Optimal Noncoherent Trellis Decoding

Dimitrios G. Chachlakis

School of Electrical and Computer Engineering
Technical University of Crete

Supervisor
Assoc. Prof. George N. Karystinos
Committee members
Assoc. Prof. Aggelos Bletsas
Assist. Prof. Panagiotis P. Markopoulos (RIT,USA)

July 2016

Contents

CHAPTER 1

Introduction

In this diploma thesis, we study the problem of optimal noncoherent trellis decoding, that is, the maximization of $|\mathbf{s}(\mathbf{x})^H \mathbf{y}|$ over $\mathbf{x}$, where $\mathbf{y}$ is a complex vector, $\mathbf{x}$ is a discrete symbol sequence, and $\mathbf{s}(\mathbf{x})$ is a vector that is produced by $\mathbf{x}$ through a trellis structure. Two example cases of noncoherent trellis decoding are noncoherent detection of a minimum-shift keying (MSK) modulated sequence and noncoherent decoding of convolutionally encoded data.

Specifically, MSK is a modulation scheme that limits problems associated with nonlinear distortion and is used in a variety of applications, like signal transmission from satellites and broadcasting [1]. Although the optimal coherent MSK receiver simplifies to constant-complexity symbol-by-symbol detection, optimal noncoherent reception of MSK takes the form of sequence detection [2–6] (due to channel-induced memory) which has exponential (in the sequence length) complexity when implemented through an exhaustive search among all possible sequences. In fact, blind sequence detection may offer significant performance gains in comparison with conventional single-symbol blind detection [7, 8]. This observation was first made in [9–12], in the context of M-ary PSK (MPSK), where it was shown that ML blind sequence detection minimizes the sequence error probability, offering significant error rate performance gains over the conventional symbol-by-symbol blind detection and attaining nearly-coherent detection performance for sufficiently long sequences. This is partly due to the fact that sequence detection exploits the correlation of the received symbols in the entire sequence (due to channel-induced memory), whereas symbol-by-symbol detection does not. However, optimal sequence detection comes at a high price when implemented through an exhaustive search among all possible transmitted data sequences; its complexity is exponential in the sequence length. Convolutional codes are used extensively to achieve reliable data transfer in numerous applications, such as digital video, radio, and satellite communications. They are modeled by a trellis structure and optimal (coherent and noncoherent) convolutional decoding also takes the form of sequence detection.

In this work, we present an algorithm that performs generalized-likelihood-ratio-test (GLRT) optimal noncoherent sequence detection of MSK signals in flat fading with log-linear (in the sequence length) complexity. Moreover, for Rayleigh fading channels, the proposed algorithm is equivalent to the maximum-likelihood (ML) noncoherent sequence detector. Our algorithm utilizes principles that have been used for polynomial-complexity optimization in [14–18] and complements efficient optimal noncoherent detection techniques that have been developed for PSK [14, 15], [19], PAM or QAM [17, 20–22], and FSK [18] signals. We then discuss how the proposed algorithm can be generalized for use on noncoherent convolutional decoding. To simplify the presentation, we consider a particular convolutional code and

modify the proposed algorithm to perform optimal noncoherent trellis decoding with empirically low complexity. Simulation studies indicate that the optimal noncoherent MSK detector attains coherent-detection performance when the sequence length is on the order of 100, offering a 5–6 dB gain over the typical single-symbol detector. Similar results are obtained for the generalized algorithm on convolutional decoding.

CHAPTER 2

MSK

2.1 Signal Model

We consider a binary information bit sequence $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T \in \{\pm 1\}^N$. During the nth single-bit period, the lowpass equivalent signal that is modulated according to MSK is given by [23]

$$s(t) = \sqrt{\frac{2E}{T}} \mathrm{e}^{j\phi[n]} \mathrm{e}^{j x_n \frac{\pi}{2T}(t-nT)}, \quad nT \le t < (n+1)T \tag{2.1}$$

where E and T denote signal strength and nominal duration, respectively. The phase of $s(t)$ at time nT is given by $\phi[n] = \phi(nT)$ where

$$\phi[0] = 0 \text{ and } \phi[n+1] = \phi[n] + x_n \frac{\pi}{2}. \tag{2.2}$$

If the modulated waveform is transmitted through a flat-fading channel, then the received signal, after downconversion, is

$$r(t) = hs(t) + n(t) \tag{2.3}$$

where h is a complex number that models signal attenuation and phase change due to the channel, and $n(t)$ is a zero-mean complex Gaussian process with variance σ_w^2, modeling noise. From the Gram-Schmidt procedure, we derive an orthonormal basis for $s(t)$ with $x_0 = 1$ and $x_0 = -1$, given by

$$\begin{aligned} \psi_1(t) &= \frac{1}{\sqrt{T}} \mathrm{e}^{j\frac{\pi t}{2T}}, \quad 0 \le t < T, \\ \psi_2(t) &= \frac{\pi}{\sqrt{T(\pi^2-4)}} \left(\mathrm{e}^{-j\frac{\pi t}{2T}} + \frac{2j}{\pi} \mathrm{e}^{j\frac{\pi t}{2T}} \right), \quad 0 \le t < T. \end{aligned} \tag{2.4}$$

By correlating $s(t)$ with $\{\psi_i(t-nT)\}_{i=1}^2$, we obtain the signal constellation

$$\mathbf{s}^1 = \begin{bmatrix} \sqrt{2E} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{s}^{-1} = \begin{bmatrix} -\frac{2j\sqrt{2E}}{\pi} \\ \frac{\sqrt{2E(\pi^2-4)}}{\pi} \end{bmatrix} \tag{2.5}$$

for $x_n = 1$ and $x_n = -1$, respectively. Hence, the nth transmitted signal vector is

$$\mathbf{s}_n = \mathbf{s}^{x_n} \mathrm{e}^{j\phi[n]}. \tag{2.6}$$

During the nth bit transmission, $n = 0, \ldots, N-1$, the optimal receiver correlates the received signal $r(t)$ with $\{\psi_i(t-nT)\}_{i=1}^2$ and produces two samples [23]

$$r_i = \int_{nT}^{nT+T} r(t) \psi_i^*(t-nT) \mathrm{d}t, \ i = 1, 2. \tag{2.7}$$

Consequently, the nth received vector becomes

$$\mathbf{r}_n \triangleq \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = h\mathbf{s}_n + \mathbf{n}_n \tag{2.8}$$

where $\mathbf{n}_n \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_2)$.

If $\mathbf{r}_0, \ldots, \mathbf{r}_{N-1}$ are the received vectors (per information bit) given by (2.8), then we may form the received vector for the entire sequence $\mathbf{x}$ as

$$\mathbf{r} \triangleq \begin{bmatrix} \mathbf{r}_0 \\ \vdots \\ \mathbf{r}_{N-1} \end{bmatrix} = h \underbrace{\begin{bmatrix} \mathbf{s}_0 \\ \vdots \\ \mathbf{s}_{N-1} \end{bmatrix}}_{\mathbf{s}} + \mathbf{w} \tag{2.9}$$

where $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I}_{2N})$, $\mathbf{s}$ is the column-wise concatenation of the N transmitted signal vectors $\mathbf{s}_0, \ldots, \mathbf{s}_{N-1}$ defined by (2.6), and $\mathbf{r}$ is the column-wise concatenation of the N received signal vectors $\mathbf{r}_0, \ldots, \mathbf{r}_{N-1}$.

2.2 Optimal Coherent Detection

The ML coherent detector maximizes the conditional probability density function (pdf) $f(\mathbf{r}|h, \mathbf{s})$ of the observation vector given the transmitted symbol sequence $\mathbf{s}$ and the channel coefficient h, i.e.,

$$\hat{\mathbf{s}}_{\text{coh}}^{\text{ML}} = \arg\max_{\mathbf{s}} f(\mathbf{r}_0, \ldots, \mathbf{r}_{N-1}|h, \mathbf{s}) = \arg\max_{\mathbf{s}} \sum_{n=0}^{N-1} \Re\left\{(h\mathbf{s}_n)^H \mathbf{r}_n\right\}. \tag{2.10}$$

The trellis diagram at the receiver is shown in Fig. 2.1.

By the trellis structure, we derive a couple of useful lemmas, the proofs of which are shown in the appendix.

Lemma 1. *The optimal coherent sequence detection is equivalent to symbol by symbol detection. To optimally decide in favor of the transmitted symbol $\mathbf{s}_n$, it suffices to know $\mathbf{r}_{n-1}, \mathbf{r}_n, \mathbf{r}_{n+1}$, and the channel coefficient h.*

Lemma 2. *At any given time, the two most likely winning paths originate from the same node.*

Due to Lemma 1, the optimal coherent MSK receiver simplifies to constant-complexity symbol-by-symbol detection.

