Maximum-likelihood Noncoherent PAM Detection

Masters Thesis
By
Georgina Abou Elkheir

Submitted to the Department of Electronic and Computer
Engineering in partial fulfillment of the requirements for the Masters Degree
Technical University of Crete

Advisor: Professor George N. Karystinos

Contents

Abstract

Sequence detection offers significant gains over conventional symbol-by-symbol detection when channel knowledge is not available at the receiver. However, maximum-likelihood (ML) noncoherent sequence detection is often intractable due to exponential complexity in the sequence length. In this work, we develop a polynomial-time ML noncoherent sequence detector for pulse-amplitude modulation (PAM) or quadrature-amplitude modulation (QAM) transmissions in Rayleigh fading. Our detector follows an auxiliary-angle approach that reduces the exponential-size space of solution vectors to a polynomial-size set of candidate sequences; we prove that this significantly smaller set always contains the ML PAM/QAM sequence. Hence, the proposed detector solves the problem of identifying the ML PAM/QAM sequence in unknown Rayleigh fading with overall polynomial complexity.

I. Introduction

Noncoherent signal detection over unknown fading channels has received significant attention, in particular for the case of the block-fading channel model [1], [2]. Since it does not need any channel knowledge, noncoherent detection is applicable even in most degraded and fast-fading channels. In [3], it was shown that conventional transmission schemes, such as phase-shift keying (PSK) or quadrature amplitude modulation (QAM), suffice to approach capacity in the single-input single-output (SISO) block Rayleigh-fading channel.

The major difference between coherent and noncoherent detection is that, in the latter, the lack of channel knowledge at the receiver induces memory in the received data sequence. Hence, to minimize the sequence error probability under this scenario, we need to perform maximum-likelihood (ML) sequence detection on the *entire* coherence interval. ML sequence detection offers significant performance gains [4]-[6] over conventional symbol-by-symbol detection but, in general, it suffers from exponential complexity with respect to the sequence length.

Interestingly, ML sequence detection does not always require exponential complexity. For example, if the data symbols take values from a PSK constellation, then ML noncoherent sequence detection under Rayleigh fading can be performed with complexity $\mathcal{O}(N \log N)$, where N is the sequence length, using the auxiliary-angle approach that was first presented in [7] and later used in [8]-[10]. However, the approach in [7] is restricted to constant-amplitude constellations. Pulse-amplitude modulation (PAM) and QAM are nonconstant-amplitude constellations, making noncoherent sequence detection more challenging.

Several attempts have been made in the direction of polynomial-complexity noncoherent PAM/QAM sequence detection. Generalized-likelihood-ratio-test (GLRT) detection algorithms of complexity $\mathcal{O}(N \log N)$ were developed in [11] for QAM transmissions over a phase-noncoherent, amplitude-coherent channel and in [12] for PAM transmissions over a phase-coherent, amplitude-noncoherent channel. The delelopments in [12] used lattice decoding techniques and also included an algorithm for GLRT PAM/QAM sequence detection over a fully noncoherent channel with complexity $\mathcal{O}\left(N^3\right)$ which was further reduced to $\mathcal{O}\left(N^2 \log N\right)$ in [13]. However, the question of whether ML noncoherent detection of a PAM/QAM sequence can also be performed in polynomial time or not has remained open.

In this work, we present the first polynomial-complexity ML noncoherent detector of PAM/QAM symbol sequences under Rayleigh fading. The proposed approach decomposes the detection problem into a polynomial-size set of simple problems. That way we effectively reduce the exponential size of the solution space to a polynomial-size subset that always contains the optimal PAM/QAM sequence.

The rest of this work is organized as follows. The problem formulation is presented in Section II where, to keep the presentation simple, we focus on PAM transmissions. The proposed method for the ML noncoherent detection of a PAM sequence is described in Subsections III.A and III.B and it is shown that its worst-case complexity is polynomial in the sequence length. In Subsection III.C, we show how the proposed technique is extended to QAM without increasing its worst-case complexity order. The performance of the proposed algorithm is tested through simulations in Section IV. A few concluding remarks are drawn in Section V.

II. Signal Model and Problem Statement

We consider a SISO system and assume transmission of a sequence of N uncoded M-ary PAM (M-PAM) data symbols $\mathbf{s} = [s_1, s_2, \ldots, s_N]^T$ where s_n, $n = 1, 2, \ldots, N$, is selected from an M-ary alphabet

$$\mathcal{A}_M \stackrel{\triangle}{=} \left\{\pm a_1, \pm a_2, \ldots, \pm a_{\frac{M}{2}}\right\} \tag{1}$$

where $0 < a_1 < a_2 < \ldots < a_{\frac{M}{2}}$.[1] The data sequence is shaped, upconverted, and transmitted over a Rayleigh fading channel. The downconverted and pulse-matched filtered received vector is

$$\mathbf{y} = h\mathbf{s} + \mathbf{n} \tag{2}$$

where $h \sim \mathcal{CN}\left(0, \sigma_h^2\right)$ represents Rayleigh flat fading channel processing and $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_n^2\mathbf{I}\right)$ denotes zero-mean additive white complex Gaussian noise.

In this work, we assume that the channel coefficient h is not available to the receiver. Hence, coherent detection cannot be utilized and the optimal receiver takes the form of a sequence detector that operates on the entire received vector $\mathbf{y}$. Given the observation vector $\mathbf{y}$, the ML detector for the M-PAM sequence $\mathbf{s}$ maximizes the conditional probability density function (pdf) of $\mathbf{y}$ given $\mathbf{s}$. Thus, the optimal decision is given by

$$\hat{\mathbf{s}}^{\mathrm{ML}} = \arg\max_{\mathbf{s} \in \mathcal{A}_M^N} f\left(\mathbf{y}|\mathbf{s}\right) \tag{3}$$

where $f(\cdot|\cdot)$ represents the pertinent vector pdf of the channel output conditioned on the transmitted symbol sequence. The conditional received vector $\mathbf{y}$ given the transmitted sequence $\mathbf{s}$ is a complex Gaussian vector with mean $E\{\mathbf{y}|\mathbf{s}\} = \mathbf{0}$ and covariance matrix

$$\mathbf{C}_{\mathbf{y}|\mathbf{s}} \stackrel{\triangle}{=} E\left\{\mathbf{y}\mathbf{y}^H|\mathbf{s}\right\} = \sigma_h^2\mathbf{s}\mathbf{s}^T + \sigma_n^2\mathbf{I}_N. \tag{4}$$

[1]In this work, we need not specify $a_1, a_2, \ldots, a_{\frac{M}{2}}$ but only let them take arbitrary different positive values. In conventional M-PAM, the positive elements of $\mathcal{A}_M$ are given by $a_i = 2i - 1$, $i = 1, 2, \ldots, \frac{M}{2}$.

