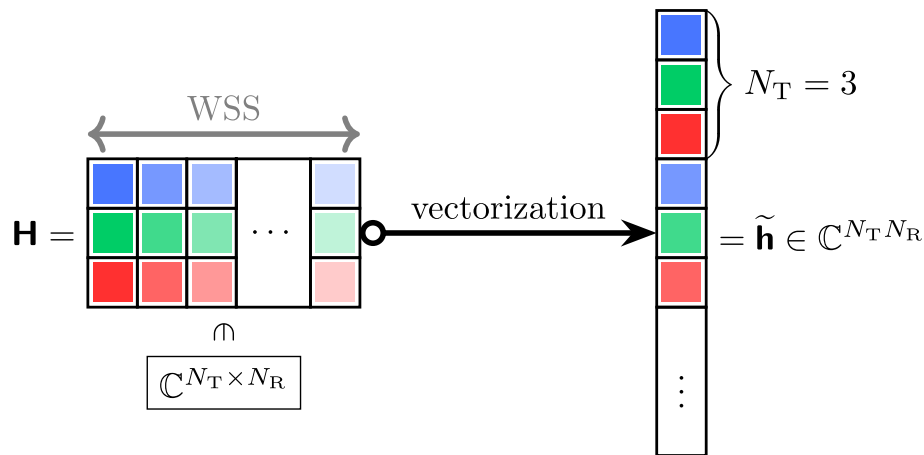


Fig. 1 Graphical representation of column-wise vectorization of $\mathbf{H}$ under the SSW-model

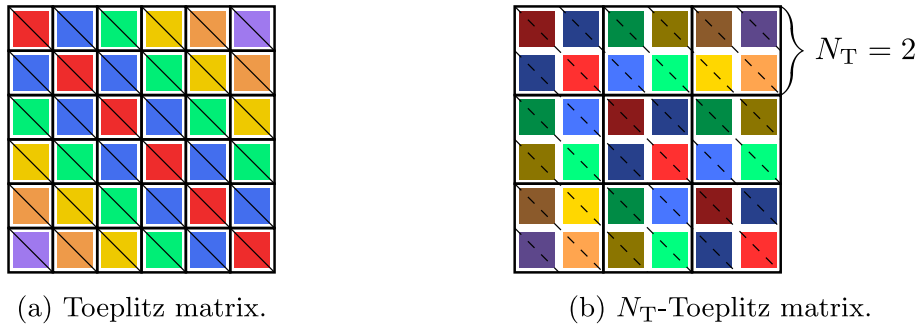


Fig. 2 Graphical representation of the covariance matrix of a WSS process and a CS process of period 2

translates into a covariance matrix that is periodic along its diagonals with period N_T , known as N_T -*Toeplitz*. These matrices are also referred to as *block-Toeplitz*, since they have constant blocks of size N_T diagonally (see Fig. 2b). They play an important role in signal processing, and have been thoroughly studied in the literature [22, 33].

2.3.1 Fading statistics in the large array regime

To conclude this section, we will present the fading correlation in the large array regime, i.e. as $N_R \rightarrow \infty$. We define the autocorrelation function (ACF) of $\{h(n) \triangleq [\tilde{h}]_n\}$, for $n \in \{-\frac{N_T N_R}{2}, \dots, \frac{N_T N_R}{2} - 1\}$ as¹

$$\mathbf{C}_H(n, m) \triangleq E[h(n+m)h(n)] = [\mathbf{C}_{\tilde{h}}]_{n+m, n}. \quad (17)$$

Under model (12), it takes the form

$$\mathbf{C}_H(n, m) = \sum_{p, q = -\frac{N_R}{2}}^{\frac{N_R}{2}-1} [\mathbf{U}_T \mathbf{C}_H(p-q) \mathbf{U}_T^H]_{n+m-pN_T, n-qN_T}, \quad (18)$$

where $\mathbf{C}_H(m) \in \mathbb{C}^{N_T \times N_T}$ is a diagonal matrix whose k th entry is the ACF of $\{\tilde{h}_k(n) \triangleq [\mathbf{H}\mathbf{F}_{N_R}]_{k,n}\}$, which is a wide-sense stationary (WSS) process. Indeed, its correlation is shift-invariant:

$$E[\tilde{h}_k(n+m)\tilde{h}_{k'}^*(n)] = \delta_{k-k'} \sum_{l=0}^{N_R-1} \frac{1}{N_R} \gamma_{k,l} e^{j2\pi m \frac{l}{N_R}} \triangleq [\mathbf{C}_H(m)]_{k,k'}. \quad (19)$$

Finally, the next result states the limit of (18) in the large array regime.

1 Proposition 2

As N_R grows without bound, the fading ACF (18) approaches

¹ We assume $N_R \in 2\mathbb{N}$ without loss of generality.

$$\lim_{N_R \rightarrow \infty} \mathbf{C}_H(n, m) = \sum_{p, q \in \mathbb{Z}} \int_0^1 [\mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H]_{n+m-pN_T, n-qN_T} e^{j2\pi(p-q)\lambda} d\lambda, \quad (20)$$

where

$$\mathbf{S}_H(\lambda) \triangleq \sum_{m \in \mathbb{Z}} \mathbf{C}_H(m) e^{-j2\pi m \lambda} \quad (21)$$

is a diagonal matrix whose k th entry, $S_H^{(k)}(\lambda)$, contains the power spectral density (PSD) of $\{\hat{h}_k(n)\}$.

1 Proof

See Appendix A. □

3 ML detection in the large array regime

This section deals with the received signal at the BS and how to exploit its massive array to obtain various insights on the detection of $\{\mathbf{X}_i \in \mathcal{X}\}$ and their design.

The received signal from (3), $\{y_i(n) \triangleq [\tilde{\mathbf{y}}|\mathbf{X}_i]_n\}$, for $n \in \{-\frac{KN_R}{2}, \dots, \frac{KN_R}{2} - 1\}$, can be expressed as

$$y_i(n) = \sum_{l=-\frac{N_R}{2}}^{\frac{N_R}{2}-1} [\mathbf{X}_i \mathbf{H}]_{n-lK, l} + [\tilde{\mathbf{z}}]_n. \quad (22)$$

Its ACF is defined as

$$C_i(n, m) \triangleq E[y_i(n+m)y_i^*(n)]. \quad (23)$$

1 Proposition 3

Under the SSW-model, the received signal at the BS is a CS process of period K , i.e. $C_i(n, m) \equiv C_i(n+rK, m)$.

1 Proof

See Appendix B. □

This result can be easily understood by considering the transformation between $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{y}}_i$, which has been illustrated in Fig. 3. By multiplying $\tilde{\mathbf{h}}$ by $\tilde{\mathbf{X}}_i$, we transform a CS process into one with period K . Then, by adding the noise term $\tilde{\mathbf{z}}$, which is WSS, the CS properties of $\tilde{\mathbf{X}}_i \tilde{\mathbf{h}}$ are preserved and $\tilde{\mathbf{y}}_i$ is a K -periodic CS process. Its K -Toeplitz covariance matrix is

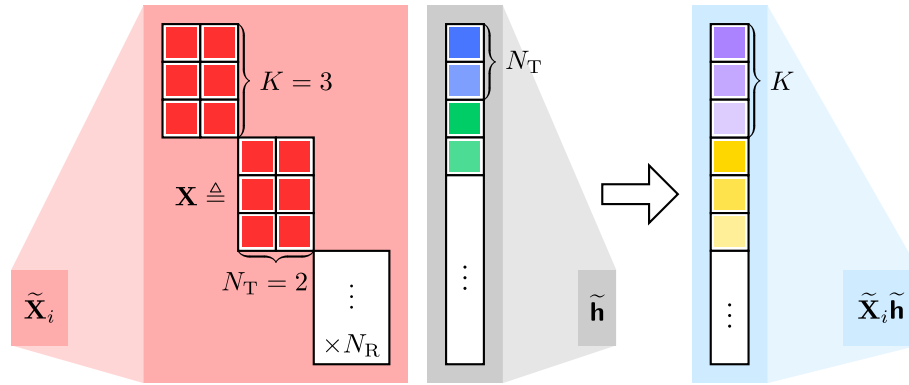


Fig. 3 Visual illustration of the change of period between $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{X}}_i \tilde{\mathbf{h}}$

$$\mathbf{C}_i \triangleq \begin{bmatrix} \mathbf{C}_i(0) & \mathbf{C}_i(-1) & \cdots & \mathbf{C}_i(-N_R + 1) \\ \mathbf{C}_i(1) & \mathbf{C}_i(0) & \cdots & \mathbf{C}_i(-N_R + 2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_i(N_R - 1) & \mathbf{C}_i(N_R - 2) & \cdots & \mathbf{C}_i(0) \end{bmatrix}, \quad (24)$$

with blocks $\mathbf{C}_i(n) \in \mathbb{C}^{K \times K}$ for $n \in \{-N_R + 1, \dots, N_R - 1\}$. Note that

$$\mathbf{C}_i(n) \equiv \mathbf{C}_i^H(-n). \quad (25)$$

Under the SSW-model, $\mathbf{C}_i(n, m)$ from (23) (see (B6)) may be developed as

$$\mathbf{C}_i(n, m) = \mathbf{P}_Z \delta_m + \sum_{l, l' = -\frac{N_R}{2}}^{\frac{N_R}{2} - 1} [\mathbf{X}_i \mathbf{U}_T \mathbf{C}_H(l - l') \mathbf{U}_T^H \mathbf{X}_i^H]_{n+m-lK, n-l'K}, \quad (26)$$

by applying (19). The obtained ACF can be asymptotically expressed as

$$\begin{aligned} \bar{\mathbf{C}}_i(n, m) &\triangleq \lim_{N_R \rightarrow \infty} \mathbf{C}_i(n, m) \\ &= \mathbf{P}_Z \delta_m + \sum_{l, l' \in \mathbb{Z}} \int_0^1 [\mathbf{X}_i \mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H \mathbf{X}_i^H]_{n+m-lK, n-l'K} e^{j2\pi(l-l')\lambda} d\lambda, \end{aligned} \quad (27)$$

where we have used (A1).