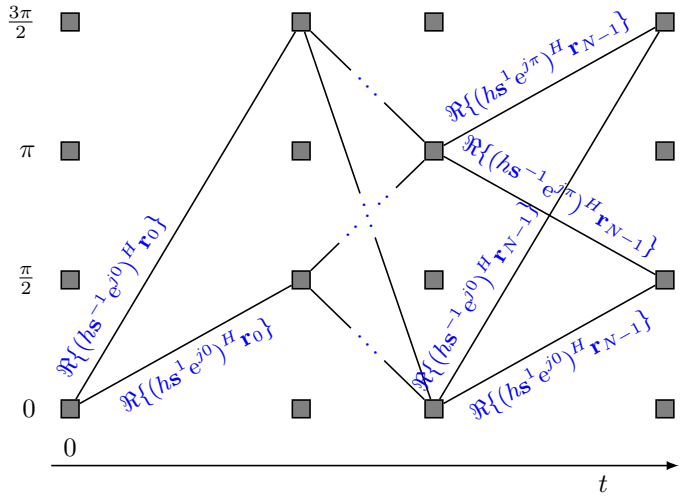

Figure 2.1: Trellis diagram at the receiver.

2.3 Optimal Noncoherent Detection

If the actual channel realization is not available to the receiver and is modeled as a random variable, then this variable appears in consecutive received vectors, hence the channel induces memory in addition to memory already present due to MSK modulation. As a result, optimal detection requires processing of the entire sequence of N received vectors.

The ML noncoherent detector maximizes the conditional pdf of $\mathbf{r}$ given $\mathbf{s}$, that is, the optimal decision is given by

$$\hat{\mathbf{s}}^{\mathrm{ML}} = \arg\max_{\mathbf{s}} f(\mathbf{r}|\mathbf{s}). \tag{2.11}$$

Note that $\|\mathbf{s}\|^2 = 2EN$ for any $\mathbf{s}$. It can be shown that, for Rayleigh fading (i.e., $h \sim \mathcal{CN}(0, \sigma_h^2)$), the received symbol vector $\mathbf{r}$ given the transmitted sequence $\mathbf{s}$ follows a proper complex Gaussian distribution with mean $\mathbb{E}[\mathbf{r}|\mathbf{s}] = \mathbf{0}$ and covariance matrix

$$\mathrm{C}_{\mathbf{r}|\mathbf{s}} \triangleq \mathbb{E}[\mathbf{r}\mathbf{r}^H|\mathbf{s}] = \sigma_h^2 \mathbf{s}\mathbf{s}^H + \sigma_w^2 \mathbf{I}_{2N}. \tag{2.12}$$

Hence, the optimization problem in (2.11) can be rewritten as

$$\hat{\mathbf{s}}^{\mathrm{ML}} = \arg\max_{\mathbf{s}} \frac{1}{\pi^{2N}|\mathrm{C}_{\mathbf{r}|\mathbf{s}}|} \mathrm{e}^{-\mathbf{r}^H \mathrm{C}_{\mathbf{r}|\mathbf{s}}^{-1} \mathbf{r}}. \tag{2.13}$$

Exploiting identities for the determinant and inverse of a rank-1 update [13] and the fact that $\|\mathbf{s}\|^2 = 2EN$, we obtain

$$|\mathrm{C}_{\mathbf{r}|\mathbf{s}}| = \sigma_w^{4N-2} \left(\sigma_w^2 + \sigma_h^2 \|\mathbf{s}\|^2\right) \tag{2.14}$$

and

$$\mathrm{C}_{\mathbf{r}|\mathbf{s}}^{-1} = \frac{1}{\sigma_w^2} \mathbf{I}_{2N} - \frac{\sigma_h^2}{\sigma_w^4 + \sigma_h^2 \sigma_w^2 \|\mathbf{s}\|^2} \mathbf{s}\mathbf{s}^H. \tag{2.15}$$

If we substitute (2.14) and (2.15) in (2.13), then we obtain

$$\hat{\mathbf{s}}^{\mathrm{ML}} = \arg\max_{\mathbf{s}} |\mathbf{s}^H \mathbf{r}|. \tag{2.16}$$

Substituting $\mathbf{s} = [\mathbf{s}_0^H, \mathbf{s}_1^H, \ldots, \mathbf{s}_{N-1}^H]^H$ in (2.16), the ML rule is rewritten in terms of the information sequence $\mathbf{x}$ as[1]

$$\hat{\mathbf{x}}^{\mathrm{ML}} = \arg\max_{\mathbf{x}\in\{\pm 1\}^N} \left|\mathbf{s}_0^H \mathbf{r}_0 + \ldots + \mathbf{s}_{N-1}^H \mathbf{r}_{N-1}\right|. \tag{2.17}$$

The equivalent expressions (2.16) and (2.17) represent the optimal decision rule for a Rayleigh-distributed channel.

If, on the other hand, the channel distribution is not Rayleigh or is unknown, then we may consider joint channel estimation and data detection, i.e., GLRT sequence detection, according to which,

$$\begin{aligned}\hat{\mathbf{s}}^{\mathrm{GLRT}} &= \arg\min_{\mathbf{s}} \left\{ \min_{h\in\mathbb{C}} \|\mathbf{r} - h\mathbf{s}\|^2 \right\} = \arg\min_{\mathbf{s}} \left\| \mathbf{r} - \frac{\mathbf{s}^H\mathbf{r}}{\|\mathbf{s}\|^2}\mathbf{s} \right\|^2 = \arg\max_{\mathbf{s}} \frac{|\mathbf{r}^H\mathbf{s}|^2}{\|\mathbf{s}\|^2} \\ &= \arg\max_{\mathbf{s}} \left|\mathbf{s}^H\mathbf{r}\right|. \end{aligned} \tag{2.18}$$

Hence, the ML optimization problem in (2.16) and the GLRT optimization problem in (2.18) are equivalent. This equivalence between noncoherent ML and GLRT has also been demonstrated in the context of equal-energy signals and uniformly distributed over $[0, 2\pi)$ channel phase [7, 20]. A straightforward approach to solve (2.16) and (2.18) (or, equivalently, (2.17)) would be an exhaustive search among all 2^N sequences $\mathbf{x} \in \{\pm 1\}^N$. However, such a solver would be impractical even for moderate values of N, since its complexity is on the order of $\mathcal{O}(2^N)$, i.e., it grows exponentially with N. In the following, we present an algorithm that solves the above problems with log-linear complexity $\mathcal{O}(N \log N)$.

2.3.1 Log-linear optimal noncoherent detection

To present the proposed algorithm, we begin by rewriting the optimal detection rule in (2.17) as

$$\begin{aligned}&\max_{\mathbf{x}\in\{\pm 1\}^N} \left|\mathbf{s}_0^H \mathbf{r}_0 + \ldots + \mathbf{s}_{N-1}^H \mathbf{r}_{N-1}\right| \\ &= \max_{\mathbf{x}\in\{\pm 1\}^N} \max_{\phi\in[0,2\pi)} \Re\left\{ \mathrm{e}^{-j\phi} \left(\mathbf{s}_0^H \mathbf{r}_0 + \ldots + \mathbf{s}_{N-1}^H \mathbf{r}_{N-1}\right)\right\} \\ &= \max_{\phi\in[0,2\pi)} \max_{\mathbf{x}\in\{\pm 1\}^N} \Re\left\{ \mathrm{e}^{-j\phi} \left(\mathbf{s}_0^H \mathbf{r}_0 + \ldots + \mathbf{s}_{N-1}^H \mathbf{r}_{N-1}\right)\right\}. \end{aligned} \tag{2.19}$$

That is, to find the optimal sequence $\hat{\mathbf{x}}^{\mathrm{ML}}$ in (2.17), we may let ϕ scan the interval $[0, 2\pi)$ and, for each value of ϕ, collect the corresponding vector

$$\hat{\mathbf{x}}(\phi) = \arg\max_{\mathbf{x}\in\{\pm 1\}^N} \Re\left\{ \mathrm{e}^{-j\phi} \left(\mathbf{s}_0^H \mathbf{r}_0 + \ldots + \mathbf{s}_{N-1}^H \mathbf{r}_{N-1}\right)\right\} \tag{2.20}$$

that solves the innermost maximization in (2.19). Then,

$$\hat{\mathbf{x}}^{\mathrm{ML}} \in \mathcal{X} \triangleq \left\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_{|\mathcal{X}|}\right\} = \bigcup_{\phi\in[0,2\pi)} \left\{\hat{\mathbf{x}}(\phi)\right\}. \tag{2.21}$$

Interestingly, as we will show below, the size of the set in (2.21) is $2N$ and can be built with log-linear complexity.

Let ϕ have a fixed value in $[0, 2\pi)$, say $\phi = 0$. Then, (2.20) can be solved efficiently utilizing the Viterbi Algorithm (VA). Note that, since ϕ is fixed (we have, without loss of generality assumed that

[1] Note that $\mathbf{s}$ depends on $\mathbf{x}$, since $\mathbf{s}_n$ contains x_n due to (2.6).