Therefore, the optimization problem in (1) is rewritten as

$$\hat{\mathbf{s}}^{\mathrm{ML}} = \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \frac{1}{\pi^N \left|\mathbf{C}_{\mathbf{y}|\mathbf{s}}\right|} \exp\left(\mathbf{y}^H \mathbf{C}_{\mathbf{y}|\mathbf{s}}^{-1}\mathbf{y}\right). \tag{5}$$

Using identities for the determinant and inverse of a rank-1 update [14], we compute

$$\left|\mathbf{C}_{\mathbf{y}|\mathbf{s}}\right| = \left|\sigma_n^2\mathbf{I}_N + \sigma_h^2\mathbf{s}\mathbf{s}^T\right| = \left|\sigma_n^2\mathbf{I}_N\right| \left(1 + \frac{\sigma_h^2}{\sigma_n^2}\|\mathbf{s}\|^2\right) = \sigma_n^{2N-2}\left(\sigma_n^2 + \sigma_h^2\|\mathbf{s}\|^2\right) \tag{6}$$

and

$$\mathbf{C}_{\mathbf{y}|\mathbf{s}}^{-1} = \left(\sigma_n^2\mathbf{I}_N + \sigma_h^2\mathbf{s}\mathbf{s}^T\right)^{-1} = \frac{1}{\sigma_n^2}\mathbf{I}_N - \frac{\sigma_h^2}{\sigma_n^4 + \sigma_h^2\sigma_n^2\|\mathbf{s}\|^2}\mathbf{s}\mathbf{s}^T. \tag{7}$$

Then, we substitute (6) and (7) in (5) and obtain

$$\begin{aligned}
\hat{\mathbf{s}}^{\mathrm{ML}} &= \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left(-\ln\left(\sigma_n^{2N-2}\left(\sigma_n^2 + \sigma_h^2\|\mathbf{s}\|^2\right)\right) - \mathbf{y}^H\left(\frac{1}{\sigma_n^2}\mathbf{I}_N - \frac{\sigma_h^2}{\sigma_n^4 + \sigma_h^2\sigma_n^2\|\mathbf{s}\|^2}\mathbf{s}\mathbf{s}^T\right)\mathbf{y}\right) \\
&= \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left(\frac{\sigma_h^2}{\sigma_n^4 + \sigma_h^2\sigma_n^2\|\mathbf{s}\|^2}\mathbf{s}^T\mathbf{y}\mathbf{y}^H\mathbf{s} - \ln\left(\sigma_n^2 + \sigma_h^2\|\mathbf{s}\|^2\right)\right) \\
&= \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left(\frac{\sigma_h^2}{\sigma_n^2}\frac{\mathcal{R}\left(\mathbf{s}^T\mathbf{y}\mathbf{y}^H\mathbf{s}\right)}{g(\|\mathbf{s}\|)} - \ln(g(\|\mathbf{s}\|))\right) = \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left\{\frac{\sigma_h^2}{\sigma_n^2}\frac{\mathbf{s}^T\mathcal{R}\left\{\mathbf{y}\mathbf{y}^H\right\}\mathbf{s}}{g(\|\mathbf{s}\|)} - \ln(g(\|\mathbf{s}\|))\right\} \\
&= \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left\{\frac{\sigma_h^2}{\sigma_n^2}\frac{\mathbf{s}^T\mathbf{V}\mathbf{V}^T\mathbf{s}}{g(\|\mathbf{s}\|)} - \ln(g(\|\mathbf{s}\|))\right\} = \arg\max_{\mathbf{s}\in\mathcal{A}_M^N} \left\{\frac{\sigma_h^2}{\sigma_n^2}\frac{\left\|\mathbf{V}^T\mathbf{s}\right\|^2}{g(\|\mathbf{s}\|)} - \ln(g(\|\mathbf{s}\|))\right\}
\end{aligned} \tag{8}$$

where $g(\|\mathbf{s}\|) \triangleq \sigma_n^2 + \sigma_h^2\|\mathbf{s}\|^2$ and $\mathbf{V} = [\mathcal{R}\{\mathbf{y}\}\ \ \mathcal{I}\{\mathbf{y}\}] \in \mathbb{R}^{N\times 2}$.

A straightforward but naive approach to (8) would be an exhaustive search among all M^N data sequences $\mathbf{s} \in \mathcal{A}_M^N$. However, such a receiver would be impractical even for moderate values of N, since its complexity grows exponentially with N. In the next section, we present an efficient algorithm that solves the above maximization in polynomial time.

III. Polynomial-complexity ML PAM/QAM Sequence Detection

A. Theoretic developments

We begin our developments by defining the set of positive amplitudes of the elements of $\mathcal{A}_M$,

$$\mathcal{A}_M^+ \triangleq \left\{a_1, a_2, \ldots, a_{\frac{M}{2}}\right\}, \tag{9}$$

and rewriting $\mathbf{s} \in \mathcal{A}_M^N$ as $\mathbf{s} = \mathbf{b} \odot \mathbf{x}$ where $\mathbf{b} \in \left(\mathcal{A}_M^+\right)^N$ is the amplitude component of $\mathbf{s}$ and has the same norm as $\mathbf{s}$, i.e. $\|\mathbf{b}\| = \|\mathbf{s}\|$, $\mathbf{x} \in \{\pm 1\}^N$ is the sign component of $\mathbf{s}$, and $\odot$ accounts for Hadamard (that is, entry-wise) product. Then, the maximization in the ML decision in (8) becomes

$$\max_{\mathbf{s}\in\mathcal{A}_M^N}\left\{\frac{\sigma_h^2}{\sigma_n^2}\frac{\left\|\mathbf{V}^T\mathbf{s}\right\|^2}{g(\|\mathbf{s}\|)}-\ln g(\|\mathbf{s}\|)\right\}=\max_{\mathbf{b}\in\left(\mathcal{A}_M^+\right)^N}\left\{\frac{\sigma_h^2}{\sigma_n^2 g(\|\mathbf{b}\|)}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|^2-\ln g\left(\|\mathbf{b}\|\right)\right\}. \quad (10)$$

Next, $\forall\ \left(t_1,t_2,\ldots,t_{\frac{M}{2}}\right)\in\mathbb{N}^{\frac{M}{2}}$ s.t. $\sum_{i=1}^{\frac{M}{2}}t_i=N$, we define the set

$$\begin{aligned}\mathcal{B}_{\left(t_1,t_2,\ldots,t_{\frac{M}{2}}\right)}\triangleq\Bigg\{&\mathbf{b}=[b_1,b_2,\ldots,b_N]^T\in\left(\mathcal{A}_M^+\right)^N\\&\text{s.t. } \left|\{b_n,\ n=1,2,\ldots,N,\ \text{s.t. } b_n=a_i\}\right|=t_i,\ \forall\, i=1,2,\ldots,\frac{M}{2}\Bigg\}.\end{aligned} \quad (11)$$