3.1 Spectral ML detector

Knowing the statistical properties of the received signal allows us to represent the ML metrics from (7) in the spectral domain asymptotically. The following formulation reaches the Cramér–Loève (CL) domain by first obtaining the Karhunen–Loève (KL) expansion of the involved signals and then applying a unitary transformation onto it [23]. The data-dependent and logarithmic terms from (8) will be studied separately. While being structurally similar to their original expressions, the spectral counterparts of (8) will feature the interplay between signal properties much more visibly and allude to the formulation of a low-complexity receiver.

3.1.1 Data-dependent term

The received signal $\{y_i(n)\}_{n \in \mathbb{Z}}$ is a second-order random process and, as such, admits an asymptotic spectral representation (as $N_R \rightarrow \infty$) known as KL expansion [34, Sec. 9.1]. It is defined as

$$y_i(n) \triangleq \int_0^1 \phi_i(n, \lambda) d\dot{y}_i(\lambda), \quad (28)$$

for $n \in \mathbb{Z}$ and $\lambda \in [0, 1)$, where $\{d\dot{y}_i(\lambda)\}$ are uncorrelated increments in the KL spectral domain and $\{\phi_i(n, \lambda)\}$ is a set of eigenfunctions that satisfy the following *orthonormality* and *completeness* relations [35, Ch. 2]:

$$\int_0^1 \phi(n, \lambda) \phi^*(n', \lambda) d\lambda = \delta_{n-n'}, \quad \sum_{n \in \mathbb{Z}} \phi(n, \lambda) \phi^*(n, \lambda') = \delta(\lambda - \lambda'). \quad (29)$$

Each class of random processes has a different eigenbasis, obtained by solving the following eigenequation:

$$\sum_{l \in \mathbb{Z}} \tilde{C}_i(l, k - l) \phi_i(l, \lambda) = S_i(\lambda) \phi_i(k, \lambda), \quad (30)$$

where $S_i(\lambda) d\lambda \triangleq E[|d\dot{y}_i(\lambda)|^2]$ is the KL spectrum of $\{y_i(n)\}$. Since we are dealing with a CS process of period K , its KL basis is

$$\phi_i^{(k)}(n, \sigma) \triangleq \phi_i\left(n, \sigma + \frac{k}{K}\right) = \sum_{q=0}^{K-1} [\mathbf{B}_i(\sigma)]_{q,k} e^{j2\pi n(\sigma + \frac{q}{K})} \quad (31)$$

for $\sigma \in [0, 1/K)$ and $k \in \{0, \dots, K-1\}$, as established in [22, 23]. These are linear combinations of complex exponentials separated by multiples of the cyclic period. They generalize the KL basis of WSS processes, which is known to be the set of Fourier eigenfunctions [34, Ch. 9]. The weights that dictate (31) are elements from $\mathbf{B}_i(\sigma) \in \mathbb{C}^{K \times K}$, whose columns are the eigenvectors of the *cyclic spectrum matrix (CSM)* of the received signal, constructed as

$$[\mathbf{S}_i(\sigma)]_{r,c} \triangleq S_i^{(\frac{r-c}{K})}\left(\sigma + \frac{r}{K}\right), \quad (32)$$

where

$$S_i^{(\frac{k}{K})}(\lambda) \triangleq \frac{1}{K} \sum_{n=0}^{K-1} \sum_{m \in \mathbb{Z}} \tilde{C}_i(n, m) e^{-j2\pi(n\frac{k}{K} + m\lambda)} \quad (33)$$

is the *cyclic spectrum* of $\{y_i(n)\}$ [34, Ch. 10]. Its eigendecomposition is then

$$\mathbf{S}_i(\sigma) \triangleq \mathbf{B}_i(\sigma) \mathcal{S}_i(\sigma) \mathbf{B}_i^H(\sigma) \in \mathbb{C}^{K \times K}, \quad (34)$$

from which we also obtain the KL spectrum contained in $\mathcal{S}_i(\sigma) \in \mathbb{R}^{K \times K}$, a diagonal matrix constructed from the eigenvalues of $\mathbf{S}_i(\sigma)$:

$$[\mathcal{S}_i(\sigma)]_{k,k} \equiv \mathcal{S}_i\left(\sigma + \frac{k}{K}\right) \triangleq \mathcal{S}_i^{(k)}(\sigma). \quad (35)$$

Recall the data-dependent quadratic term of the ML detector (8):

$$Q_j(\{y_i(n)\}) = \frac{1}{KN_R} \sum_{r,c=-\frac{KN_R}{2}}^{\frac{KN_R}{2}-1} y_i^*(r) [\mathbf{C}_j^{-1}]_{r,c} y_i(c). \quad (36)$$

The following result states its limit in the large array regime using the KL spectral representation.

1 Proposition 4

When $N_R \rightarrow \infty$, the quadratic term from (8) can be expressed in the KL spectral domain as

$$\tilde{Q}_j(\{y_i(n)\}) \triangleq \lim_{N_R \rightarrow \infty} Q_j(\{y_i(n)\}) = \int_0^{\frac{1}{K}} d\mathbf{\check{y}}_i^H(\sigma) \mathbf{B}_i^H(\sigma) \mathbf{S}_j^{-1}(\sigma) \mathbf{B}_i(\sigma) d\mathbf{\check{y}}_i(\sigma), \quad (37)$$

where $d\mathbf{\check{y}}_i(\sigma) \in \mathbb{C}^K$ is a vector constructed by stacking K samples of $\{\mathbf{\check{y}}_i(\sigma)\}$ spaced $\frac{1}{K}$, i.e.

$$d\mathbf{\check{y}}_i(\sigma) \triangleq [d\mathbf{\check{y}}_i(\sigma), d\mathbf{\check{y}}_i(\sigma + \frac{1}{K}), \dots, d\mathbf{\check{y}}_i(\sigma + \frac{K-1}{K})]^T. \quad (38)$$

1 Proof

See Appendix C. □

Notice the structural resemblances between the original quadratic form (8) and the one in the KL spectral domain (37). We can now use results from [23] to obtain it as a more compact expression. The CL representation of $\{y_i(n)\}$ is defined as

$$d\check{y}_i(\lambda) \triangleq \lim_{N_R \rightarrow \infty} \frac{1}{KN_R} \sum_{n=-\frac{KN_R}{2}}^{\frac{KN_R}{2}-1} y_i(n) e^{-j2\pi n\lambda}, \quad (39)$$

which we can stack in a K -dimensional vector $[d\check{\mathbf{y}}_i(\sigma)]_r \triangleq d\check{\mathbf{y}}_i(\sigma + \frac{r}{K})$. By [23, Corollary 1], we know that $d\check{\mathbf{y}}_i(\sigma) = \mathbf{B}_i(\sigma) d\mathbf{\check{y}}_i(\sigma)$. Therefore, (37) reduces to

$$\tilde{Q}_j(\{y_i(n)\}) = \int_0^{\frac{1}{K}} d\check{\mathbf{y}}_i^H(\sigma) \mathbf{S}_j^{-1}(\sigma) d\check{\mathbf{y}}_i(\sigma), \quad (40)$$

which is the integral of a quadratic form in the CL (Fourier) spectral domain.

It is clear that, in order to analyze (40), we need to obtain $\{\mathbf{S}_j(\sigma)\}$, which is stated next.

1 Proposition 5

The CSM of $\{y_j(n)\}$ from (22) under SSW-model is

$$\mathbf{S}_j(\sigma) = \mathbf{F}_K^H \mathbf{\Theta}^H(\sigma) \mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(K\sigma) \mathbf{U}_T^H \mathbf{X}_j^H \mathbf{\Theta}(\sigma) \mathbf{F}_K + \mathbf{P}_Z \mathbf{I}_K, \quad (41)$$

where $\mathbf{\Theta}(\sigma) \triangleq \text{Diag}(1, e^{j2\pi\sigma}, \dots, e^{j2\pi\sigma(K-1)})$.

1 Proof

See Appendix D. □

Let $\mathbf{G}_j(\sigma) \in \mathbb{C}^{K \times K}$ be the signal contribution in (41), i.e. $\mathbf{G}_j(\sigma) \triangleq \mathbf{S}_j(\sigma) - \mathbf{P}_Z \mathbf{I}_K$. Notice that its rank is $\leq N_T$. A similar idea was encountered in [15], where it was shown that linearly modulated pulse-amplitude modulation (PAM) signals exhibit a rank-1 structure in the CL domain, leaving margin for noise power estimation in the null space in the case of oversampling. Hence, a clear duality manifests between the *time* and *combined time-space* domains. For instance, in a SIMO scheme (see [21]) $\text{rank}(\mathbf{G}_j(\sigma)) = 1$, implying that $\{y_i(n)\}$ in (3) would adopt a structure in combined time-space equivalent to a PAM in the time domain. For a general MIMO case with $N_T > 1$, the signal received with the N_R antennas is being treated herein as a linear combination of PAM signals, and this is what leads to an upper bound of N_T for the rank of $\mathbf{G}_j(\sigma)$.

3.1.2 Logarithmic term

We have successfully expressed the data-dependent part of the ML detector in the CL spectral domain. To do so with the remaining logarithmic term, we will use some results on the asymptotic log-determinant of K -Toeplitz matrices [36]. It is known that

$$\ell_j \triangleq \lim_{N_R \rightarrow \infty} \ell_j = \lim_{N_R \rightarrow \infty} \frac{1}{KN_R} \ln |\mathbf{C}_j| = \frac{1}{K} \int_0^1 \ln |\dot{\mathbf{C}}_j(\lambda)| d\lambda, \quad (42)$$

where

$$\dot{\mathbf{C}}_j(\lambda) \triangleq \sum_{n \in \mathbb{Z}} \mathbf{C}_j(n) e^{-j2\pi n\lambda} \quad (43)$$

is the discrete-time Fourier transform of the $K \times K$ blocks defined in (24). For the purposes of this work, it will be referred to as *Rihaczek spectrum matrix (RSM)* [37, Sec. 1.4.1.3]. We next state how it can be obtained from a unitary transformation of the CSM of $\{y_j(n)\}$.