$\phi = 0$),((2.20)) is equivalent to performing coherent detection on $\mathbf{x}$, i.e., it is equivalent to (2.10). Due to (2.21), the decision sequence $\hat{\mathbf{x}}\,(\phi = 0)$ will be equal to $\hat{\mathbf{x}}_m$ for some $m \in \{1, 2, \ldots, |\mathcal{X}|\}$, where

$$\hat{\mathbf{x}}_m = [\hat{x}_0, \ldots, \hat{x}_n, \hat{x}_{n+1}, \ldots, \hat{x}_{N-1}]^T \tag{2.22}$$

and $\hat{x}_i$, $i = 0, 1, \ldots, N-1$, is the decision on the ith information bit x_i of the sequence of N consecutive bits. Based on $\hat{\mathbf{x}}_m$ given by (2.22), we construct the phase vector

$$\hat{\boldsymbol{\phi}}_m = \left[\hat{\phi}[0] = 0, \ldots, \hat{\phi}[n], \hat{\phi}[n+1], \ldots, \hat{\phi}[N-1]\right]^T, \tag{2.23}$$

using the iteration in (2.2).

As ϕ scans the interval $[0, 2\pi)$, the decision on $\hat{\mathbf{x}}\,(\phi)$ will change from $\hat{\mathbf{x}}_m$ to $\hat{\mathbf{x}}_k$ for some $k \in \{1, 2, \ldots, |\mathcal{X}|\} \setminus \{m\}$. An important observation is that successive decision sequences (in this case, $\hat{\mathbf{x}}_m$ and $\hat{\mathbf{x}}_k$) will differ in exactly two consecutive bit decisions $\hat{x}_n$ and $\hat{x}_{n+1}$, i.e.,[2]

$$\begin{aligned}
\hat{\mathbf{x}}_k &= \left[\hat{x}_0, \ldots, \hat{x}_{n-1}, \hat{x}_n^{\mathrm{c}}, \hat{x}_{n+1}^{\mathrm{c}}, \hat{x}_{n+2}, \ldots, \hat{x}_{N-1}\right]^T, \\
\hat{\boldsymbol{\phi}}_k &= \left[\hat{\phi}[0], \ldots, \hat{\phi}[n], \hat{\phi}[n+1] + \pi, \hat{\phi}[n+2], \ldots, \hat{\phi}[N-1]\right]^T,
\end{aligned} \tag{2.24}$$

for some $n = 0, 1, \ldots, N-1$, or in the last one ($\hat{x}_{N-1}$).[3] The decision on $\mathbf{x}$ changes (from $\hat{\mathbf{x}}_m$ to $\hat{\mathbf{x}}_k$), according to (2.20) and utilizing (2.6), for some $n = 0, 1, \ldots, N-1$, only when[4,5]

$$\begin{aligned}
&\sum_{i=0}^{N-1} \Re\left\{ \mathrm{e}^{-j\phi} \left(\mathbf{s}^{\hat{x}_m[i]} \mathrm{e}^{j\hat{\phi}_m[i]} - \mathbf{s}^{\hat{x}_k[i]} \mathrm{e}^{j\hat{\phi}_k[i]} \right)^H \mathbf{r}_i \right\} = 0 \\
\Leftrightarrow &\sum_{i=n,n+1} \Re\left\{ \mathrm{e}^{-j\phi} \left(\mathbf{s}^{\hat{x}_m[i]} \mathrm{e}^{j\hat{\phi}_m[i]} - \mathbf{s}^{\hat{x}_m^{\mathrm{c}}[i]} \mathrm{e}^{j\hat{\phi}_k[i]} \right)^H \mathbf{r}_i \right\} = 0 \\
\Leftrightarrow &\cos\left(\phi - \underline{\angle v_1 - v_2}\right) = 0 \\
\Leftrightarrow &\phi = \underbrace{\pm\frac{\pi}{2} + \underline{\angle v_1 - v_2} \pmod{2\pi}}_{\phi_n^{(1)}, \phi_n^{(2)}}
\end{aligned} \tag{2.25}$$

where

$$\begin{aligned}
v_1 &\triangleq \left(\mathbf{s}^{\hat{x}_m[n]} \mathrm{e}^{j\hat{\phi}_m[n]} \right)^H \mathbf{r}_n + \left(\mathbf{s}^{\hat{x}_m[n+1]} \mathrm{e}^{j\hat{\phi}_m[n+1]} \right)^H \mathbf{r}_{n+1}, \\
v_2 &\triangleq \left(\mathbf{s}^{\hat{x}_m^{\mathrm{c}}[n]} \mathrm{e}^{j\hat{\phi}_m[n]} \right)^H \mathbf{r}_n + \left(\mathbf{s}^{\hat{x}_m^{\mathrm{c}}[n+1]} \mathrm{e}^{j\hat{\phi}_m[n+1]+j\pi} \right)^H \mathbf{r}_{n+1}.
\end{aligned} \tag{2.26}$$

Hence, for $n = 0, 1, \ldots, N-1$, from (2.25) we collect $2N$ distinct phases

$$\phi_0^{(1)}, \phi_0^{(2)}, \ldots, \phi_{N-1}^{(1)}, \phi_{N-1}^{(2)}. \tag{2.27}$$

Lemma 3. *The produced $2N$ phases among any consecutive pair of sequence decision on $\mathbf{x}$, are identical to those in* (2.27).

---

[2]Since the constellation is binary, the notation $(\cdot)^{\mathrm{c}}$ is utilized for the relation $x^{\mathrm{c}} = -x$. That is, $1^{\mathrm{c}} = -1$ and $(-1)^{\mathrm{c}} = 1$.

[3]In the case where $n = N-1$, $\hat{\mathbf{x}}_k = [\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{N-2}, \hat{x}_{N-1}^{\mathrm{c}}]^T$ and $\hat{\boldsymbol{\phi}}_k = \hat{\boldsymbol{\phi}}_m$.

[4]We use the notation $\hat{x}_k[i]$ and $\hat{\phi}_k[i]$ to refer to the ith element of the $N \times 1$ vectors $\hat{\mathbf{x}}_k$ and $\hat{\boldsymbol{\phi}}_k$, respectively.

[5]For simplicity purposes, we abuse the notation in the following sense; in the case where $n = N-1$, $\mathbf{r}_{n+1}, \hat{\phi}_k[n+1], \hat{\phi}_m[n+1], \hat{x}_m[n+1], \hat{x}_m^{\mathrm{c}}[n+1]$, and $\hat{x}_k[n+1]$ are not defined and hence are assumed to be equal to 0 or $\mathbf{0}$, depending on their dimension. Hence, when $n = N-1$, (2.26) is simplified since the second term of each summation is equal to 0.

Lemma 3 (the proof of which is shown at the appendix) implies that it suffices to compute the $2N$ phases only once. Hence, if we sort the phases in (2.27) in ascending order, i.e.,

$$(\theta_0, \theta_1, \ldots, \theta_{2N-1}) = \text{sort}\left(\phi_0^{(1)}, \phi_0^{(2)}, \ldots, \phi_{N-1}^{(1)}, \phi_{N-1}^{(2)}\right), \tag{2.28}$$

then we partition the entire interval $[0, 2\pi)$ into $2N$ disjoint intervals, i.e.,

$$[0, \theta_0), (\theta_0, \theta_1), \ldots, (\theta_{2N-2}, \theta_{2N-1}), \tag{2.29}$$

in each of which the sequence decision $\hat{\mathbf{x}}$ remains constant. Note that we ignore the last interval $(\theta_{2N-1}, 2\pi)$ because it produces the same decision sequence $\hat{\mathbf{x}}$ with the interval $[0, \theta_0)$.

All the above lead to the following algorithm.

1. Set $\phi = 0$ and identify $\hat{\mathbf{x}} = [\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{N-1}]^T$ by (2.10) with $h = 1$. This step has complexity $\mathcal{O}(N)$.

2. For any $n = 0, 1, \ldots, N-1$, based on $\hat{x}_n$, construct $\phi[n]$ through (2.2). Then, $\hat{x}_n$ and $\phi[n]$ are used to construct $\mathbf{s}_n$ through (2.6). Once $\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_{N-1}$ are available, calculate the metric of interest in (2.17) with complexity $\mathcal{O}(N)$.

3. Set the current $\hat{\mathbf{x}}$ as the best sequence, i.e., $\hat{\mathbf{x}}^{\text{ML}} = \hat{\mathbf{x}}$, and store its metric from Step 2 as the ML_value.

4. For each $n = 0, 1, \ldots, N-1$, use $\hat{x}_n$ and $\phi[n]$ from Step 1 and evaluate (2.26) to obtain v_1 and v_2 which are plugged into (2.25). As a result, from (2.25), obtain the $2N$ phases $\phi_0, \phi_1, \ldots, \phi_{2N-1}$.