Apparently, the ith element of $\left(t_1,t_2,\ldots,t_{\frac{M}{2}}\right)$, denoted by t_i, equals the number of times the amplitude element a_i appears in vector $\mathbf{b}$. Hence, all vectors in set $\mathcal{B}_{\left(t_1,t_2,\ldots,t_{M/2}\right)}$ contain the same number of values $a_1,a_2,\ldots,a_{\frac{M}{2}}$ with each other. Consequently, all vectors in $\mathcal{B}_{\left(t_1,t_2,\ldots,t_{M/2}\right)}$ have identical norm which equals $\sum_{i=1}^{\frac{M}{2}}t_ia_i^2$. We say that set $\mathcal{B}_{\left(t_1,t_2,\ldots,t_{M/2}\right)}$ *is associated with* this norm. The number of all possible sets $\mathcal{B}_{\left(t_1,t_2,\ldots,t_{M/2}\right)}$ that we can construct equals $T=\binom{N+\frac{M}{2}-1}{\frac{M}{2}-1}$ which is the number of possible ways one can choose a set of size N from $\frac{M}{2}$ elements when repetitions are possible. Interestingly, T is polynomial in N.

Let sets $\mathcal{B}_1,\mathcal{B}_2\ldots,\mathcal{B}_T$ be all possible sets of the form $\mathcal{B}_{\left(t_1,t_2,\ldots,t_{M/2}\right)}$ that we can construct. Also, let $\beta_1,\beta_2,\ldots,\beta_T$ denote their associated norms, i.e. $\beta_k=\|\mathbf{b}\|,\ \forall\ \mathbf{b}\in\mathcal{B}_k,\ k=1,2,\ldots,T$. Since, by construction, sets $\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_{\mathcal{T}}$ are disjoint and their union equals $\left(\mathcal{A}_M^+\right)^N$, they constitute a partition of $\left(\mathcal{A}_M^+\right)^N$, i.e. $\bigcup_{k=1}^{T}\mathcal{B}_k=\left(\mathcal{A}_M^+\right)^N$, and the maximization in (10) is rewritten as

$$\begin{aligned}&\max_{k=1,2,\ldots,T}\max_{\mathbf{b}\in\mathcal{B}_k}\left\{\frac{\sigma_h^2}{\sigma_n^2 g(\|\mathbf{b}\|)}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|^2-\ln g(\|\mathbf{b}\|)\right\}\\&=\max_{k=1,2,\ldots,T}\left\{\frac{\sigma_h^2}{\sigma_n^2 g(\beta_k)}\max_{\mathbf{b}\in\mathcal{B}_k}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T(\mathbf{b}\odot\mathbf{x})\right\|^2-\ln g(\beta_k)\right\}.\end{aligned} \quad (12)$$

To find the ML symbol sequence $\hat{\mathbf{s}}^{\text{ML}}$, we need to compute the optimal amplitude and sign vectors $\mathbf{b}$, $\mathbf{x}$, respectively, for each set $\mathcal{B}_k,\ k=1,2,\ldots,T$. The key idea is to find a set of candidate pairs $(\mathbf{b},\mathbf{x})\in\mathcal{B}_k\times\{\pm1\}^N$, for each $k=1,2,\ldots,T$, by focusing on the inner maximization

$$\max_{\mathbf{b}\in\mathcal{B}_k}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|^2=\left(\max_{\mathbf{b}\in\mathcal{B}_k}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|\right)^2. \quad (13)$$

Hence, for a fixed $k=1,2,\ldots,T$, our problem simplifies to the identification of the pair $(\mathbf{b},\mathbf{x})\in\mathcal{B}_k\times\{\pm1\}^N$ that maximizes $\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|$, that is

$$\max_{\mathbf{b}\in\mathcal{B}_k}\max_{\mathbf{x}\in\{\pm1\}^N}\left\|\mathbf{V}^T\left(\mathbf{b}\odot\mathbf{x}\right)\right\|. \quad (14)$$

We introduce the angle $\phi \in (-\pi, \pi]$ and define the 2×1 polar vector

$$\mathbf{c}(\phi) \triangleq \begin{bmatrix} \sin(\phi) \\ \cos(\phi) \end{bmatrix}. \tag{15}$$

From Cauchy-Schwarz inequality, we know that for any vector $\mathbf{a} \in \mathbb{R}^D$,

$$\mathbf{a}^T \mathbf{c}(\phi) \leq \|\mathbf{a}\| \underbrace{\|\mathbf{c}(\phi)\|}_{=1} = \|\mathbf{a}\| \tag{16}$$

with equality if and only if ϕ is the angle of $\mathbf{a}$. Hence, $\max\limits_{\phi \in (-\pi,\pi]} \left\{\mathbf{a}^T \mathbf{c}(\phi)\right\} = \|\mathbf{a}\|$, since we can always assign a value to $\phi \in (-\pi, \pi]$ such that $\mathbf{c}(\phi)$ is codirectional with $\mathbf{a}$. Substituting $\mathbf{a}$ with $\mathbf{V}^T(\mathbf{b} \odot \mathbf{x})$, we obtain $\max\limits_{\phi \in (-\pi,\pi]} \left\{\left(\mathbf{b}^T \odot \mathbf{x}^T\right) \mathbf{V}\mathbf{c}(\phi)\right\} = \left\|\mathbf{V}^T(\mathbf{b} \odot \mathbf{x})\right\|$. Then, (14) becomes

$$\max_{\mathbf{b} \in \mathcal{B}_k} \max_{\mathbf{x} \in \{\pm 1\}^N} \left\|\mathbf{V}^T(\mathbf{b} \odot \mathbf{x})\right\| = \max_{\mathbf{b} \in \mathcal{B}_k} \max_{\mathbf{x} \in \{\pm 1\}^N} \max_{\phi \in (-\pi,\pi]} \left\{\left(\mathbf{x}^T \odot \mathbf{b}^T\right) \mathbf{V}\mathbf{c}(\phi)\right\}. \tag{17}$$

If we set

$$\mathbf{u}(\phi) \triangleq \mathbf{V}\mathbf{c}(\phi) = \begin{bmatrix} |V_{1,1}\sin(\phi) + V_{1,2}\cos(\phi)| \\ \vdots \\ |V_{N,1}\sin(\phi) + V_{N,2}\cos(\phi)| \end{bmatrix}, \tag{18}$$

then (17) becomes

$$\begin{aligned} \max_{\mathbf{b} \in \mathcal{B}_k} \max_{\mathbf{x} \in \{\pm 1\}^N} \left\|\mathbf{V}^T(\mathbf{b} \odot \mathbf{x})\right\| &= \max_{\phi \in (-\pi,\pi]} \max_{\mathbf{b} \in \mathcal{B}_k} \max_{\mathbf{x} \in \{\pm 1\}^N} \left\{\mathbf{x}^T \mathrm{diag}(\mathbf{b})\mathbf{u}(\phi)\right\} \\ &= \max_{\phi \in (-\pi,\pi]} \max_{\mathbf{b} \in \mathcal{B}_k} \max_{\mathbf{x} \in \{\pm 1\}^N} \left\{\mathbf{x}^T(\mathbf{b} \odot \mathbf{u}(\phi))\right\}. \end{aligned} \tag{19}$$