1 Proposition 6

The RSM and CSM of a CS process $\{y_j(n)\}$ of period K are related by

$$\mathbf{S}_j(\lambda) = \mathbf{F}_K^H \mathbf{\Theta}^H(\lambda) \dot{\mathbf{C}}_j(K\lambda) \mathbf{\Theta}(\lambda) \mathbf{F}_K. \quad (44)$$

1 Proof

By the definition of the CSM (32), we know that

$$[\mathbf{S}_j(\lambda)]_{r,c} = \frac{1}{K} \sum_{n=0}^{K-1} \sum_{l \in \mathbb{Z}} \mathbb{E}[y_j(l)y_j^*(n)] e^{j2\pi(n(\lambda + \frac{c}{K}) - l(\lambda + \frac{r}{K}))}. \quad (45)$$

With some algebraic manipulations, we can express it in terms of the RSM:

$$\begin{aligned} [\mathbf{S}_j(\lambda)]_{r,c} &= \frac{1}{K} \sum_{n,p=0}^{K-1} e^{j2\pi(n(\lambda + \frac{c}{K}) - p(\lambda + \frac{r}{K}))} \sum_{l \in \mathbb{Z}} \mathbb{E}[y_j(p + lK)y_j^*(n)] e^{-j2\pi lK(\lambda + \frac{r}{K})} \\ &= \frac{1}{K} \sum_{n,p=0}^{K-1} e^{-j2\pi p(\lambda + \frac{r}{K})} [\dot{\mathbf{C}}_j(K\lambda)]_{p,n} e^{j2\pi n(\lambda + \frac{c}{K})}. \end{aligned} \quad (46)$$

We have used the fact that, by definition, the RSM is periodic in λ with period 1. $\square$

Applying Proposition 6 onto (42), we can represent the asymptotic logarithmic term as a function of the CSM:

$$\begin{aligned} \ell_j &= \frac{1}{K} \int_0^1 \ln |\boldsymbol{\Theta}(\frac{\lambda}{K}) \mathbf{F}_K \mathbf{S}_j(\frac{\lambda}{K}) \mathbf{F}_K^H \boldsymbol{\Theta}^H(\frac{\lambda}{K})| d\lambda \\ &= \frac{1}{K} \int_0^1 \ln |\mathbf{S}_j(\frac{\lambda}{K})| d\lambda = \int_0^{\frac{1}{K}} \ln |\mathbf{S}_j(\sigma)| d\sigma. \end{aligned} \quad (47)$$

3.1.3 ML metrics in the spectral domain

Using the results derived in the previous sections, we now display the total form of the asymptotic ML metrics in the spectral domain:

$$\tilde{\mathcal{L}}_j(\{y_i(n)\}) \triangleq \lim_{N_R \rightarrow \infty} \mathcal{L}_j(\tilde{\mathbf{Y}}|\mathbf{X}_i) = \int_0^{\frac{1}{K}} d\tilde{\mathbf{Y}}_i^H(\sigma) \mathbf{S}_j^{-1}(\sigma) d\tilde{\mathbf{Y}}_i(\sigma) + \ln |\mathbf{S}_j(\sigma)| d\sigma. \quad (48)$$

Notice the structural similarities between this expression and the one from (7).

In detection theory, the Kullback–Leibler divergence (KL-Div.) between two hypotheses $\mathbf{X}_i$ and $\mathbf{X}_j$ plays a fundamental role in assessing their discrimination. It dictates the fastest error probability decay rate that can be achieved by any statistical test based on observations of $\{[\tilde{\mathbf{Y}}|\mathbf{X}_i]_n\}$ as $N_R \rightarrow \infty$ [38, Sec. 10.4]. To obtain it, we first define the normalized log-likelihood ratio (LLR) between the distributions of $\mathbf{Y}|\mathbf{X}_i$ and $\mathbf{Y}|\mathbf{X}_j$ as

$$L_{i,j}^{(KN_R)}(\tilde{\mathbf{Y}}) \triangleq \frac{1}{KN_R} \ln \frac{f_{\mathbf{Y}|\mathbf{X}_i}(\tilde{\mathbf{Y}})}{f_{\mathbf{Y}|\mathbf{X}_j}(\tilde{\mathbf{Y}})}. \quad (49)$$

Then, the normalized KL-Div is

$$\mathcal{D}_{\text{KL}}(i||j) \triangleq \mathbb{E}[L_{i,j}^{(KN_R)}(\tilde{\mathbf{Y}}|\mathbf{X}_i)] = \mathbb{E}[\mathcal{L}_j(\tilde{\mathbf{Y}}|\mathbf{X}_i) - \mathcal{L}_i(\tilde{\mathbf{Y}}|\mathbf{X}_i)], \quad (50)$$

i.e. the expectation of their LLR when $\mathbf{X}_i$ has been transmitted. It yields the value toward which $L_{i,j}^{(KN_R)}(\tilde{\mathbf{Y}}|\mathbf{X}_i)$ converges asymptotically (in the mean-square sense), so larger values of $\mathcal{D}_{\text{KL}}(i||j)$ are usually linked to improved separation of the two hypotheses.

Having derived (48), we can state the normalized KL-Div asymptotically in the CL spectral domain:

$$\begin{aligned}\bar{\mathcal{D}}_{\text{KL}}(i||j) &\triangleq \lim_{N_R \rightarrow \infty} \mathcal{D}_{\text{KL}}(i||j) = \mathbb{E} \left[\bar{\mathbf{Q}}_j(\{y_i(n)\}) - \bar{\mathbf{Q}}_i(\{y_i(n)\}) \right] + \bar{\ell}_j - \bar{\ell}_i \\ &= \int_0^{\frac{1}{K}} \mathbb{E} \left[d\check{\mathbf{y}}_i^H(\sigma) (\mathbf{S}_j^{-1}(\sigma) - \mathbf{S}_i^{-1}(\sigma)) d\check{\mathbf{y}}_i(\sigma) \right] + \ln |\mathbf{S}_j(\sigma)| - \ln |\mathbf{S}_i(\sigma)| d\sigma.\end{aligned}\quad (51)$$

The first term is obtained by using the cyclic property and linearity of the trace, as well as some identities from [23]:

$$\bar{\mathcal{D}}_{\text{KL}}(i||j) = \int_0^{\frac{1}{K}} d\sigma \left(\text{Tr}[\mathbf{S}_j^{-1}(\sigma) \mathbf{S}_i(\sigma) - \mathbf{I}_K] - \ln \frac{|\mathbf{S}_i(\sigma)|}{|\mathbf{S}_j(\sigma)|} \right). \quad (52)$$

As expected, this expression is reminiscent of the KL-Div between two zero-mean multivariate normal distributions, as noted by the structure of its integrand [38, Sec. 3.2]. Representing $\{\mathbf{S}_i(\sigma), \mathbf{S}_j(\sigma)\}$ in terms of their eigendecompositions (through the use of Proposition 5), will prove useful in Sect. 4 to examine the fundamental aspects of the codeword detection problem.

3.2 Low-complexity implementation of ML detection

The main advantage of using the spectral representation of the ML detector derived previously is that it can be approximated efficiently for finite N_R . Throughout this section, we assume the receiver is aware of $\{\mathbf{C}_j\}$ for all $\mathbf{X}_j \in \mathcal{X}$. The acquisition of this second-order statistical CSI is a topic of active research and outside the scope of this work [39]. Nevertheless, it being available at the BS is a reasonable premise due to its slow-varying nature² [28, Sec. 4.5] and the fact it can be estimated in the presence of data [40].

3.2.1 Received data

To implement the ML detector in the spectral domain, the receiver must first obtain the CL expansion of the received signal, $\{d\check{\mathbf{y}}_i(\sigma)\}$. It can be approximated by

$$\left[d\hat{\mathbf{y}}_i\left(\frac{l}{KN_R}\right) \right]_k \triangleq \frac{1}{KN_R} \text{FFT}_{KN_R} [y_i(n)](l + kN_R), \quad (53)$$

where $\text{FFT}_{KN_R}[\cdot]$ is the fast Fourier transform (FFT) of KN_R points. Therefore, the set of $\{d\hat{\mathbf{y}}_i(\frac{l}{KN_R})\}_{l \in \{0, \dots, N_R-1\}}$ can be computed with $O(KN_R \ln(KN_R))$ operations. Then, the quadratic form in (48) is approximated as

$$\hat{\mathbf{Q}}_j(\{y_i(n)\}) \triangleq \sum_{l=0}^{N_R-1} d\hat{\mathbf{y}}_i^H\left(\frac{l}{KN_R}\right) \hat{\mathbf{S}}_j^{-1}\left(\frac{l}{KN_R}\right) d\hat{\mathbf{y}}_i\left(\frac{l}{KN_R}\right) \quad (54)$$

with $O(N_R K^2)$ operations, where $\hat{\mathbf{S}}_j(\frac{l}{KN_R})$ is an estimation of $\mathbf{S}_j(\frac{l}{KN_R})$.

² This long-term CSI includes those parameters related to the channel geometry.

3.2.2 Statistics

Beside processing the received data with (53) and (54), approximating the ML metrics in spectral form requires the computation of various statistical structures from $\{\mathbf{C}_j\}$. While direct implementation of the ML detector (7) demands $\{\mathbf{C}_j^{-1}\}$, with a cost of $O(K^3 N_R^3)$ operations for each $\mathbf{X}_j \in \mathcal{X}$, the spectral ML approximation uses inverse CSMs.