5. Sort the $2N$ phases from Step 4 in ascending order to obtain $\theta_0, \theta_1, \ldots, \theta_{2N-1}$.

6. For each $i = 0, 1, \ldots, 2N-1$, repeat the following Steps 7–9.

7. Move to θ_i, get the index k for which $\theta_i = \phi_k$ and set it equal to $k \pmod N$. Update the current $\hat{\mathbf{x}}$ by setting $\hat{x}_k$ and $\hat{x}_{k+1}$ equal to $\hat{x}_k^{\text{c}}$ and $\hat{x}_{k+1}^{\text{c}}$, respectively. Also set $\phi[k+1] = \phi[k+1] + \pi \pmod{2\pi}$.

8. Update $\mathbf{s}_k$ and $\mathbf{s}_{k+1}$ through (2.6) using the modified $\hat{\mathbf{x}}$ and $\phi[k+1]$ from Step 7. Then, evaluate (2.17) and store the new metric as $\text{value\_}\hat{\mathbf{x}}$.

9. Compare $\text{value\_}\hat{\mathbf{x}}$ with ML_value. If $\text{value\_}\hat{\mathbf{x}}$ is greater than ML_value, then set the current vector $\hat{\mathbf{x}}$ as $\hat{\mathbf{x}}^{\text{ML}}$ and set the ML_value to $\text{value\_}\hat{\mathbf{x}}$.

The pseudo-code of the proposed ML/GLRT noncoherent MSK sequence detection algorithm is illustrated in Fig. 2.2. The overall complexity of the proposed algorithm is dominated by the computational cost of the phase sorting at line 19, which is $\mathcal{O}(N \log N)$.

Algorithm 1 Optimal Blind MSK Detection in Time $\mathcal{O}(N \log N)$

Input: $\mathbf{r}, \mathbf{s}^1, \mathbf{s}^{-1}$

1: Get $\hat{\mathbf{x}} = [\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{N-1}]^T$ for $\phi = 0$.
2: $\phi[0] = 0$
3: **for** $n = 1 : N-1$ **do**
4: $\quad \phi[n] = \phi[n-1] + \hat{x}_{n-1}\frac{\pi}{2}$.
5: **end for**
6: value_$\hat{\mathbf{x}} = \sum_{n=0}^{N-1} \left(\mathbf{s}^{\hat{x}_n} \mathrm{e}^{j\phi[n]}\right)^H \mathbf{r}_n$
7: ML_value $= |$value_$\hat{\mathbf{x}}|$
8: $\hat{\mathbf{x}}^{\mathrm{ML}} = \hat{\mathbf{x}}$
9: **for** $n = 0 : N-1$ **do**
10: $\quad$ **if** $n < N-1$ **then**
11: $\quad\quad v_1 = \left(\mathbf{s}^{\hat{x}_n} \mathrm{e}^{j\phi[n]}\right)^H \mathbf{r}_n + \left(\mathbf{s}^{\hat{x}_{n+1}} \mathrm{e}^{j\phi[n+1]}\right)^H \mathbf{r}_{n+1}$
12: $\quad\quad v_2 = \left(\mathbf{s}^{\hat{x}^c_n} \mathrm{e}^{j\phi[n]}\right)^H \mathbf{r}_n + \left(\mathbf{s}^{\hat{x}^c_{n+1}} \mathrm{e}^{j\phi[n+1]+j\pi}\right)^H \mathbf{r}_{n+1}$
13: $\quad$ **else**
14: $\quad\quad v_1 = \left(\mathbf{s}^{\hat{x}_n} \mathrm{e}^{j\phi[n]}\right)^H \mathbf{r}_n$
15: $\quad\quad v_2 = \left(\mathbf{s}^{\hat{x}^c_n} \mathrm{e}^{j\phi[n]}\right)^H \mathbf{r}_n$
16: $\quad$ **end if**
17: $\quad (\phi_n, \phi_{n+N}) = \pm\frac{\pi}{2} + \underline{/v_1 - v_2} \pmod{2\pi}$
18: **end for**
19: $(\theta_0, \theta_1, \ldots, \theta_{2N-1}) = \mathrm{sort}\,(\phi_0, \phi_1, \ldots, \phi_{2N-1})$
20: **for** $i = 0 : 2N-1$ **do**
21: $\quad$ Let k be the index for which $\theta_i = \phi_k$ at line 19.
22: $\quad k = k \pmod{N}$
23: $\quad$ **if** $k < N-1$ **then**
24: $\quad\quad$ value_$\hat{\mathbf{x}} =$ value_$\hat{\mathbf{x}} - \left(\mathbf{s}^{\hat{x}_k} \mathrm{e}^{j\phi[k]}\right)^H \mathbf{r}_k - \left(\mathbf{s}^{\hat{x}_{k+1}} \mathrm{e}^{j\phi[k+1]}\right)^H \mathbf{r}_{k+1}$
$\qquad\qquad + \left(\mathbf{s}^{\hat{x}^c_k} \mathrm{e}^{j\phi[k]}\right)^H \mathbf{r}_k + \left(\mathbf{s}^{\hat{x}^c_{k+1}} \mathrm{e}^{j\phi[k+1]+j\pi}\right)^H \mathbf{r}_{k+1}$
25: $\quad\quad \hat{\mathbf{x}}(k:k+1) = \hat{\mathbf{x}}^c(k:k+1)$
26: $\quad\quad \phi[k+1] = \phi[k+1] + \pi \pmod{2\pi}$
27: $\quad$ **else**
28: $\quad\quad$ value_$\hat{\mathbf{x}} =$ value_$\hat{\mathbf{x}} - \left(\mathbf{s}^{\hat{x}_k} \mathrm{e}^{j\phi[k]}\right)^H \mathbf{r}_k + \left(\mathbf{s}^{\hat{x}^c_k} \mathrm{e}^{j\phi[k]}\right)^H \mathbf{r}_k$
29: $\quad\quad \hat{\mathbf{x}}(k) = \hat{\mathbf{x}}^c(k)$
30: $\quad$ **end if**
31: $\quad$ best_value $= |$value_$\hat{\mathbf{x}}|$
32: $\quad$ **if** best_value$>$ ML_value **then**
33: $\quad\quad$ ML_value=best_value
34: $\quad\quad \hat{\mathbf{x}}^{\mathrm{ML}} = \hat{\mathbf{x}}$
35: $\quad$ **end if**
36: **end for**

Output: $\hat{\mathbf{x}}^{\mathrm{ML}}$

Figure 2.2: ML/GLRT noncoherent sequence detection of MSK.

2.4 Simulation Results

We consider MSK transmissions through a Rayleigh flat-fading channel with $\sigma_h^2 = 1$. In Fig. 2.3, we plot the bit error rate (BER) of the optimal noncoherent sequence detector as a function of the transmitted signal-to-noise ratio (SNR), for sequence length $N = 1, 2, 4, 100$. We include the BER of the conventional ML coherent detector, as a reference. We observe that, as the sequence length increases, the noncoherent detector approaches the coherent one in terms of BER. Moreover, the BER of the conventional noncoherent detector (i.e., $N = 1$) is 5–6dB far from the coherent one; as the sequence length N increases, the BER gap

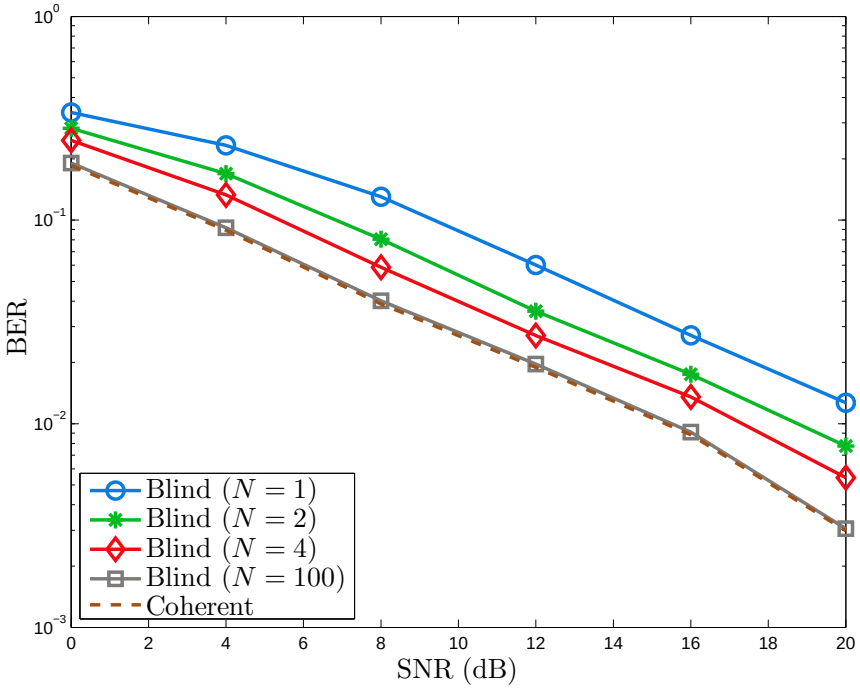

Figure 2.3: MSK BER versus bit SNR for ML/GLRT noncoherent detection with sequence length $N = 1, 2, 4, 100$ and ML coherent detection.