For a given point $\phi \in (-\pi, \pi]$ and a given vector $\mathbf{b} \in \mathcal{B}_k$, $\mathbf{b} \odot \mathbf{u}(\phi)$ is a fixed vector and the inner maximization $\max\limits_{\mathbf{x} \in \{\pm 1\}^N} \left\{\mathbf{x}^T(\mathbf{b} \odot \mathbf{u}(\phi))\right\}$ in (19) is achieved for

$$\mathbf{x}(\phi) \triangleq \mathrm{sgn}(\mathbf{u}(\phi)), \tag{20}$$

resulting in the value

$$\max_{\mathbf{x} \in \{\pm 1\}^N} \left\{\mathbf{x}^T(\mathbf{b} \odot \mathbf{u}(\phi))\right\} = \mathbf{b}^T|\mathbf{u}(\phi)|, \tag{21}$$

and

$$\mathbf{b}_k(\phi) \triangleq \arg\max_{\mathbf{b} \in \mathcal{B}_k} \mathbf{b}^T|\mathbf{u}(\phi)|. \tag{22}$$

Apparently, for any $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$, since $\left(t_1, t_2, \ldots, t_{\frac{M}{2}}\right)$ is common for all vectors in $\mathcal{B}_k$, the maximization problem in (22) is shifted to determining the t_1, t_2, $\ldots$, and $t_{\frac{M}{2}}$ indices of $\mathbf{b}$ that correspond to

a_1, a_2, $\ldots$, and $a_{\frac{M}{2}}$, respectively. Indeed, for any $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$, the optimal amplitude vector $\mathbf{b}_k(\phi)$ in (22) is achieved by sorting the elements of $|\mathbf{u}(\phi)|$ and assigning the t_1 smallest values to a_1, the next t_2 smallest values to a_2, etc. Since $\mathbf{c}(\phi+\pi) = -\mathbf{c}(\phi)$, we obtain $\mathbf{u}(\phi+\pi) = -\mathbf{u}(\phi) \Rightarrow |\mathbf{u}(\phi+\pi)| = |\mathbf{u}(\phi)|$. Hence, the ordering of the elements of $|\mathbf{u}(\phi + \pi)|$ is identical to the one of $|\mathbf{u}(\phi)|$ which implies that $\phi + \pi$ and ϕ are associated with equal amplitude vectors $\mathbf{b}_k(\phi + \pi) = \mathbf{b}_k(\phi)$ and opposite sign vectors $\mathbf{x}(\phi + \pi) = -\mathbf{x}(\phi)$. Since opposite sign vectors result in the same metric value in the maximization in (14), we can ignore the values of $\phi \in \left(-\pi, -\frac{\pi}{2}\right] \cup \left(\frac{\pi}{2}, \pi\right]$ and restrict ϕ to the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

To gain some intuition into the purpose of inserting the auxiliary variable ϕ, we notice that every element of $\mathbf{u}(\phi)$ is actually a continuous function (curve) of ϕ. When, due to (20), we examine the sign of $\mathbf{u}(\phi)$ at a given point ϕ, we actually examine the signs of the curves $u_1(\phi)$, $u_2(\phi)$, $\ldots$, $u_N(\phi)$ according to their values at ϕ. Similarly, when, due to (22), we sort the elements of $|\mathbf{u}(\phi)|$ at a given point ϕ, we actually sort the curves according to their magnitudes at point ϕ. The optimal pair $(\mathbf{b}, \mathbf{x}) \in \mathcal{B}_k \times \{\pm 1\}^N$ in (19) is met if we scan the entire interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and collect the "locally optimal" pair $(\mathbf{b}(\phi), \mathbf{x}(\phi))$ for any point $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The key observation in this work is that, due to the continuity of the curves in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$, we expect that in an area "around" ϕ the curves will retain their signs and magnitude sorting. Hence, we expect the formation of intervals in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ within which both the signs and the magnitude sorting of the curves will remain unaltered, irrespectively of the fact that the value of each curve changes. Our objective, in the sequel, is to determine all these intervals, show that their number is less than or equal to N^2, and present an algorithm that identifies the pairs $(\mathbf{b}, \mathbf{x}) \in \mathcal{B}_k \times \{\pm 1\}^N$ that are associated with these intervals and runs in polynomial time. Once we have collected all candidate pairs $(\mathbf{b}, \mathbf{x}) \in \mathcal{B}_k \times \{\pm 1\}^N$ for any $k = 1, 2, \ldots, T$, the overall optimal vector $\hat{\mathbf{s}}^{\mathrm{ML}}$ will be determined through exhaustive search among all candidate pairs $(\mathbf{b}, \mathbf{x})$.

To identify all intervals in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ that retain their sign and amplitude vectors, it suffices to examine when *(i)* one of the curves that correspond to the elements of $\mathbf{u}(\phi)$ changes its sign, which happens when it becomes zero, or *(ii)* two of the curves that correspond to the elements of $|\mathbf{u}(\phi)|$ change their magnitude sorting, which happens when the two curves intersect. Interestingly, the second case is a necessary but not sufficient condition for a change in $\mathbf{b}(\phi)$, since the intersecting curves may correspond to elements of $|\mathbf{u}(\phi)|$ that are assigned to the same amplitude value a_i, $i = 1, 2, \ldots, \frac{M}{2}$. In the latter case, although the magnitude sorting of the curves changes, the amplitude vector $\mathbf{b}(\phi)$ does not.

When a curve, say the one that corresponds to element i of $\mathbf{u}(\phi)$, becomes zero, we have

$$u_i(\phi) = 0 \Leftrightarrow \mathbf{V}_{i,:}\mathbf{c}(\phi) = 0 \Leftrightarrow V_{i,1}\sin\phi + V_{i,2}\cos\phi = 0 \\ \Leftrightarrow \phi = \tan^{-1}\left(-\frac{V_{i,2}}{V_{i,1}}\right). \quad (23)$$

At the intersection of two curves, say those that correspond to elements j and k of $|\mathbf{u}(\phi)|$, we have

$$|u_j(\phi)| = |u_k(\phi)| \Leftrightarrow |\mathbf{V}_{j,:}\mathbf{c}(\phi)| = |\mathbf{V}_{k,:}\mathbf{c}(\phi)| \Leftrightarrow V_{j,1}\sin(\phi) + V_{j,2}\cos(\phi) = \pm\left(V_{k,1}\sin(\phi) + V_{k,2}\cos(\phi)\right) \\ \Leftrightarrow \phi = \tan^{-1}\left(\frac{\pm V_{k,2} - V_{j,2}}{V_{j,1} \mp V_{k,1}}\right). \quad (24)$$