The set of $\{\hat{\mathbf{S}}_j^{-1}(\frac{l}{KN_R})\}$ can be obtained as follows. Recall the RSM is computed as

$$\dot{\mathbf{C}}_j(\lambda) = \sum_{n=0}^{\infty} \mathbf{C}_j(n) e^{-j2\pi n\lambda} + \sum_{n'=-\infty}^{-1} \mathbf{C}_j(n') e^{-j2\pi n'\lambda}. \quad (55)$$

With some simple algebraic manipulations and considering (25), it may be obtained as

$$\dot{\mathbf{C}}_j(\lambda) = \sum_{n=0}^{\infty} \mathbf{C}_j(n) e^{-j2\pi n\lambda} + \left(\sum_{n'=0}^{\infty} \mathbf{C}_j(n') e^{-j2\pi n'\lambda} \right)^H - \mathbf{C}_j(0). \quad (56)$$

Therefore, it is straightforward to see it can be approximated as

$$\hat{\mathbf{C}}_j(\frac{l}{N_R}) \triangleq \text{FFT}_{N_R}[\mathbf{C}_j(n)](l) + \text{FFT}_{N_R}[\mathbf{C}_j(n)]^H(l) - \mathbf{C}_j(0), \quad (57)$$

with which the CSM is estimated as

$$\hat{\mathbf{S}}_j(\frac{l}{KN_R}) \triangleq \mathbf{F}_K^H \boldsymbol{\Theta}^H(\frac{l}{KN_R}) \hat{\mathbf{C}}_j(\frac{l}{N_R}) \boldsymbol{\Theta}(\frac{l}{KN_R}) \mathbf{F}_K, \quad (58)$$

based on Proposition 6. Notice that this expression yields a hermitian matrix but does not guarantee it to be positive semi-definite³ in general for finite N_R . Therefore, following the same idea behind the *Blackman-Tukey spectral estimator* [41, Sec. 4.5], we weight the correlation blocks in (57) with a triangular window:

$$\hat{\mathbf{C}}_j^{(\text{BT})}(\frac{l}{N_R}) = \text{FFT}_{N_R} \left[\left(1 - \frac{n}{N_R}\right) \mathbf{C}_j(n) \right](l) + \text{FFT}_{N_R} \left[\left(1 - \frac{n}{N_R}\right) \mathbf{C}_j(n) \right]^H(l) - \mathbf{C}_j(0). \quad (59)$$

In summary, computing $\{\hat{\mathbf{S}}_j^{-1}(\frac{l}{KN_R})\}_{l \in \{0, \dots, N_R-1\}}$ involves $O(K^2 N_R \ln N_R)$ operations for the RSM, plus $O(K^3 N_R)$ additional for (58) and the matrix inversions. Fortunately, these statistics can be pre-calculated at the receiver before data transmission and are valid for the detection of a certain number T of consecutive codewords, in which the channel statistics remain stable [28, Sec. 4.5]. The same can be said for the logarithmic terms, whose direct computation involves $O(K^3 N_R^3)$ operations each, whereas their spectral approximations

$$\hat{\ell}_j = \sum_{l=0}^{N_R-1} \frac{1}{KN_R} \ln |\hat{\mathbf{S}}_j(\frac{l}{KN_R})| \quad (60)$$

are computed with just $O(K^3 N_R)$.

³ We know from [23] that $\mathbf{S}_j(\sigma)$ is a covariance matrix and thus, positive semi-definite.

Table 1 Computational complexity of ML detection

Once every T codewords		
	Direct ML	Spectral ML approx
$\{\mathbf{C}_j^{-1}\} \mid \{\mathbf{S}_j^{-1}(\sigma)\}$	$O(MK^3N_R^3)$	$O(MK^2N_R(\ln N_R + K))$
$\{\ell_j\}$	$O(MK^3N_R^3)$	$O(MK^3N_R)$
Once for each codeword detection		
	Direct ML	Spectral ML approx
$\{Q_j\}$	$O(MK^2N_R^2)$	$O(KN_R(\ln(KN_R) + MK))$
ML detection	$O(M)$	$O(M)$

3.2.3 Approximation validity

Table 1 contains a summary of the computational complexity associated with each implementation. The advantages of using the spectral approximation are clear when employing large arrays at the BS, since the number of operations to compute both data statistics and ML metrics scale slower with N_R . Nevertheless, this approach is based on an asymptotic representation of the ML detector as $N_R \rightarrow \infty$, and its validity for finite N_R remains to be seen.

The performance of the spectral ML approximation described in the previous section will be assessed next against the regular ML detector from Sect. 2.2. We consider a channel under SSW-model, whose ACF is constructed from (18), by setting

$$[\mathbf{C}_H(m)]_{k,k} \triangleq \frac{1}{k+2} \text{sinc}^2\left(\frac{m}{k+2}\right), \quad k = 0, \dots, N_T - 1. \quad (61)$$

This yields a set of spectra $\{S_H^{(k)}(\lambda)\}$ with triangular shape of widths $\{\frac{1}{k+2}\}$. This fading profile has been adopted motivated by two main reasons. On the one hand, it represents a simple model constructed from piece-wise linear functions of compact support. On the other hand, it characterizes a sufficiently well-behaved, worst-case scenario. Since triangular functions belong to differentiability class C^0 , the correlation drop-off between adjacent antennas is assumed to decay quadratically, much slower than one corresponding to smooth spectra [42, Sec. 4.4.5]. Note, however, that $\{S_H^{(k)}(\lambda) \in C^0\}$ avoids the occurrence of disturbing artifacts such as Gibbs phenomenon.

The channel correlation matrix is then

$$\mathbf{C}_{\tilde{H}} = \eta \cdot (\mathbf{I}_{N_R} \otimes \mathbf{U}_T) \begin{bmatrix} \mathbf{C}_H(0) & \mathbf{C}_H(1) & \cdots & \mathbf{C}_H(N_R - 1) \\ \mathbf{C}_H(1) & \mathbf{C}_H(0) & \cdots & \mathbf{C}_H(N_R - 2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_H(N_R - 1) & \mathbf{C}_H(N_R - 2) & \cdots & \mathbf{C}_H(0) \end{bmatrix} (\mathbf{I}_{N_R} \otimes \mathbf{U}_T^H), \quad (62)$$

where η is a power normalization parameter that ensures that $\text{Tr}[\mathbf{C}_{\tilde{H}}] \triangleq N_T N_R$, as stated in the signal model from Sect. 2.1. The transmitter basis $\mathbf{U}_T$ is selected randomly from

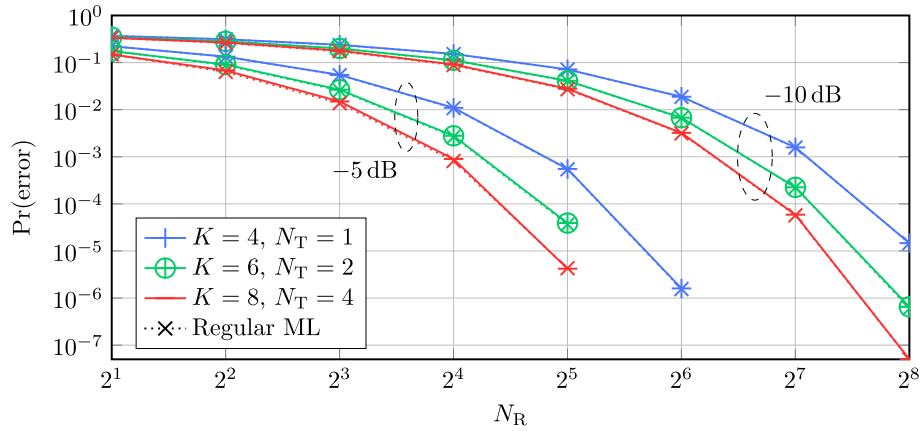


Fig. 4 Error probability in terms of N_R , for various K and N_T at $\text{SNR} = -10$ dB and $\text{SNR} = -5$ dB

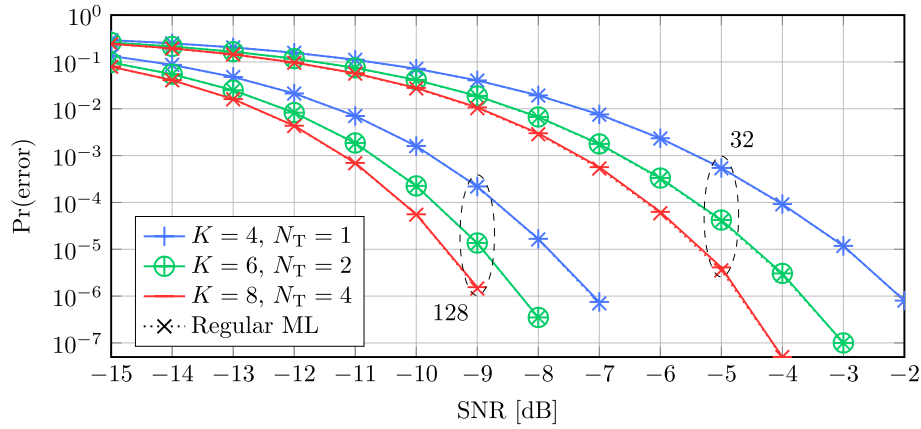


Fig. 5 Error probability in terms of SNR, for various K and N_T , with $N_R = 32$ and $N_R = 128$

the set of unitary $N_T \times N_T$ matrices [43]. Noise power is $\frac{1}{K} E[\|\mathbf{Z}\|_F^2] = P_Z N_R$. SNR at the receiver is defined as

$$\text{SNR} \triangleq \frac{E[\|\mathbf{X}\mathbf{H}\|_F^2]}{E[\|\mathbf{Z}\|_F^2]} = \frac{\text{Tr}[\mathbf{C}_h E[\tilde{\mathbf{X}}^H \tilde{\mathbf{X}}]]}{P_Z K N_R}. \quad (63)$$

Setting $K \geq 2N_T$, the pair of tested codewords is constructed as follows. An element $\mathbf{U} \in \mathbb{C}^{K \times K}$ is selected randomly from the set of unitary $K \times K$ matrices. Its first N_T columns are named $\mathbf{U}_i \in \mathbb{C}^{K \times N_T}$, and the following N_T ones are named $\mathbf{U}_j \in \mathbb{C}^{K \times N_T}$. Then, codewords are $\mathbf{X}_i \triangleq \sqrt{\frac{K}{N_T}} \mathbf{U}_i$ and $\mathbf{X}_j \triangleq \sqrt{\frac{K}{2N_T}} \mathbf{U}_j$, thus differing both in power (1 and 1/2, respectively) and spanned subspace (they are orthogonal), as to illustrate the effects caused by both discrepancies. Under this transmission scheme, (63) simplifies to $\text{SNR} = \frac{3/4}{P_Z}$.