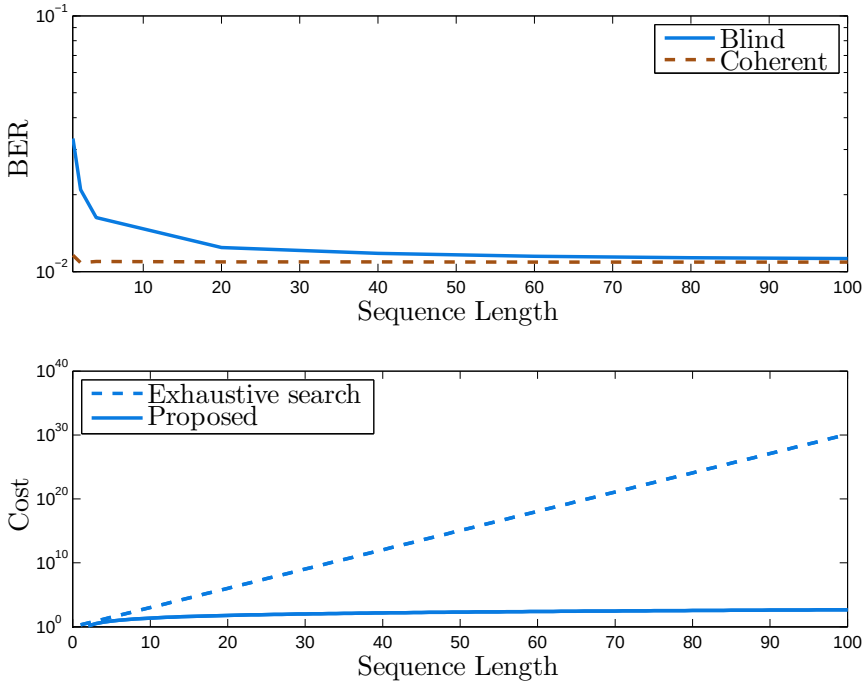

Figure 2.4: Noncoherent MSK detection BER and cost versus sequence length (SNR=15dB).

decreases to zero. To demonstrate the rate of convergence to coherent detection performance, in Fig. 2.4, we set the SNR to 15dB and plot the BER of the optimal noncoherent detector and the computational cost of the proposed algorithm and the conventional exhaustive-search approach as a function of the sequence length N. We note that the BER of the noncoherent scheme with $N = 100$ is nearly equal to the BER of the coherent one with perfect channel knowledge. Interestingly, this is achieved with complexity on the

order of $100\log_2 100 \simeq 700$ computations while the conventional exhaustive search would require 2^{100} computations.

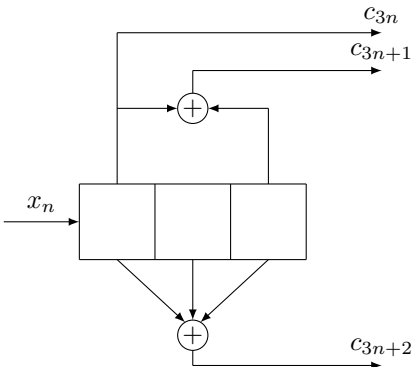

Figure 3.1: $(K = 3, k = 1, n = 3)$ convolutional encoder.

CHAPTER 3

Convolutional Coding

3.1 Signal Model

We consider an information bit sequence $\mathbf{x} = \left[x_0, x_1, \ldots, x_{N-1}\right]^T \in \{0,1\}^N$ which undergoes convolutional encoding according to the encoder shown in Fig. 3.1 [23]. The convolutional encoder actually encodes the information sequence $\tilde{\mathbf{x}} = \left[\mathbf{x}^T, 0, 0\right]^T$ in order to empty the register of the encoder in Fig. 3.1. The sequence to be transmitted is the resulting encoded vector $\mathbf{c} \triangleq \left[c_0, c_1, \ldots, c_{L-1}\right]^T \in \{0,1\}^L$ where $L = 3N + 6$. If we define the encoded subvector $\mathbf{c}_n$ (that corresponds to the information bit x_n) as

$$\mathbf{c}_n = \begin{bmatrix} c_{3n} \\ c_{3n+1} \\ c_{3n+2} \end{bmatrix} \in \{0,1\}^3 \quad , \quad n = 0, 1, \ldots, N+1, \tag{3.1}$$

then we may rewrite the encoded sequence $\mathbf{c}$ as

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{N+1} \end{bmatrix}. \tag{3.2}$$

For the transmission of vector $\mathbf{c}$, we consider binary phase-shift keying (BPSK) [23]. During the lth coded-bit period, $l = 0, 1, \ldots, L-1$, the low-pass equivalent BPSK signal is given by

$$s(t) = \sqrt{\frac{2E}{T}}(1 - 2c_l)p(t - lT) \quad , \quad lT \leq t < (l+1)T, \tag{3.3}$$

where E and T represent signal energy and nominal duration, respectively, and $p(t)$ is a shaping pulse with duration from 0 to T, normalized to unit energy. If the modulated waveform is transmitted through a flat-fading channel, then the received signal, after downconversion, is

$$r(t) = hs(t) + n(t) \tag{3.4}$$

where h is a complex number that models signal attenuation and phase change due to the channel and $n(t)$ is a zero-mean complex Gaussian process with variance σ_w^2 modelling noise. The signal space can be represented from the single basis function $p(t)$.

During the lth coded-bit transmission, the optimal receiver correlates the received signal $r(t)$ with the basis function $p(t)$ and produces the received symbol

$$r_l = \int_{lT}^{lT+T} r(t)p(t - lT)dt, \quad l = 0, 1, \ldots, L-1. \tag{3.5}$$

Consequently, the lth received symbol becomes

$$r_l = h(1 - 2c_l)\sqrt{2E} + n_l = \begin{cases} h\sqrt{2E} + n_l & , c_l = 0, \\ -h\sqrt{2E} + n_l & , c_l = 1, \end{cases} \tag{3.6}$$

where $n_l \sim \mathcal{CN}(0, \sigma_w^2)$.

From the received symbols (per coded bit) $r_0, r_1, \ldots, r_{L-1}$ given by (3.6), we may form the received vector for the entire received coded sequence $\mathbf{c}$ as

$$\mathbf{r} = \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{L-1} \end{bmatrix} = h\sqrt{2E}(\mathbf{1} - 2\mathbf{c}) + \mathbf{w} \tag{3.7}$$

where $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_w^2\mathbf{I}_L)$ and $\mathbf{r}$ is the column-wise concatenation of the L received signal symbols $r_0, r_1, \ldots, r_{L-1}$. If we define the received vector per information sequence bit

$$\mathbf{r}_n \triangleq \begin{bmatrix} r_{3n} \\ r_{3n+1} \\ r_{3n+2} \end{bmatrix}, \quad n = 0, 1, \ldots, N+1, \tag{3.8}$$

then we may rewrite the received vector $\mathbf{r}$ in correspondence to the information bit sequence $(x_0, x_1, \ldots, x_{N-1}, 0, 0)$ as

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_{N+1} \end{bmatrix}. \tag{3.9}$$

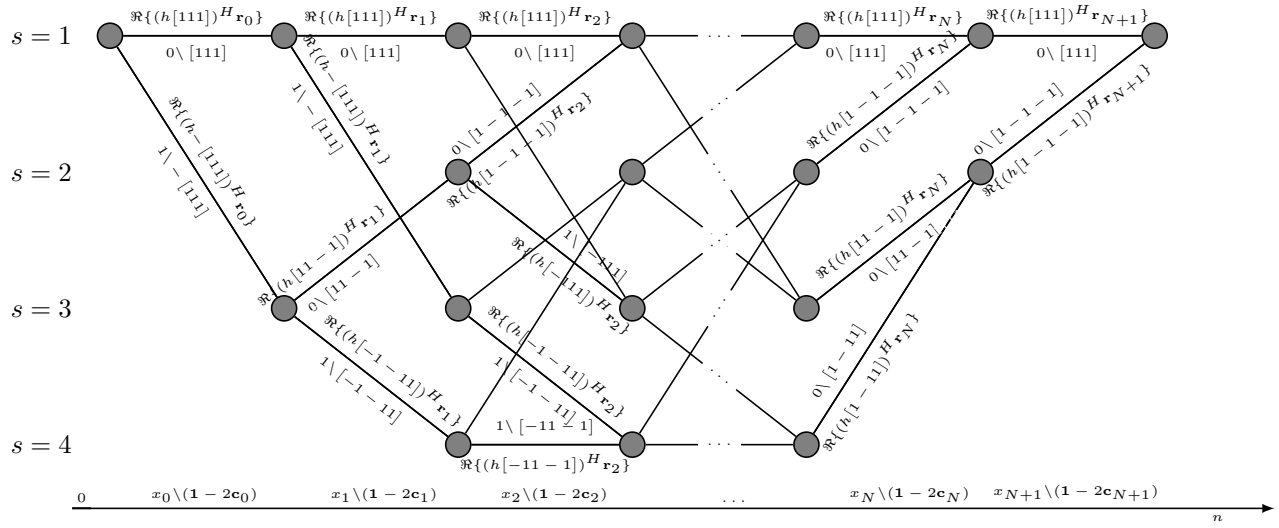

Figure 3.2: Trellis diagram at the receiver.