We observe that the solution $\phi = \tan^{-1}\left(-\frac{V_{i,2}}{V_{i,1}}\right)$ in (23) determines a point that partitions $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ into two intervals; in one of these intervals $u_i(\phi)$ is positive, while in the other interval it is negative. Since there are N curves $u_1(\phi)$, $u_2(\phi)$, $\dots$, $u_N(\phi)$, we obtain N corresponding sign-change points through (23). We also observe that the two solutions $\tan^{-1}\left(\frac{\pm V_{k,2} - V_{j,2}}{V_{j,1} \mp V_{k,1}}\right)$ in (24) determine two points that partition $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ into three intervals; in successive intervals the inequality $|u_j(\phi)| > |u_k(\phi)|$ becomes $|u_j(\phi)| < |u_k(\phi)|$ or vice versa. The two points originate from a pair of elements of $|\mathbf{u}(\phi)|$. Since the N elements of $|\mathbf{u}(\phi)|$ can be combined in $\binom{N}{2} = \frac{N^2-2}{2}$ pairs and each pair yields two points, the number of points that correspond to intersections of curves in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ equals $2\binom{N}{2} = N^2 - N$. Hence, the total number of points is N^2 and $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is partitioned into N^2 disjoint intervals so that each such interval corresponds to a single amplitude vector $\mathbf{b} \in \mathcal{B}_k$ and a single sign vector $\mathbf{x} \in \{\pm 1\}^N$. Of course, more than one intervals may correspond to the same amplitude vector, as explained before. In the following subsection, we efficiently identify these candidate pairs $(\mathbf{b}, \mathbf{x})$, since one of them is the optimal pair for set $\mathcal{B}_k$.

B. Algorithmic developments

We begin by computing and sorting in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ all N^2 points where curves change their signs or intersect. Then, we consider $\phi = -\frac{\pi}{2}$, compute $\mathbf{u}\left(-\frac{\pi}{2}\right)$, find $\mathbf{x}\left(-\frac{\pi}{2}\right) = \text{sgn}\left(\mathbf{u}\left(-\frac{\pi}{2}\right)\right)$, sort the elements of $\left|\mathbf{u}\left(-\frac{\pi}{2}\right)\right|$, and, for each $k = 1, 2, \dots, T$, assign amplitudes to its elements -as described before- to create the corresponding amplitude vector $\mathbf{b}_k\left(-\frac{\pi}{2}\right)$. For any $k = 1, 2, \dots, T$, the pair $\left(\mathbf{b}_k\left(-\frac{\pi}{2}\right), \mathbf{x}\left(-\frac{\pi}{2}\right)\right)$ is optimal for the interval that starts at $-\frac{\pi}{2}$ and ends at the first sign-change or intersection point from the sorted list.

We continue by considering the first point from the sorted list. If it corresponds to a sign change, then we update $\mathbf{x}$ by switching the sign of its element that corresponds to the curve that changes its sign at the point we consider. Otherwise, it is an intersection point and we consider the two elements of $|\mathbf{u}(\phi)|$ that correspond to the two curves that intersect at this point. If these two elements are assigned to different amplitudes in $\mathcal{A}_M^+$, then we switch these two amplitudes and define a new vector $\mathbf{b}$ accordingly. Otherwise, we do nothing.

Next, we move to the next point from the sorted list and repeat the above procedure, i.e. we update the amplitude vector $\mathbf{b}$ or the sign vector $\mathbf{x}$ if the point corresponds to an intersection or a sign change, respectively. We continue similarly until we visit all N^2 points. In the end, for each $k = 1, 2, \ldots, T$, we have collected a set of at most N^2 candidate pairs $(\mathbf{b}, \mathbf{x})$. Finally, we consider (10) which is our original ML optimization problem. The optimal pair $(\mathbf{b}, \mathbf{x})$ is obtained by comparing all collected pairs $(\mathbf{b}, \mathbf{x})$, for all $k = 1, 2, \ldots, T$, against the ML metric in (10). The ML M-PAM vector $\hat{\mathbf{s}}^{\mathrm{ML}}$ is obtained from the optimal pair $(\mathbf{b}, \mathbf{x})$ through $\mathbf{s} = \mathbf{b} \odot \mathbf{x}$.

As a brief summary, the sequence of calculations of the proposed ML noncoherent receiver is as follows. The entire received vector $\mathbf{y}$ is utilized to form matrix $\mathbf{V} = [\mathcal{R}\{\mathbf{y}\}\ \mathcal{I}\{\mathbf{y}\}]$. Then, the N^2 sign-change or intersection points are calculated from (23) or (24), respectively, and sorted in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Subsequently, for any $k = 1, 2, \ldots, T$, we start at $\phi = -\frac{\pi}{2}$ and visit the N^2 points to create the corresponding candidate pairs $(\mathbf{b}, \mathbf{x}) \in \mathcal{B}_k \times \{\pm 1\}^N$ which are less than or equal to N^2. All collected pairs $(\mathbf{b}, \mathbf{x}) \in \bigcup_{k=1}^{T} \mathcal{B}_k \times \{\pm 1\}^N$ are computed against the ML metric in (10). Finally, the optimal symbol sequence $\hat{\mathbf{s}}^{\mathrm{ML}}$ is obtained from the optimal pair $(\mathbf{b}, \mathbf{x})$ through $\mathbf{s} = \mathbf{b} \odot \mathbf{x}$.

The complexity of the proposed ML noncoherent receiver is dominated by the computational cost needed to generate all candidate pairs $(\mathbf{b}, \mathbf{x})$. For any $k = 1, 2, \ldots, T$, we create at most N^2 pairs $(\mathbf{b}, \mathbf{x})$, each with constant complexity (except from the first pair that requires $\mathcal{O}(N)$ calculations, given the sorting of the elements of $\left|\mathbf{u}\left(-\frac{\pi}{2}\right)\right|$ which costs $\mathcal{O}(N \log N)$ but is valid for all $k = 1, 2, \ldots, T$). Hence, for any $k = 1, 2, \ldots, T$, the worst-case complexity for the generation of all pairs $(\mathbf{b}, \mathbf{x}) \in \mathcal{B}_k \times \{\pm 1\}^N$ is $\mathcal{O}(N^2)$. Since $T = \binom{N+\frac{M}{2}-1}{\frac{M}{2}-1} = \mathcal{O}\left(N^{\frac{M}{2}-1}\right)$, the worst-case overall complexity of the proposed receiver is $\mathcal{O}\left(N^{\frac{M}{2}+1}\right)$.

C. Extension to QAM

For simplicity of the presentation, we concentrated our analysis and developments of polynomial-complexity ML noncoherent sequence detection in Rayleigh fading to PAM transmissions, as described in (1) and (2). In the case of QAM, with a slight modification that is presented below, our proposed algorithm still solves, with polynomial complexity, the ML noncoherent sequence detection problem.