Figure 4 and Fig. 5 assess the performance of detectors from Sect. 2.2 (Regular ML) and Sect. 3.2 (Spectral ML) for different configurations of K and N_T , at various

regimes of N_R and SNR. The metric of choice is the average error probability in the binary detection problem involving $\mathbf{X}_i$ and $\mathbf{X}_j$, computed as

$$\Pr(\text{error}) \triangleq \frac{\Pr(\text{detect } \mathbf{X}_j | \text{transmit } \mathbf{X}_i) + \Pr(\text{detect } \mathbf{X}_i | \text{transmit } \mathbf{X}_j)}{2}. \quad (64)$$

It has been averaged over 2×10^7 Monte Carlo simulations.

On the one hand, Fig. 4 contains the detection error probability for increasing values of N_R , at SNR = −10 dB and SNR = −5 dB. All six curves display similar decreasing trends, with overall better results for higher K and N_T . Remarkably, the difference between regular and spectral ML detectors is diminutive, even for moderate receiver arrays. It is most noticeable for $N_R < 2^5$ in the higher SNR regime. On the other hand, Fig. 5 shows the same metric in terms of SNR, for $N_R = 32$ and $N_R = 128$. The performance loss incurred from using the spectral approximation is only noticeable for the smaller array at lower error probability values (higher SNR). Based on these results, we can state the low-complexity detector developed through Sect. 3.2 can be used instead of (7) in a wide range of configurations (particularly with large arrays) with negligible performance degradation.

4 Codeword structure analysis

In this section, we will get some final insights on the problem considered in this work, as well as connect it to various ideas from the literature on detection theory. The figure of merit will be the asymptotic normalized KL-Div, whose relevance in evaluating hypothesis testing has been discussed in Sect. 3.1.3. More precisely, it will be linked to various results on singular (i.e. error-free) detection: understanding if two codewords can be detected unmistakably allows to characterize the existence of error floors under various asymptotic regimes. This is of particular interest when assessing the gains achievable by pouring more resources into a system.

We study the transmission of structured codewords within the presented spectral domain. Each transmitted codeword, taken from M -ary alphabet $\mathcal{X}$, can be constructed from its singular value decomposition as $\mathbf{X}_i \triangleq \Phi_i \mathbf{W}_i^{\frac{1}{2}} \Psi_i^H$, where:

- $\Phi_i \in \mathbb{C}^{K \times K}$ is a unitary matrix.
- $\mathbf{W}_i^{\frac{1}{2}} \in \mathbb{R}^{K \times N_T}$ is a power-loading diagonal matrix with nonnegative real entries $\{[\mathbf{W}_i^{\frac{1}{2}}]_{p,p} \triangleq (w_i^{(p)})^{\frac{1}{2}}\}_{p=0,\dots,N_T-1}$. The first $N_i \leq N_T$ ones are non-null (*active*), while the rest are 0 (*inactive*).
- $\Psi_i \in \mathbb{C}^{N_T \times N_T}$ is a unitary precoding matrix.

The rank of $\mathbf{X}_i$ is N_i , as stated by the definition of $\mathbf{W}_i^{\frac{1}{2}}$. Therefore, $\Phi_i \mathbf{W}_i^{\frac{1}{2}}$ spans a N_i -dimensional subspace V_i . This formulation allows us to straightforwardly get the CSM basis $\mathbf{B}_i(\sigma)$ in some relevant scenarios that will be explored next.

4.1 Statistical CSI at the transmitter

When the transmitter is aware of $\mathbf{U}_T$ in (9) (i.e. the channel geometry), it can implement a precoding suitable to the fading correlation geometry: $\Psi_i \triangleq \mathbf{U}_T$, which is to be understood as transmitting along the channel basis. This reduces the difficulty of

eigendecomposing $\mathbf{S}_i(\sigma)$ immensely. We can clearly see from (41) that its eigenbasis and KL spectrum are

$$\mathbf{B}_i(\sigma) = \mathbf{F}_K^H \mathbf{\Theta}^H(\sigma) \mathbf{\Phi}_i, \quad \mathbf{S}_i(\sigma) = \mathbf{W}_i^{\frac{1}{2}} \mathbf{S}_H(K\sigma) (\mathbf{W}_i^{\frac{1}{2}})^H + \mathbf{P}_Z \mathbf{I}_K. \quad (65)$$

The normalized KL-Div resulting from using this configuration onto (52) is

$$\tilde{\mathcal{D}}_{\text{KL}}(i||j) = \int_0^{\frac{1}{K}} d\sigma \left(\text{Tr}[\mathbf{\Phi}_j \mathbf{S}_j^{-1}(\sigma) \mathbf{\Phi}_j^H \mathbf{\Phi}_i \mathbf{S}_i(\sigma) \mathbf{\Phi}_i^H - \mathbf{I}_K] - \ln \frac{|\mathbf{S}_i(\sigma)|}{|\mathbf{S}_j(\sigma)|} \right). \quad (66)$$

In the case in which $\mathbf{\Phi}_i = \mathbf{\Phi}_j$, i.e. both codewords use the same signal basis, information is purely modulated on the transmitted energy through $\mathbf{W}_i$ and $\mathbf{W}_j$. The asymptotic normalized KL-Div simplifies into

$$\tilde{\mathcal{D}}_{\text{KL}}(i||j) = \int_0^{\frac{1}{K}} d\sigma \left(\text{Tr}[\mathbf{S}_j^{-1}(\sigma) \mathbf{S}_i(\sigma) - \mathbf{I}_K] - \ln \frac{|\mathbf{S}_i(\sigma)|}{|\mathbf{S}_j(\sigma)|} \right) \triangleq \sum_{k=0}^{N_T-1} \mathcal{D}_{\text{IS}}^{(k)}(\mathcal{S}_i || \mathcal{S}_j), \quad (67)$$

where

$$\mathcal{D}_{\text{IS}}^{(k)}(\mathcal{S}_i || \mathcal{S}_j) \triangleq \int_0^{\frac{1}{K}} d\sigma \left(\frac{w_i^{(k)} \mathbf{S}_H^{(k)}(K\sigma) + \mathbf{P}_Z}{w_j^{(k)} \mathbf{S}_H^{(k)}(K\sigma) + \mathbf{P}_Z} - 1 - \ln \frac{w_i^{(k)} \mathbf{S}_H^{(k)}(K\sigma) + \mathbf{P}_Z}{w_j^{(k)} \mathbf{S}_H^{(k)}(K\sigma) + \mathbf{P}_Z} \right). \quad (68)$$

This is the *Itakura-Saito divergence* (IS-Div.) [38, Sec. 10.4.2] between $\mathcal{S}_i^{(k)}(\sigma)$ and $\mathcal{S}_j^{(k)}(\sigma)$ in the k th KL sub-band. This measure of dissimilarity is nonnegative and only null when $\mathcal{S}_i^{(k)}(\sigma) = \mathcal{S}_j^{(k)}(\sigma)$ (except in a support of measure 0). Moreover, (67) is the IS-Div between the full KL spectra corresponding to $\{y_i(n)\}$ and $\{y_j(n)\}$:

$$\mathcal{D}_{\text{IS}}(\mathcal{S}_i || \mathcal{S}_j) \triangleq \int_0^1 d\lambda \left(\frac{\mathcal{S}_i(\lambda)}{\mathcal{S}_j(\lambda)} - 1 - \ln \frac{\mathcal{S}_i(\lambda)}{\mathcal{S}_j(\lambda)} \right). \quad (69)$$

It is of interest to study the tendency of (68) in the high SNR regime, i.e. as $\mathbf{P}_Z \rightarrow 0$ (see (63)), since it reveals the properties of $\mathbf{X}_i$ and $\mathbf{X}_j$ that are relevant for their detection asymptotically. Given $N_j > N_i$, the corresponding IS-Div in the k th KL sub-band for $N_i \leq k < N_j$ is

$$\lim_{\mathbf{P}_Z \rightarrow 0} \mathcal{D}_{\text{IS}}^{(k)}(\mathcal{S}_i || \mathcal{S}_j) \geq -\frac{1}{K} - \int_0^{\frac{1}{K}} d\sigma \left(\lim_{\mathbf{P}_Z \rightarrow 0} \ln \frac{\mathbf{P}_Z}{w_j^{(k)} \mathbf{S}_H^{(k)}(K\sigma) + \mathbf{P}_Z} \right) = +\infty, \quad (70)$$

where we have applied Fatou's Lemma [44, Lemma B.2.4] in order to swap the order of limit and integration. The same divergence is observed when $N_i > N_j$ for $N_j \leq k < N_i$. On the contrary, when $N_i = N_j$,

$$\lim_{\mathbf{P}_Z \rightarrow 0} \tilde{\mathcal{D}}_{\text{KL}}(i||j) = \frac{1}{K} \sum_{k=0}^{N_i-1} \frac{w_i^{(k)}}{w_j^{(k)}} - 1 - \ln \frac{w_i^{(k)}}{w_j^{(k)}} \quad (71)$$

by Lebesgue's Dominated Convergence Theorem (LDCT) [44, Th. 11.3.13]. These results are consistent with [26], which states that singular detection is achievable in the high

SNR regime if and only if $\mathbf{X}_i$ and $\mathbf{X}_j$ span different column spaces. Within our framework, this property translates into (70) diverging, since $\mathbf{X}_i$ and $\mathbf{X}_j$ belong to different subspaces, whereas (71) is bounded due to both codewords spanning the same column space.