3.2 Optimal Coherent Decoder

The ML coherent decoder maximizes the conditional probability density function (pdf) $f(\mathbf{r} \mid h, \mathbf{c})$ of the observation vector given the transmitted coded symbol sequence $\mathbf{c}$ and the channel coefficient h, i.e.,

$$\begin{aligned}\hat{\mathbf{c}}_{\text{coh}}^{\text{ML}} &= \arg\max_{\mathbf{c}} f(\mathbf{r}_0, \ldots, \mathbf{r}_{N+1} \mid h, \mathbf{c}) = \arg\max_{\mathbf{c}} \left\{ (\mathbf{1}-2\mathbf{c})^T \Re\left\{h^* \mathbf{r}\right\} \right\} \\ &= \arg\max_{\mathbf{c}} \sum_{n=0}^{N+1} (\mathbf{1}-2\mathbf{c}_n)^T \Re\left\{h^* \mathbf{r}_n\right\}.\end{aligned} \tag{3.10}$$

Note that $\mathbf{c}$ depends on $\mathbf{x}$, since $\mathbf{c}_n$ contains x_n due to the convolutional encoder and thus, we can acquire $\hat{\mathbf{x}}_{\text{coh}}^{\text{ML}}$ from $\hat{\mathbf{c}}_{\text{coh}}^{\text{ML}}$. This is illustrated in Fig. 3.2. The above problem can be easily solved with linear complexity utilizing the Viterbi Algorithm. The trellis diagram at the receiver is shown in Fig. 3.2 [23].

3.3 Optimal Noncoherent Decoding

If the actual channel coefficient is not available at the receiver and is modelled as a random variable, then this variable appears in consecutive received vectors, hence the channel induces memory. As a result, optimal decoding requires processing of the entire sequence of $N+2$ received vectors.

The ML noncoherent decoder maximizes the conditional pdf of $\mathbf{r}$ given $\mathbf{c}$, that is, the optimal decision is given by

$$\hat{\mathbf{c}}^{\text{ML}} = \arg\max_{\mathbf{c}} f(\mathbf{r} \mid \mathbf{c}). \tag{3.11}$$

Note that $\|\mathbf{1}-2\mathbf{c}\|^2 = L$ for any $\mathbf{c}$. It can be shown that, for Rayleigh fading (i.e., $h \sim \mathcal{CN}(0, \sigma_h^2)$), the received symbol vector $\mathbf{r}$ given the transmitted coded sequence $\mathbf{c}$ follows a proper complex Gaussian distribution with mean $\mathbb{E}[\mathbf{r} \mid \mathbf{c}] = 0$ and covariance matrix

$$\mathrm{C}_{\mathbf{r}|\mathbf{c}} \triangleq \mathbb{E}[\mathbf{r}\mathbf{r}^H \mid \mathbf{c}] = 2E\sigma_h^2(\mathbf{1}-2\mathbf{c})(\mathbf{1}-2\mathbf{c})^T + \sigma_w^2 \mathbf{I}_L. \tag{3.12}$$

Hence, the optimization problem in (3.11) can be rewritten as

$$\hat{\mathbf{c}}^{\text{ML}} = \arg\max_{\mathbf{c}} \frac{1}{\pi^L |\mathrm{C}_{\mathbf{r}|\mathbf{c}}|} \mathrm{e}^{-\mathbf{r}^H \mathrm{C}_{\mathbf{r}|\mathbf{c}}^{-1} \mathbf{r}}. \tag{3.13}$$

Exploiting identities for the determinant and inverse of a rank-1 update [13] and the fact that $\|\mathbf{1}-2\mathbf{c}\|^2 = L$, we obtain

$$|\mathrm{C}_{\mathbf{r}|\mathbf{c}}| = |\sigma_w^2 \mathbf{I}_L| \left(1 + 2E\frac{\sigma_h^2}{\sigma_w^2}(\mathbf{1}-2\mathbf{c})^T(\mathbf{1}-2\mathbf{c})\right) = \sigma_w^{2L-2}\left(\sigma_w^2 + 2E\sigma_h^2\|\mathbf{1}-2\mathbf{c}\|^2\right) \tag{3.14}$$

and

$$\mathrm{C}_{\mathbf{r}|\mathbf{c}}^{-1} = \sigma_w^{-2}\mathbf{I}_L - \frac{2E\sigma_h^2}{\sigma_w^4 + 2E\sigma_h^2\sigma_w^2\|\mathbf{1}-2\mathbf{c}\|^2}(\mathbf{1}-2\mathbf{c})(\mathbf{1}-2\mathbf{c})^T. \tag{3.15}$$

If we substitute (3.14) and (3.15) in (3.13), then we obtain

$$\hat{\mathbf{c}}^{\mathrm{ML}} = \arg\max_{\mathbf{c}} |(\mathbf{1}-2\mathbf{c})^T\mathbf{r}|. \tag{3.16}$$

Substituting $\mathbf{c} = \left[\mathbf{c}_0^T, \mathbf{c}_1^T, \ldots, \mathbf{c}_{N+1}^T\right]^T$ in (3.16), the ML rule is rewritten in terms of the information sequence $\tilde{\mathbf{x}} = \begin{bmatrix}\mathbf{x}^T & 0 & 0\end{bmatrix}^T$ as

$$\hat{\tilde{\mathbf{x}}}^{\mathrm{ML}} = \arg\max_{\tilde{\mathbf{x}}\in\{0,1\}^{N+2}} |(\mathbf{1}-2\mathbf{c}_0)^T\mathbf{r}_0 + (\mathbf{1}-2\mathbf{c}_1)^T\mathbf{r}_1 + \ldots + (\mathbf{1}-2\mathbf{c}_{N+1})^T\mathbf{r}_{N+1}| \tag{3.17}$$

from which we can acquire the optimal information sequence $\mathbf{x}^{\mathrm{ML}}$ which is the N first elements of $\hat{\tilde{\mathbf{x}}}^{\mathrm{ML}}$. The equivalent expressions (3.16) and (3.17) represent the optimal decision rule for a Rayleigh-distributed channel.

If, on the other hand, the channel distribution is not Rayleigh or is unknown, then we may consider joint channel estimation and data decoding, i.e., GLRT sequence decoding according to which,

$$\begin{aligned}\hat{\mathbf{c}}^{\mathrm{GLRT}} &= \arg\min_{\mathbf{c}}\left\{\min_{h\in\mathbb{C}}\|\mathbf{r}-\sqrt{2E}h(\mathbf{1}-2\mathbf{c})\|^2\right\} = \arg\min_{\mathbf{c}}\left\|\mathbf{r} - \frac{(\mathbf{1}-2\mathbf{c})^T\mathbf{r}}{\|\mathbf{1}-2\mathbf{c}\|^2}(\mathbf{1}-2\mathbf{c})\right\| \\ &= \arg\max_{\mathbf{c}} |(\mathbf{1}-2\mathbf{c})^T\mathbf{r}|.\end{aligned} \tag{3.18}$$

Hence, the ML optimization problem in (3.16) and the GLRT optimization problem in (3.18) are equivalent. This equivalence between noncoherent ML and GLRT has also been demonstrated in the context of equal-energy signals and uniformly distributed over $[0, 2\pi)$ channel phase [7, 20]. A straightforward approach to solve (3.16) and (3.18) (or, equivalently, (3.17)) would be an exhaustive search among all 2^N information sequences $\mathbf{x} \in \{0,1\}^N$. However such a solver would be impractical even for moderate values of N, since its complexity is on the order of $\mathcal{O}(2^N)$, i.e., it grows exponentially with N. In the following we present an algorithm that solves the above problems with empirically low complexity.