We recall that, for M-PAM, the ML sequence detection problem is given by (8). The two important quantities that appear in (8) are $\mathbf{y}^H\mathbf{s}$ and $\|\mathbf{s}\|$. In the case of M-QAM, the transmitted signal vector takes the form

$$\mathbf{s} = \mathbf{s}_\mathcal{R} + j\mathbf{s}_\mathcal{I} = [\mathbf{I} \;\; j\mathbf{I}] \begin{bmatrix} \mathbf{s}_\mathcal{R} \\ \mathbf{s}_\mathcal{I} \end{bmatrix} \tag{25}$$

where the $\sqrt{M}$-PAM vectors $\mathbf{s}_\mathcal{R} \in \mathcal{A}_{\sqrt{M}}$ and $\mathbf{s}_\mathcal{I} \in \mathcal{A}_{\sqrt{M}}$ are the real and imaginary, respectively, parts of the M-QAM vector $\mathbf{s}$. We define the $\sqrt{M}$-PAM vector $\tilde{\mathbf{s}} \triangleq \left[\mathbf{s}_\mathcal{R}^T \;\; \mathbf{s}_\mathcal{I}^T\right]^T$ of length $2N$. Then, (25) becomes

$$\mathbf{s} = [\mathbf{I} \;\; j\mathbf{I}]\, \tilde{\mathbf{s}} \tag{26}$$

and the two quantities that affect the optimization problem in (8) become

$$\|\mathbf{s}\| = \|\tilde{\mathbf{s}}\| \tag{27}$$

and

$$\mathbf{y}^H\mathbf{s} = \mathbf{y}^H [\mathbf{I} \;\; j\mathbf{I}]\, \tilde{\mathbf{s}} = \left[\mathbf{y}^H \;\; j\mathbf{y}^H\right] \tilde{\mathbf{s}} = \tilde{\mathbf{y}}^H \tilde{\mathbf{s}} \tag{28}$$

where $\tilde{\mathbf{y}} \triangleq \left[\mathbf{y}^T \;\; -j\mathbf{y}^T\right]^T \in \mathbb{C}^{2N}$. Due to (8), (27), and (28), it turns out that the ML noncoherent detection of an M-QAM sequence of length N simplifies to the ML noncoherent detection of an equivalent $\sqrt{M}$-PAM sequence of length $2N$. Hence, the optimal M-QAM sequence of length N is obtained by the proposed algorithm with worst-case complexity $\mathcal{O}\left(N^{\frac{\sqrt{M}}{2}+1}\right)$.

IV. Simulation Results

As an illustration, we consider PAM transmissions in an unknown Rayleigh fading channel environment and present the symbol error rate (SER) of the GLRT and ML sequence detectors.

In Fig. 1, we present the SER as a function of the transmitted signal-to-noise ratio (SNR), for 4-PAM and sequence length $N = 3$, 7, and 9. We note that GLRT detection is implemented by the algorithm developed in [12] while ML detection is implemented by the proposed algorithm. Interestingly, both algorithms have the same complexity $\mathcal{O}\left(N^3\right)$. We observe the performance gains of ML detection over GLRT detection, for all three values of sequence length N.

In Fig. 2, we present the SER as a function of the SNR, for 8-PAM and sequence length $N = 3$ and 7. Here, GLRT detection has the same complexity as in 4-PAM while ML detection requires $\mathcal{O}\left(N^5\right)$ computational cost. The performance gains of MLSD over GLRT detection are maintained.

In Fig. 3, we present a similar study for 16-QAM and sequence length $N = 3$ and 7. In this case, both algorithms have the same complexity $\mathcal{O}\left(N^3\right)$. Finally, in Fig. 4 we present the SER as a function of the sequence length, for 4-PAM and SNR = 60 dB. The performance gains of MLSD over GLRT detection are maintained for all values of the sequence length.

V. Conclusions

We developed a ML noncoherent sequence detector for PAM or QAM transmissions in Rayleigh fading that operates in polynomial time. Our detector follows an auxiliary-angle approach that reduces the exponential-size space of solution vectors to a polynomial-size set of candidate sequences; we proved that this significantly smaller set always contains the ML PAM/QAM sequence. Hence, the proposed detector solves the problem of identifying the ML PAM/QAM sequence in unknown Rayleigh fading with overall polynomial complexity.

Finally, we illustrated the performance superiority of the proposed ML detector to the polynomial-complexity GLRT detector [12], [13] for several PAM/QAM constellation sizes and several sequence lengths.

References

[1] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, pp. 139-157, Jan. 1999.

[2] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, pp. 359-383, Feb. 2002.

[3] R.-R. Chen, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, pp. 1676-1689, Oct. 2003.

[4] D. Makrakis and P. T. Mathiopoulos, "Optimal decoding in fading channels: A combined envelope, multiple differential and coherent detection approach," in *Proc. IEEE GLOBECOM 1989*, Dallas, TX, Nov. 1989, vol. 3, pp. 1551-1557.

[5] S. G. Wilson, J. Freebersyser, and C. Marshall, "Multi-symbol detection of M-DPSK," in *Proc. IEEE GLOBECOM 1989*, Dallas, TX, Nov. 1989, vol. 3, pp. 1692-1697.

[6] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, pp. 300-308, Mar. 1990.

[7] K. M. Mackenthun, Jr., "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 42, pp. 1471-1474, Feb./Mar./Apr. 1994.

[8] W. Sweldens, "Fast block noncoherent decoding," *IEEE Comm. Letters*, vol. 5, pp. 132-134, Apr. 2001.

[9] I. Motedayen-Aval, A. Krishnamoorthy, and A. Anastasopoulos, "Optimal joint detection/estimation in fading channels with polynomial complexity," *IEEE Trans. Inf. Theory*, vol. 53, pp. 209-223, Jan. 2007.

[10] G. N. Karystinos and D. A. Pados, "Rank-2-optimal adaptive design of binary spreading codes," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3075-3080, Sept. 2007.

[11] I. Motedayen-Aval and A. Anastasopoulos, "Polynomial-complexity noncoherent symbol-by-symbol detection with application to adaptive iterative decoding of Turbo-like codes," *IEEE Trans. Commun.*, vol. 51, pp. 197-207, Feb. 2003.

[12] D. J. Ryan, I. B. Collings, and I. V. L. Clarkson, "GLRT-optimal noncoherent lattice decoding," *IEEE Trans. Signal. Proc.*, vol. 55, pp. 3773-3786, July 2007.

[13] R. G. McKilliam, D. J. Ryan, I. V. L. Clarkson, and I. B. Collings, "An improved algorithm for optimal noncoherent QAM detection," in *Proc. 2008 Australian Communications Theory Workshop*, Christchurch, New Zealand, Feb. 2008, pp. 64-68.

[14] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.