A similar analysis can be performed for the low SNR regime, i.e. $P_{\mathbf{Z}} \rightarrow \infty$ (see (63)). Using LDCT once again, we obtain the following second-order Taylor approximation:

$$\lim_{P_{\mathbf{Z}} \rightarrow \infty} \bar{\mathcal{D}}_{\text{KL}}(i||j) = \lim_{P_{\mathbf{Z}} \rightarrow \infty} \frac{1}{P_{\mathbf{Z}}} \sum_{k=0}^{N_T-1} (w_i^{(k)} - w_j^{(k)})^2 \frac{1}{2} \int_0^{\frac{1}{K}} (\mathbf{S}_{\mathbf{H}}^{(k)}(K\sigma))^2 d\sigma. \quad (72)$$

By [45, Prop. 6.A.3.], sorting $\{w_i^{(k)}\}$ in descending order and $\{w_j^{(k)}\}$ in ascending order (or vice-versa) yields the maximum value of (72) asymptotically. Remarkably, this strategy is also optimal in (71) for the high SNR regime: since it is a Schur-convex function of $\{w_i^{(k)}/w_j^{(k)}\}$ [45, Prop. 3.C.1.], it is maximized when $\{w_i^{(k)}\}$ and $\{w_j^{(k)}\}$ are arranged in opposite orders [45, Prop. 6.A.1.]. This idea suggests a design approach on the power profiles of $\mathcal{X}$ based on the joint majorization properties of $\{\{w_i^{(p)}\}, \{w_j^{(p')}\}\}$, for every pair $\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}$.

4.2 No CSI at the transmitter

We now consider the case in which the UE is not aware of the channel basis nor spectrum. When the transmitter is not aware of the channel statistics, it is reasonable to set an agnostic precoding ($\Psi_i \triangleq \mathbf{I}_{N_T}$) and no energy modulation in the power-loading matrix ($\{(w_i^{(p)})^{\frac{1}{2}} = 1\}_{p=0, \dots, N_i-1}$), so that no radiated direction is favored. For a clearer exposition, we set $N_i \triangleq N_T$ for all $\mathbf{X}_i \in \mathcal{X}$. This yields $\mathcal{X}$ to be a set of semi-unitary matrices:

$$\mathbf{X}_i \mathbf{X}_i^H \triangleq \mathbf{P}_i, \quad \mathbf{X}_i^H \mathbf{X}_i = \mathbf{I}_{N_T}, \quad (73)$$

where $\mathbf{P}_i \in \mathbb{C}^{K \times K}$ is the projection matrix onto V_i . If every codeword in $\mathcal{X}$ spans a distinct subspace, this alphabet is known as a *Grassmannian constellation*, since each V_i corresponds to a different point on the Grassmann manifold $\mathcal{G}(N_T, \mathbb{C}^K)$ [46].

Under these premises, the eigendecomposition of the CSM of $\{y_i(n)\}$ is

$$\mathbf{B}_i(\sigma) = \mathbf{F}_i^H \mathbf{\Theta}^H(\sigma) [\mathbf{U}_i, \mathbf{U}_{\bar{i}}], \quad \mathbf{S}_i(\sigma) = \text{Diag}(\mathbf{S}_{\mathbf{H}}(K\sigma), \mathbf{0}_{K-N_T}) + P_{\mathbf{Z}} \mathbf{I}_K, \quad (74)$$

where $\mathbf{U}_i \triangleq \Phi_i \mathbf{W}_i \mathbf{U}_T$ and $\mathbf{U}_{\bar{i}} \in \mathbb{C}^{K \times (K-N_T)}$ spans the orthogonal complement to V_i (its corresponding projector is $\mathbf{P}_{\bar{i}} \triangleq \mathbf{U}_{\bar{i}} \mathbf{U}_{\bar{i}}^H$). It is clear that $\mathbf{S}_i(\sigma)$ is common for all $\mathbf{X}_i$, so the logarithmic term in (52) is null. By applying (74) onto (52), the resulting KL-Div is

$$\bar{\mathcal{D}}_{\text{KL}}(i||j) = \int_0^{\frac{1}{K}} d\sigma \text{Tr}[\mathbf{S}_i^{-1}(\sigma) [\mathbf{U}_j, \mathbf{U}_{\bar{j}}]^H [\mathbf{U}_i, \mathbf{U}_{\bar{i}}] \mathbf{S}_i(\sigma) [\mathbf{U}_i, \mathbf{U}_{\bar{i}}]^H [\mathbf{U}_j, \mathbf{U}_{\bar{j}}] - \mathbf{I}_K]. \quad (75)$$

With simple algebraic manipulations, this is developed as

$$\bar{\mathcal{D}}_{\text{KL}}(i||j) = \frac{1}{K} \int_0^1 \text{Tr}[\Sigma^{-1}(\lambda) \mathbf{U}_j^H \mathbf{U}_i \Sigma(\lambda) \mathbf{U}_i^H \mathbf{U}_j] d\lambda + \frac{1}{K} \|\mathbf{U}_{\bar{j}}^H \mathbf{U}_{\bar{i}}\|_{\text{F}}^2 - 1 \quad (76a)$$

$$+ \frac{1}{K} \int_0^1 d\lambda (P_{\mathbf{Z}} \text{Tr}[\mathbf{U}_j \Sigma^{-1}(\lambda) \mathbf{U}_j^H \mathbf{P}_{\bar{i}}] + \frac{1}{P_{\mathbf{Z}}} \text{Tr}[\mathbf{U}_{\bar{j}}^H \mathbf{U}_i \Sigma(\lambda) \mathbf{U}_i^H \mathbf{U}_{\bar{j}}]) \quad (76b)$$

where $\Sigma(\lambda) \triangleq \mathbf{S}_{\mathbf{H}}(\lambda) + \mathbf{P}_{\mathbf{Z}}\mathbf{I}_{N_T}$. The first term in (76a) compares the signal subspaces between hypotheses i and j , whereas the second term compares the noise subspaces. The two terms in (76b) compare the crossed subspaces.

This new expression allows to analyze the behavior of the KL-Div in the high and low SNR regimes as in Sect. 4.1. When $\mathbf{P}_{\mathbf{Z}} \rightarrow 0$, the previous KL-Div becomes

$$\begin{aligned} \lim_{\mathbf{P}_{\mathbf{Z}} \rightarrow 0} \mathcal{D}_{\text{KL}}^{\infty}(i||j) &= \frac{1}{K} \int_0^1 \text{Tr}[\mathbf{S}_{\mathbf{H}}^{-1}(\lambda) \mathbf{U}_j^H \mathbf{U}_i \mathbf{S}_{\mathbf{H}}(\lambda) \mathbf{U}_i^H \mathbf{U}_j] d\lambda + \frac{1}{K} \|\mathbf{U}_j^H \mathbf{U}_i\|_{\text{F}}^2 - 1 \\ &+ \lim_{\mathbf{P}_{\mathbf{Z}} \rightarrow 0} \frac{1}{K \mathbf{P}_{\mathbf{Z}}} \text{Tr}[\mathbf{U}_j^H \mathbf{U}_i (\int_0^1 \mathbf{S}_{\mathbf{H}}(\lambda) d\lambda) \mathbf{U}_i^H \mathbf{U}_j]. \end{aligned} \quad (77)$$

Therefore, it will diverge as long as $V_i \neq V_j$ which is consistent with singular detection conditions in the high SNR regime [26]. On the contrary, when $\mathbf{P}_{\mathbf{Z}} \rightarrow \infty$, the second-order Taylor approximation of the KL-Div is a function of correlations of spectra, similarly to what was observed in Sect. 4.1 (see (72)). This implies high spectral correlation generally benefits detection when noise obscures the signal of interest, which is a fairly common behavior displayed at low SNR by classical techniques like the *estimator-correlator* [47].

5 Concluding remarks

This work has presented a novel approach to analyze a noncoherent uplink massive MIMO system, by exploiting the cyclostationarity emerging in the joint treatment of the time and space domains. The use of large arrays has allowed the derivation of asymptotic representations of the received signals, from which relevant attributes for codeword detection become manifest. As a result of this theoretical study, a low-complexity implementation of the ML detector has been inspired, which benefits from the low dimensional spectral correlation structures.

A straightforward extension of the presented work may involve examining the down-link scenario using similar techniques as the ones described herein. Although the signal received by the UE does not display a CS nature, a representation in terms of the channel spectrum would be accessible (due to the massive array at the BS) and provide insights on the problem, as well as potentially lead to simplified precoding strategies based on the FFT in the space domain. Another clear line of research may employ the asymptotic framework developed in this paper to design alphabets tailored to specific fading profiles. The limit expressions obtained in the high and low SNR regimes provide particularly useful design criteria. Additionally, the study of the spectral interplay between users in a multiple access setting may provide deep and novel perspectives on challenging problems.