3.3.1 Noncoherent decoding with empirically low complexity

To present the proposed algorithm, we begin by rewriting the optimal decoding rule in (3.17) as

$$\begin{aligned}\hat{\tilde{\mathbf{x}}}^{\mathrm{ML}} &= \max_{\tilde{\mathbf{x}}\in\{0,1\}^{N+2}} |(\mathbf{1}-2\mathbf{c}_0)^T\mathbf{r}_0 + (\mathbf{1}-2\mathbf{c}_1)^T\mathbf{r}_1 + \ldots + (\mathbf{1}-2\mathbf{c}_{N+1})^T\mathbf{r}_{N+1}| \\ &= \max_{\tilde{\mathbf{x}}\in\{0,1\}^{N+2}} \max_{\phi\in[0,2\pi)} \Re\left\{\mathrm{e}^{-j\phi}\left((\mathbf{1}-2\mathbf{c}_0)^T\mathbf{r}_0 + \ldots + (\mathbf{1}-2\mathbf{c}_{N+1})^T\mathbf{r}_{N+1}\right)\right\} \\ &= \max_{\phi\in[0,2\pi)} \max_{\tilde{\mathbf{x}}\in\{0,1\}^{N+2}} \Re\left\{\mathrm{e}^{-j\phi}\left((\mathbf{1}-2\mathbf{c}_0)^T\mathbf{r}_0 + \ldots + (\mathbf{1}-2\mathbf{c}_{N+1})^T\mathbf{r}_{N+1}\right)\right\}.\end{aligned} \tag{3.19}$$

That is, to find the optimal sequence $\tilde{\mathbf{x}}^{\mathrm{ML}}$ in (3.17), we may let ϕ scan the interval $[0, 2\pi)$ and, for each value of ϕ, collect the corresponding vector

$$\tilde{\mathbf{x}}(\phi) = \underset{\tilde{\mathbf{x}}\in\{0,1\}^{N+2}}{\arg\max}\ \Re\left\{\mathrm{e}^{-j\phi}\left((\mathbf{1}-2\mathbf{c}_0)^T\mathbf{r}_0 + \ldots + (\mathbf{1}-2\mathbf{c}_{N+1})^T\mathbf{r}_{N+1}\right)\right\} \tag{3.20}$$

that solves the innermost maximization in (3.19). Then,

$$\hat{\tilde{\mathbf{x}}}^{\mathrm{ML}} \in \mathcal{X} \triangleq \{\hat{\tilde{\mathbf{x}}}_1, \hat{\tilde{\mathbf{x}}}_2, \ldots, \hat{\tilde{\mathbf{x}}}_{|\mathcal{X}|}\} = \bigcup_{\phi\in[0,2\pi)} \left\{\hat{\tilde{\mathbf{x}}}(\phi)\right\}. \tag{3.21}$$

The size of the set in (3.21) is not fixed and varies according to the received data. Although the size of the set is not fixed, it is, in most cases small and can be built with significantly lower complexity as compared to $\mathcal{O}(2^N)$.

Let ϕ have a fixed value in $[0, 2\pi)$, say $\phi = 0$. Then, (3.20) can be solved efficiently utilizing the VA. Note that, since ϕ is fixed (we have, without loss of generality assumed that $\phi = 0$), 3.20 is equivalent to performing coherent decoding on $\tilde{\mathbf{x}}$, i.e., it is equivalent to (3.10). Due to (3.21), the decision sequence $\hat{\tilde{\mathbf{x}}}(\phi = 0)$ will be equal to $\hat{\tilde{\mathbf{x}}}_m$ for some $m \in \{1, 2, \ldots, |\mathcal{X}|\}$, where

$$\hat{\tilde{\mathbf{x}}}_m = \left[\hat{x}_0, \ldots, \hat{x}_{N-1}, 0, 0\right]^T \tag{3.22}$$

and $\hat{x}_i$, $i = 0, 1, \ldots, N-1$, is the decision on the ith information bit x_i of the sequence of N consecutive bits. Moreover, if we (without loss of generality) assume that $h = \mathrm{e}^{j\phi}$ in (3.10), then, at $\phi = 0$ we have two paths arriving at each node on the trellis diagram for $n \geq 2$. We choose an arbitrary node at an arbitrary time $n \geq 2$ on the trellis diagram. We define the winning (w) and the failing (f) coded bit vectors arriving at node state s at time n as

$$\mathbf{c}^{w,s,n} \triangleq \begin{bmatrix}\mathbf{c}_0^{w,s,n}\\ \vdots \\ \mathbf{c}_n^{w,s,n}\end{bmatrix} \quad , \quad \mathbf{c}^{f,s,n} \triangleq \begin{bmatrix}\mathbf{c}_0^{f,s,n}\\ \vdots \\ \mathbf{c}_n^{f,s,n}\end{bmatrix}, \tag{3.23}$$

respectively. The corresponding complex-valued metrics of each path arriving at the node state s at time n are defined as

$$W_s^n \triangleq \sum_{i=0}^{n}(\mathbf{1}-2\mathbf{c}_i^{w,s,n})^T\mathbf{r}_i \quad , \quad F_s^n \triangleq \sum_{i=0}^{n}(\mathbf{1}-2\mathbf{c}^{f,s,n})^T\mathbf{r}_i \tag{3.24}$$

The decision on the winning path arriving at node state s at time n changes only when

$$\begin{aligned}
&\sum_{i=0}^{n}\Re\left\{\mathrm{e}^{-j\phi}\left((\mathbf{1}-2\mathbf{c}_i^{w,s,n}) - (\mathbf{1}-2\mathbf{c}_i^{f,s,n})\right)^T\mathbf{r}_i\right\} = 0\\
&\Leftrightarrow \cos(\phi - \angle(W_s^n - F_s^n)) = 0\\
&\Leftrightarrow \phi = \underbrace{\pm\frac{\pi}{2} + \angle(W_s^n - F_s^n)}_{\phi_s^{n(1)},\phi_s^{n(2)}} \pmod{2\pi}
\end{aligned} \tag{3.25}$$

The complex-valued metrics W_s^n, F_s^n can be calculated on the fly during the Viterbi algorithm run. Hence, for $n = 2, 3, \ldots, N+1$ and for $s = 1, 2, 3, 4$, from (3.25) we collect $8N - 10$ distinct phases

$$\phi_1^{2^{(1)}}, \phi_1^{2^{(2)}}, \phi_2^{2^{(1)}}, \phi_2^{2^{(2)}}, \ldots, \phi_1^{N+1^{(1)}}, \phi_1^{N+1^{(2)}}. \tag{3.26}$$

The first decision change on a node will happen at $\phi_1 = \min\{\phi_1^{2^{(1)}}, \phi_1^{2^{(2)}}, \phi_2^{2^{(1)}}, \phi_2^{2^{(2)}}, \ldots, \phi_1^{N+1^{(1)}}, \phi_1^{N+1^{(2)}}\}$. If we perform coherent decoding at $\phi = \phi_1$ and we repeat the above process we will get a new set of $8N - 10$ distinct phases

$$\tilde{\phi}_1^{2^{(1)}}, \tilde{\phi}_1^{2^{(2)}}, \tilde{\phi}_2^{2^{(1)}}, \tilde{\phi}_2^{2^{(2)}}, \ldots, \tilde{\phi}_1^{N+1^{(1)}}, \tilde{\phi}_1^{N+1^{(2)}}.. \tag{3.27}$$

Thus, we can repeat the above process until we build the set in (3.21).

All the above lead to the following algorithm.

1. Set $\phi_{\text{cur}} = 0$.
2. For $h = 1$ in (3.10) acquire $\hat{\mathbf{c}}$ and for $s \in \{1, 2, 3, 4\}, n \in \{2, 3, \ldots, N+1\}$ calculate W_s^n, F_s^n according to 3.24.
3. Set best_value $= |W_1^{N+1}|$.
4. Set $\hat{\mathbf{x}}^{\text{ML}} = \hat{\mathbf{x}}$.
5. Repeat
   a) For all W_s^n, F_s^n calculate the distinct phases $\phi_1^{2^{(1)}}, \phi_1^{2^{(2)}}, \phi_2^{2^{(1)}}, \phi_2^{2^{(2)}}, \ldots, \phi_1^{N+1^{(1)}}, \phi_1^{N+1^{(2)}}$ and discard all ϕs less or equal to ϕ_{cur}. This step has complexity $\mathcal{O}(N)$.
   b) If all phases are discarded in (a) stop, else set ϕ_{cur} as the minimum of the remaining ϕs in (a) and save node state and time for which ϕ_{cur} was obtained. This step has constant complexity.
   c) For $h = \mathrm{e}^{j\phi_{\text{cur}}}$ in (3.10) acquire $\hat{\mathbf{c}}$ and for $s \in \{1, 2, 3, 4\}, n \in \{2, 3, \ldots, N+1\}$ calculate W_s^n, F_s^n according to (3.24). To avoid ambiguity on the node for which ϕ_{cur} was obtained the decision on this node should not be the same as in the previous iteration. This step has complexity $\mathcal{O}(N)$.
   d) If $|W_s^n| >$ best_value
      i. Set best_value $= |W_s^n|$.
      ii. set $\hat{\mathbf{x}}^{\text{ML}} = \hat{\mathbf{x}}$.