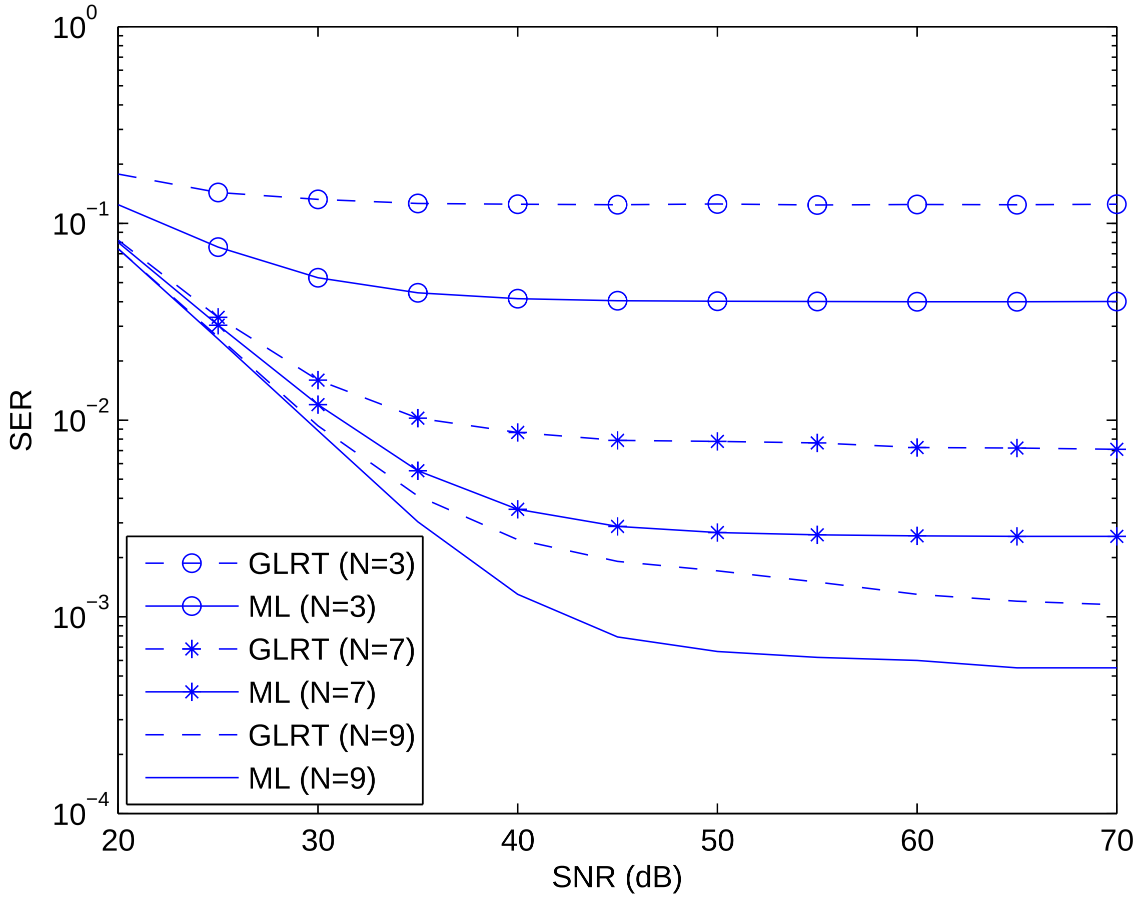

Fig. 1. 4-PAM SER versus SNR for GLRT and ML noncoherent receivers with sequence length $N = 3$, 7, and 9.

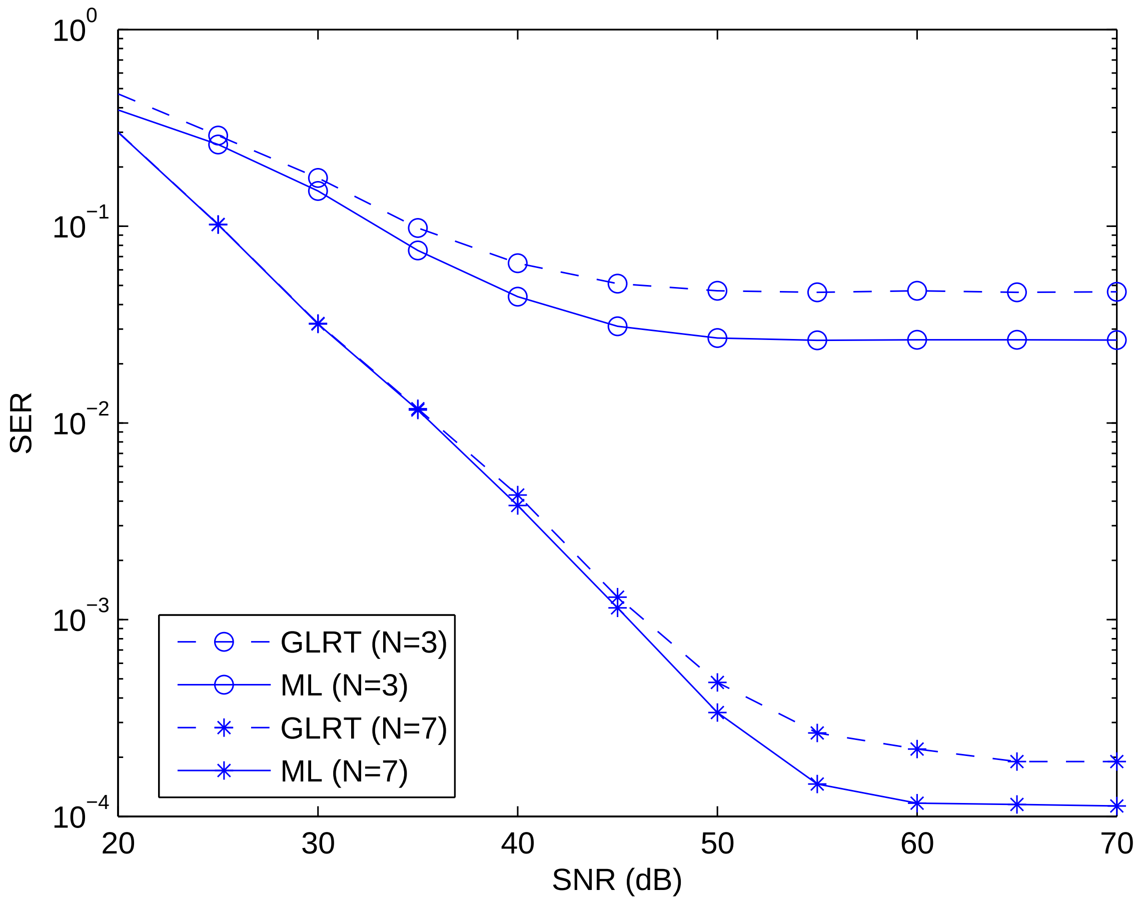

Fig. 2. 8-PAM SER versus SNR for GLRT and ML noncoherent receivers with sequence length $N = 3$ and 7.