Appendix A Proof of Proposition 2

From (19), recall the expression for the k th diagonal entry of $\mathbf{C}_{\mathbf{H}}(m)$, which is a Darboux sum. It converges to an integral by defining $S_{\mathbf{H}}^{(k)}(\frac{l}{N_R}) \triangleq \gamma_{k,l}$ and setting $\frac{l}{N_R} \rightarrow \lambda \in [0, 1)$ and $\frac{1}{N_R} \rightarrow d\lambda$:

$$\lim_{N_R \rightarrow \infty} [\mathbf{C}_H(m)]_{k,k} = \int_0^1 S_H^{(k)}(\lambda) e^{j2\pi m\lambda} d\lambda. \quad (\text{A1})$$

Since $\{\hat{h}_k(n)\}$ is a WSS process, by the Wiener-Khinchin theorem [34, Sec. 9.2.2], we know that $S_H^{(k)}(\lambda)$ is in fact its PSD. By defining $\mathbf{S}_H(\lambda) \triangleq \text{Diag}(\{S_H^{(k)}(\lambda)\}_{k \in \{0, \dots, N_T-1\}})$ as in (21), the resulting ACF from (18) can be expressed in the limit as

$$\lim_{N_R \rightarrow \infty} \mathbf{C}_H(n, m) = \sum_{p, q \in \mathbb{Z}} \int_0^1 [\mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H]_{n+m-pN_T, n-qN_T} e^{j2\pi(p-q)\lambda} d\lambda. \quad (\text{A2})$$

Appendix B Proof of Proposition 3

From (23), the ACF of $\{y_i(n)\}$ is

$$C_i(n, m) = \sum_{l, l' = -\frac{N_R}{2}}^{\frac{N_R}{2}-1} \sum_{k, k' = 0}^{N_T-1} [\mathbf{X}_i]_{n+m-lK, k} [\mathbf{X}_i]_{n-l'K, k'}^* E[\mathbf{H}]_{k, l} [\mathbf{H}]_{k', l'}^* + \mathbf{P}_Z \delta_m. \quad (\text{B3})$$

We have used the fact that $\mathbf{H}$ and $\mathbf{Z}$ are independent and the ACF of $\{\tilde{\mathbf{Z}}_n\}$ from its covariance matrix in Sect. 2.1. For $r \in \mathbb{Z}$,

$$C_i(n + rK, m) = \sum_{l, l', k, k'} [\mathbf{X}_i]_{n+m-(l-r)K, k} [\mathbf{X}_i]_{n-l'K, k'}^* E[\mathbf{H}]_{k, l} [\mathbf{H}]_{k', l'}^* + \mathbf{P}_Z \delta_m. \quad (\text{B4})$$

Then, by setting $s = l - r$ and $s' = l' - r$, such that $s, s' \in \{-\frac{N_R}{2} - r, \dots, \frac{N_R}{2} - 1 - r\}$,

$$C_i(n + rK, m) = \sum_{s, s'} \sum_{k, k'} [\mathbf{X}_i]_{n+m-sK, k} [\mathbf{X}_i]_{n-s'K, k'}^* E[\mathbf{H}]_{k, s+r} [\mathbf{H}]_{k', s'+r}^* + \mathbf{P}_Z \delta_m. \quad (\text{B5})$$

Finally, by the spatial stationarity of $\mathbf{H}$ (refer to (10)), we can see that

$$\begin{aligned} C_i(n + rK, m) &= \sum_{s, s' = -\frac{N_R}{2}}^{\frac{N_R}{2}-1} \sum_{k, k'} [\mathbf{X}_i]_{n+m-sK, k} [\mathbf{X}_i]_{n-s'K, k'}^* E[\mathbf{H}]_{k, s} [\mathbf{H}]_{k', s'}^* + \mathbf{P}_Z \delta_m \\ &\equiv C_i(n, m). \end{aligned} \quad (\text{B6})$$

Appendix C Proof of Proposition 4

Let $\Omega_j(n, m) \triangleq \lim_{N_R \rightarrow \infty} [\mathbf{C}_j^{-1}]_{n+m, n}$ be the *precision function* of $\{y_j(n)\}$. By definition, we know that $[\mathbf{C}_j \mathbf{C}_j^{-1}]_{r, c} = [\mathbf{C}_j^{-1} \mathbf{C}_j]_{r, c} \triangleq \delta_{r-c}$. Therefore, in the limit of $N_R \rightarrow \infty$,

$$\lim_{N_R \rightarrow \infty} [\mathbf{C}_j \mathbf{C}_j^{-1}]_{r, c} = \sum_{k \in \mathbb{Z}} \tilde{\mathbf{C}}_j(k, r - k) \Omega_j(c, k - c) \triangleq \delta_{r-c} \quad (\text{C7})$$

must hold. From (29) and (30) we know that

$$\tilde{\mathcal{C}}_j(c, r - c) = \int_0^1 \phi_j(r, \lambda) \mathcal{S}_j(\lambda) \phi_j^*(c, \lambda) d\lambda. \quad (\text{C8})$$

It can then be verified that

$$\Omega_j(c, r - c) = \int_0^1 \phi_j(r, \lambda) \mathcal{S}_j^{-1}(\lambda) \phi_j^*(c, \lambda) d\lambda \quad (\text{C9})$$

does indeed fulfill (C7).

We may express the quadratic form (36) in the limit as

$$\tilde{\mathcal{Q}}_j(\{y_i(n)\}) = \sum_{r, c \in \mathbb{Z}} y_i^*(r) \Omega_j(c, r - c) y_i(c) d\sigma, \quad (\text{C10})$$

where $\frac{1}{KN_R} \rightarrow d\sigma$. Using (C9) and the KL expansion of $\{y_i(n)\}$ (28), the previous expression can be represented in the KL spectral domain:

$$\tilde{\mathcal{Q}}_j(\{y_i(n)\}) = d\sigma \iint \int_0^1 d\mathbf{y}_i^*(\alpha) \frac{\langle \phi_i, \phi_j \rangle(\alpha, \lambda) \langle \phi_j, \phi_i \rangle(\lambda, \beta)}{\mathcal{S}_j(\lambda)} d\mathbf{y}_i(\beta) d\lambda, \quad (\text{C11})$$

where

$$\langle \phi_a, \phi_b \rangle(\alpha, \beta) \triangleq \sum_{n \in \mathbb{Z}} \phi_a^*(n, \alpha) \phi_b(n, \beta) \quad (\text{C12})$$

is the inner product between two basis components. Since $\{y_i(n)\}$ and $\{y_j(n)\}$ are both CS processes, it can be particularized as

$$\langle \phi_a^{(k)}, \phi_b^{(k')} \rangle(\sigma, \sigma') \triangleq \langle \phi_a, \phi_b \rangle\left(\sigma + \frac{k}{K}, \sigma' + \frac{k'}{K}\right) = \delta(\sigma - \sigma') [\mathbf{B}_a^H(\sigma) \mathbf{B}_b(\sigma')]_{k, k'}. \quad (\text{C13})$$

This yields

$$\tilde{\mathcal{Q}}_j(\{y_i(n)\}) = \sum_{k, m} \int_0^{\frac{1}{K}} [\mathbf{B}_i^H(\sigma) \mathcal{S}_j^{-1}(\sigma) \mathbf{B}_i(\sigma)]_{k, m} d\mathbf{y}_i^*\left(\sigma + \frac{k}{K}\right) d\mathbf{y}_i\left(\sigma + \frac{m}{K}\right), \quad (\text{C14})$$

where we have used (34) and the sifting property of the Dirac delta distribution.

Appendix D Proof of Proposition 5

Recall the cyclic spectrum from (33). We compute it for $\{y_j(n)\}$ by using (27):

$$\begin{aligned} \mathcal{S}_j^{(k)}(\sigma) &= \frac{1}{K} \sum_{n'=0}^{K-1} \sum_{m' \in \mathbb{Z}} \mathbf{P}_z \delta_{m'} e^{-j2\pi(n' \frac{k}{K} + m'\sigma)} + \frac{1}{K} \sum_{n=0}^{K-1} e^{-j2\pi n \frac{k}{K}} \sum_{m \in \mathbb{Z}} e^{-j2\pi m \sigma} \\ &\quad \times \sum_{l \in \mathbb{Z}} \int_0^1 [\mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H \mathbf{X}_j^H]_{n+m-lK, n} e^{j2\pi l \lambda} d\lambda. \end{aligned} \quad (\text{D15})$$

Notice we have set $l' = 0$ since $n \in \{0, \dots, K-1\}$. With $p = n + m - lK$, for $p \in \{0, \dots, K-1\}$, we have

$$S_j^{(\frac{k}{K})}(\sigma) = \frac{P_Z}{K} \sum_{n=0}^{K-1} e^{-j2\pi n' \frac{k}{K}} + \frac{1}{K} \sum_{n,p=0}^{K-1} \int_0^1 e^{-j2\pi p\sigma} [\mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H \mathbf{X}_j^H]_{p,n} e^{j2\pi n(\sigma - \frac{k}{K})} \sum_{l \in \mathbb{Z}} e^{j2\pi l(\lambda - K\sigma)} d\lambda. \quad (D16)$$

We may represent the previous expression more compactly by defining the steering vector $\boldsymbol{\theta}(\sigma) \triangleq [1, e^{j2\pi\sigma}, \dots, e^{j2\pi\sigma(K-1)}]^T$, which yields

$$S_j^{(\frac{k}{K})}(\sigma) = \frac{1}{K} \int_0^1 \boldsymbol{\theta}^H(\sigma) \mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(\lambda) \mathbf{U}_T^H \mathbf{X}_j^H \boldsymbol{\theta}(\sigma - \frac{k}{K}) \delta(\lambda - K\sigma) d\lambda + P_Z \delta_k. \quad (D17)$$

We have also converted the sum of exponentials with respect to n' into a Kronecker delta [34, Sec. 10.4.3]. Similarly, the sum with respect to l has been treated as a Dirac delta distribution [35, Sec. 2.11]. Using its sifting property we are left with

$$S_j^{(\frac{k}{K})}(\sigma) = \frac{1}{K} \boldsymbol{\theta}^H(\sigma) \mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(K\sigma) \mathbf{U}_T^H \mathbf{X}_j^H \boldsymbol{\theta}(\sigma - \frac{k}{K}) + P_Z \delta_k. \quad (D18)$$

With the cyclic spectrum, we are able to construct the CSM using (32):

$$[\mathbf{S}_j(\sigma)]_{r,c} = \frac{1}{K} \boldsymbol{\theta}^H(\sigma + \frac{r}{K}) \mathbf{X}_j \mathbf{U}_T \mathbf{S}_H(K\sigma) \mathbf{U}_T^H \mathbf{X}_j^H \boldsymbol{\theta}(\sigma + \frac{c}{K}) + P_Z \delta_{r-c}, \quad (D19)$$

for $r, c \in \{0, \dots, K-1\}$. We have used the fact that $\mathbf{S}_H(\lambda)$ is periodic in λ with period 1, as seen in (21), so that $\mathbf{S}_H(K\sigma + r) \equiv \mathbf{S}_H(K\sigma)$.

Acknowledgements

Not applicable.

Author Contributions

MV-I Original idea, research development, implementation of numerical simulations, preparation of the manuscript and revision. JR Original idea, research development, assessment of numerical simulations, preparation of the manuscript and revision.

Funding

This work was (partially) funded by project MAYTE (PID2022-136512OB-C21) by MICIU/AEI/10.13039/501100011033 and ERDF/EU, grant 2021 SGR 01033 and grant 2022 FI SDUR 00164 by Departament de Recerca i Universitats de la Generalitat de Catalunya.

Data Availability

No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest

The authors declare no Conflict of interest.

Received: 25 April 2025 Accepted: 31 December 2025

Published online: 11 February 2026

References

1. W.A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: half a century of research. *Signal Process.* **86**(4), 639–697 (2006)
2. R.F. Schkoda, R.B. Lund, J.R. Wagner, Clustering of cyclostationary signals with applications to climate station sitings, eliminations, and merges. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **7**(5), 1754–1762 (2014)
3. E. Broszkiewicz-Suwaj, A. Makagon, R. Weron, A. Wyłomańska, On detecting and modeling periodic correlation in financial data. *Physica A* **336**(1–2), 196–205 (2004)

4. G. Hellbourn, R. Weber, C. Capdessus, A.-J. Boonstra, Cyclostationary approaches for spatial RFI mitigation in radio astronomy. *C R Phys.* **13**(1), 71–79 (2011)
5. R. Arora, W.A. Sethares, J.A. Bucklew, Latent periodicities in genome sequences. *IEEE J. Sel. Top. Signal Process.* **2**(3), 332–342 (2008)
6. A. Napolitano, Generalizations of cyclostationarity: a new paradigm for signal processing for mobile communications, radar, and sonar. *IEEE Signal Process. Mag.* **30**(6), 53–63 (2013)
7. W.A. Gardner, *Statistical Spectral Analysis. Prentice Hall Information and System Sciences Series*, 1st edn. (Prentice-Hall Inc, Englewood Cliffs, 1988)
8. A. Napolitano, Cyclostationarity: new trends and applications. *Signal Process.* **120**, 385–408 (2016)
9. A. Napolitano, Cyclostationarity: limits and generalizations. *Signal Process.* **120**, 323–347 (2016)
10. J. Antoni, F. Bonnardot, A. Raad, M. El Badaoui, Cyclostationary modelling of rotating machine vibration signals. *Mech. Syst. Signal Process.* **18**(6), 1285–1314 (2004)
11. W.A. Gardner, *Cyclostationarity in Communications and Signal Processing* (IEEE Press, Piscataway, 1994)
12. M. Shi, Y. Bar-Ness, W. Su, Revisiting the timing and frequency offset estimation based on cyclostationarity with new improved method. *IEEE Commun. Lett.* **13**(7), 537–539 (2009)
13. S. Horstmann, D. Ramirez, P.J. Schreier, Joint detection of almost-cyclostationary signals and estimation of their cycle period. *IEEE Signal Process. Lett.* **25**(11), 1695–1699 (2018)
14. L. Tong, G. Xu, B. Hassibi, T. Kailath, Blind channel identification based on second-order statistics: a frequency-domain approach. *IEEE Trans. Inf. Theory* **41**(1), 329–334 (1995)
15. J. Riba, J. Villares, G. Vázquez, A nondata-aided SNR estimation technique for multilevel modulations exploiting signal cyclostationarity. *IEEE Trans. Signal Process.* **58**(11), 5767–5778 (2010)
16. A. Pries, D. Ramirez, P.J. Schreier, LMPIT-inspired tests for detecting a cyclostationary signal in noise with spatio-temporal structure. *IEEE Trans. Wirel. Commun.* **17**(9), 6321–6334 (2018)
17. M. Chafii, L. Bariah, S. Muhaidat, M. Debbah, Twelve scientific challenges for 6G: rethinking the foundations of communications theory. *IEEE Commun. Surv. Tutor.* **25**(2), 868–904 (2023)
18. L. Jing, E.D. Carvalho, P. Popovski, A.O. Martinez, Design and performance analysis of noncoherent detection systems with massive receiver arrays. *IEEE Trans. Signal Process.* **64**(19), 5000–5010 (2016)
19. T.L. Marzetta, *Spatially-stationary propagating random field model for massive MIMO small-scale fading*, in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, Vail, CO, USA (2018)
20. W. Weichselberger, M. Herdin, H. Ozelik, E. Bonek, A stochastic MIMO channel model with joint correlation of both link ends. *IEEE Trans. Wirel. Commun.* **5**(1), 90–100 (2006)
21. M. Vilà-Insa, J. Riba, *A cyclostationary perspective on noncoherent SIMO communications*, in *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)* (IEEE, Lucca, 2024), pp. 71–75
22. J. Riba, M. Vilà, On infinite past predictability of cyclostationary signals. *IEEE Signal Process. Lett.* **29**, 647–651 (2022)
23. M. Vilà-Insa, J. Riba, Asymptotic analysis of synchronous signal processing (2024) [arXiv:2403.18445](https://arxiv.org/abs/2403.18445) [eess.SP]
24. A. Lapidoth, *A Foundation in Digital Communication* (Cambridge University Press, Cambridge, 2017)
25. P. Stoica, A. Nehorai, Performance study of conditional and unconditional direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.* **38**(10), 1783–1795 (1990)
26. M. Vilà-Insa, J. Riba, Singular detection in noncoherent communications. *IEEE Wirel. Commun. Lett.* **14**(4), 1164–1168 (2025)
27. R. Feng, C.-X. Wang, J. Huang, X. Gao, S. Salous, H. Haas, Classification and comparison of massive MIMO propagation channel models. *IEEE Internet Things J.* **9**(23), 23452–23471 (2022)
28. R.W. Heath Jr., A. Lozano, *Foundations of MIMO Communication*, 1st edn. (Cambridge University Press, Cambridge, 2018)
29. L. Sanguinetti, E. Bjornson, J. Hoydis, Toward massive MIMO 2.0: understanding spatial correlation, interference suppression, and pilot contamination. *IEEE Trans. Commun.* **68**(1), 232–257 (2020)
30. R.M. Gray, Toeplitz and circulant matrices: a review. *Found. Trends Commun. Inform. Theory* **2**(3), 155–239 (2006)
31. P.P. Vaidyanathan, *The Theory of Linear Prediction. Synthesis Lectures on Signal Processing*, vol. 3 (Springer, San Rafael, 2008)
32. J. Zhang, E. Bjornson, M. Matthaiou, D.W.K. Ng, H. Yang, D.J. Love, Prospective multiple antenna technologies for beyond 5G. *IEEE J. Sel. Areas Commun.* **38**(8), 1637–1660 (2020)
33. D. Ramirez, P.J. Schreier, J. Via, I. Santamaria, L.L. Scharf, Detection of multivariate cyclostationarity. *IEEE Trans. Signal Process.* **63**(20), 5395–5408 (2015)
34. P.J. Schreier, L.L. Scharf, *Statistical Signal Processing of Complex-Valued Data* (Cambridge University Press, Cambridge, 2010)
35. R.A. Kennedy, P. Sadeghi, *Hilbert Space Methods in Signal Processing* (Cambridge University Press, Cambridge, 2013)
36. H. Widom, Asymptotic behavior of block Toeplitz matrices and determinants. *Adv. Math.* **13**(3), 284–322 (1974)
37. F. Hlawatsch, *Wireless Communications over Rapidly Time-Varying Channels* (Elsevier Science & Technology, San Diego, 2011)
38. B.C. Levy, *Principles of Signal Detection and Parameter Estimation* (Springer, New York, 2008)
39. T. Yang, M. Barzegar Khalilsarai, S. Haghighatshoar, G. Caire, Structured channel covariance estimation from limited samples for large antenna arrays. *EURASIP J. Wirel. Commun. Netw.* **2023**(1) (2023)
40. T. Kim, M. Bengtsson, E. Larsson, M. Skoglund, Transactions letters—combining long-term and low-rate short-term channel state information over correlated MIMO channels. *IEEE Trans. Wirel. Commun.* **7**(7), 2409–2414 (2008)
41. S.M. Kay, *Modern Spectral Estimation. Prentice-Hall Signal Processing Series* (Prentice Hall, Englewood Cliffs, 1988)
42. M. Vetterli, J. Kovačević, V.K. Goyal, *Foundations of Signal Processing* (Cambridge University Press, Cambridge, 2014)

43. F. Mezzadri, How to generate random matrices from the classical compact groups. *Not. Am. Math. Soc.* **54**(5), 592–604 (2007)
44. A.D.R. Choudary, C.P. Niculescu, *Real Analysis on Intervals. Springer eBook Collection*, 1st edn. (Springer, New Delhi, 2014)
45. A.W. Marshall, I. Olkin, B.C. Arnold, *Inequalities: Theory of Majorization and Its Applications* (Springer, New York, 2011)
46. L. Zheng, D.N.C. Tse, Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inf. Theory* **48**(2), 359–383 (2002)
47. W. Zhang, H.V. Poor, Z. Quan, Frequency-domain correlation: an asymptotically optimum approximation of quadratic likelihood ratio detectors. *IEEE Trans. Signal Process.* **58**(3), 969–979 (2010)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.