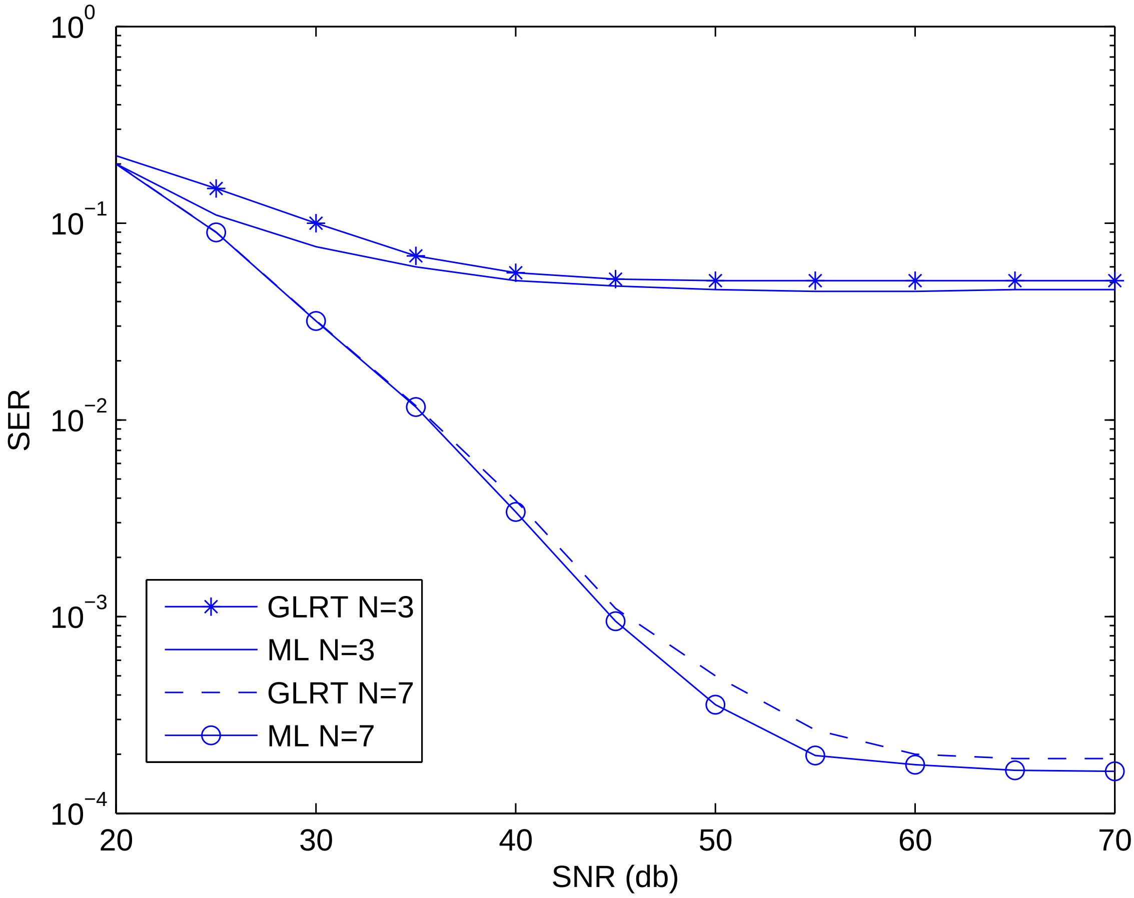

Fig. 3. 16-QAM SER versus SNR for GLRT and ML noncoherent receivers with sequence length $N = 3$ and 7.

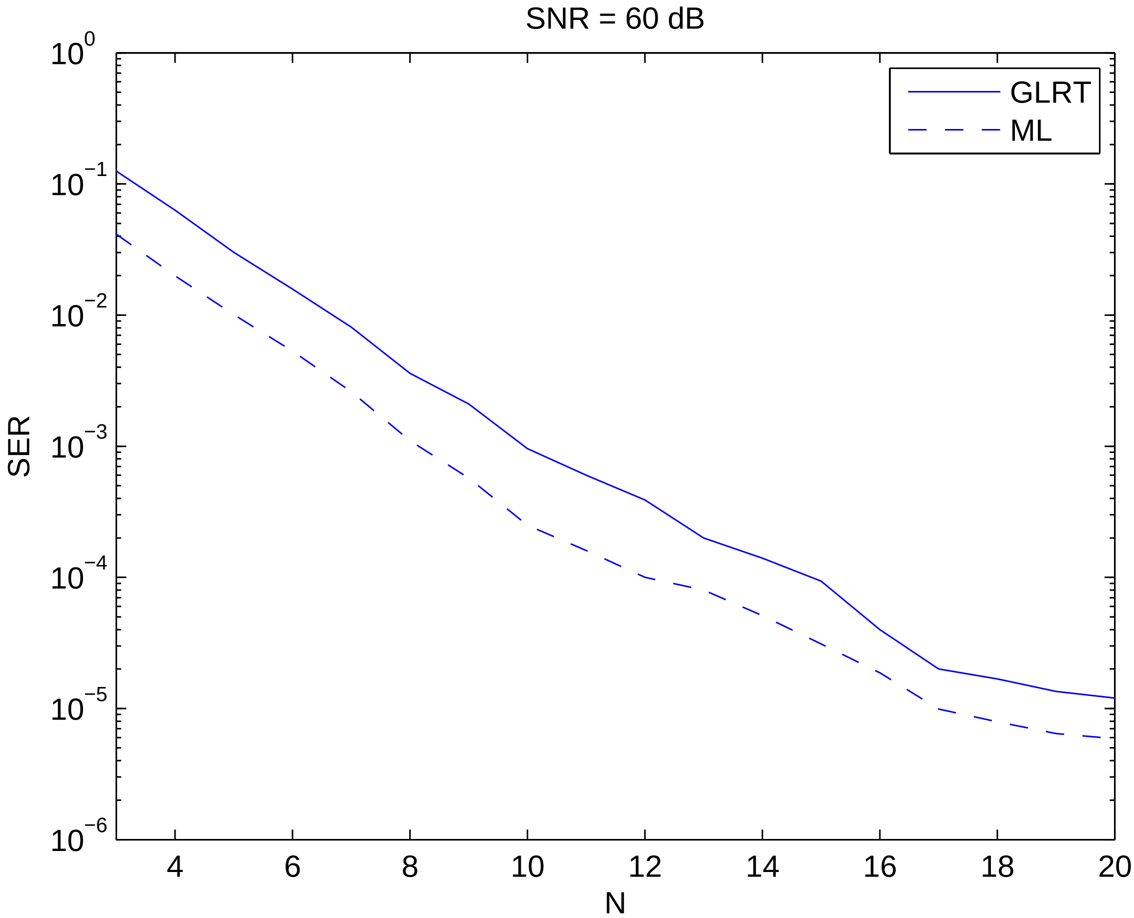

Fig. 4. 4-PAM SER versus N for GLRT and ML noncoherent receivers for SNR = 60 dB.