      This step has complexity $\mathcal{O}(N)$.

The pseudo-code of the proposed ML/GLRT noncoherent trellis/convolutional decoding algorithm is illustrated in Fig. 3.3.

Algorithm 2 Optimal Noncoherent Trellis/Convolutional Decoding

Input: r

$\phi_{\text{cur}} = 0$.
Acquire $\hat{\mathbf{c}}, \hat{\mathbf{x}}$ for $h = 1$ according to (3.10) and calculate on the fly W_s^n,F_s^n for $n \in \{2, 3, \ldots, N+1\}$ and $s \in \{1, 2, 3, 4\}$.
best_value $= |W_1^{N+1}|$.
$\hat{\mathbf{x}}^{\text{ML}} = \hat{\mathbf{x}}$.
while True **do**
    for $n = 2$ to $N+1$ **do**
        for $s = 1$ to 4 **do**
            $\left(\phi_s^{n^{(1)}}, \phi_s^{n^{(2)}}\right) = \frac{\pi}{2} \pm \underline{\angle W_s^n - F_s^n} \mod (2\pi)$.
        end for
    end for
    $\Phi = \{\phi_1^{2^{(1)}}, \phi_1^{2^{(2)}}, \phi_2^{2^{(1)}}, \phi_2^{2^{(2)}}, \ldots, \phi_1^{N+1^{(1)}}, \phi_1^{N+1^{(2)}}\}$
    $\hat{\Phi} = \{\phi \in \Phi \mid \phi > \phi_{\text{cur}}\}$
    if $\hat{\Phi}$ is empty **then**
        break
    else
        $[\phi_{\text{cur}}, s, n] = \min\{\hat{\Phi}\}$
    end if
    Acquire $\hat{\mathbf{c}}, \hat{\mathbf{x}}$ for $h = \mathrm{e}^{j\phi_{\text{cur}}}$ according to (3.10) and calculate on the fly W_s^n,F_s^n for $n \in \{2, 3, \ldots, N+1\}$ and $s \in \{1, 2, 3, 4\}$. To avoid ambiguity on the node for which ϕ_{cur} was obtained the decision on this node should not be the same as in the previous iteration.
    if $|W_1^{N+1}| >$ best_value **then**
        best_value $= |W_1^{N+1}|$.
        $\hat{\mathbf{x}}^{\text{ML}} = \hat{\mathbf{x}}$.
    end if
end while

Output: $\hat{\mathbf{x}}^{\text{ML}}$

Figure 3.3: ML/GLRT noncoherent trellis/convolutional decoding algorithm.

3.4 Simulation results

We consider BPSK transmissions of coded-bit vectors for fixed SNR through a Rayleigh flat-fading channel with $\sigma_h^2 = 1$. In Fig. 3.4, we plot the cost of the optimal noncoherent convolutional decoder as a function of the information sequence length N for $N = 1, 2, 4, 10, 20, 40, 60, 80, 100$. For 1000 runs of the algorithm per sequence length we plot the worst case line cost, the average case line cost and the best case line cost. We include the lines $11N^2$ and $13N^2$, as a reference. We observe that the cost of the proposed noncoherent convolutional decoding algorithm follows a complexity trend of $\mathcal{O}(N^2)$.

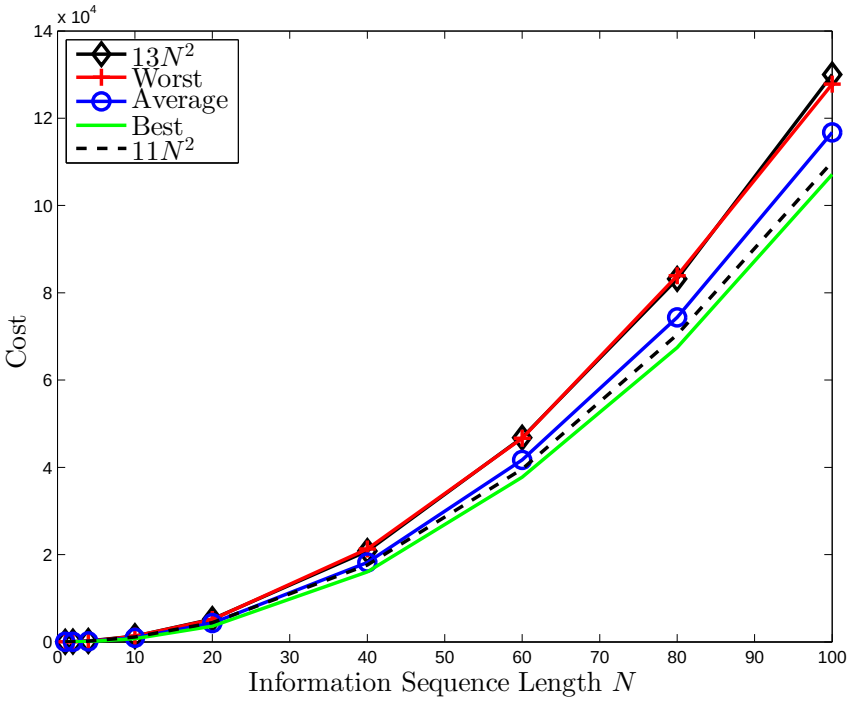

Figure 3.4: Optimal noncoherent convolutional decoding cost (order of number of operations) versus information sequence length N.

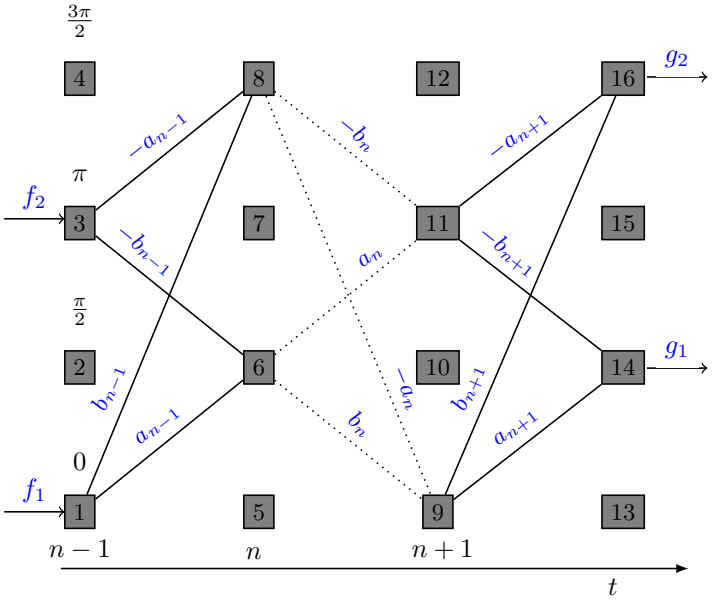

Figure 4.1: Portion of Trellis diagram at a receiver

CHAPTER 4

Appendix

4.1 Proof of Lemma 1

Proof. Fig. 4.1 represents a portion of a trellis diagram at a receiver after a random sequence has been transmitted, where we assume that n is odd and we define

$$\begin{aligned} a_n &\triangleq \begin{cases} \Re\{(h\mathbf{s}^1 \mathrm{e}^{j0})^H \mathbf{r}_n\}, & n = \text{odd} \\ \Re\{(h\mathbf{s}^1 \mathrm{e}^{j\frac{\pi}{2}})^H \mathbf{r}_n\}, & n = \text{even}, \end{cases} \\ b_n &\triangleq \begin{cases} \Re\{(h\mathbf{s}^{-1} \mathrm{e}^{j0})^H \mathbf{r}_n\}, & n = \text{odd} \\ \Re\{(h\mathbf{s}^{-1} \mathrm{e}^{j\frac{\pi}{2}})^H \mathbf{r}_n\}, & n = \text{even}. \end{cases} \end{aligned} \tag{4.1}$$

In addition we define $f_1(\mathbf{r}_0, \ldots, \mathbf{r}_{n-2}, \mathbf{s}^1, \mathbf{s}^{-1}, h) \in \mathbb{R}$ as the metric of the path priori to node 1, $f_2(\mathbf{r}_0, \ldots, \mathbf{r}_{n-2}, \mathbf{s}^1, \mathbf{s}^{-1}, h) \in \mathbb{R}$ as the metric of the path priori to node 3, $g_1(\mathbf{r}_{n+2}, \ldots, \mathbf{r}_{N-1}, \mathbf{s}^1, \mathbf{s}^{-1}, h) \in \mathbb{R}$ as the metric of the path posteriori to node 14, and, $g_2(\mathbf{r}_{n+2}, \ldots, \mathbf{r}_{N-1}, \mathbf{s}^1, \mathbf{s}^{-1}, h) \in \mathbb{R}$ as the metric of the path posteriori to node 16. We choose an arbitrary winning path. The winning path passes through node 1 and node 16. Somehow the nodes of the path at times $n, n+1$ are lost and we need to retrieve

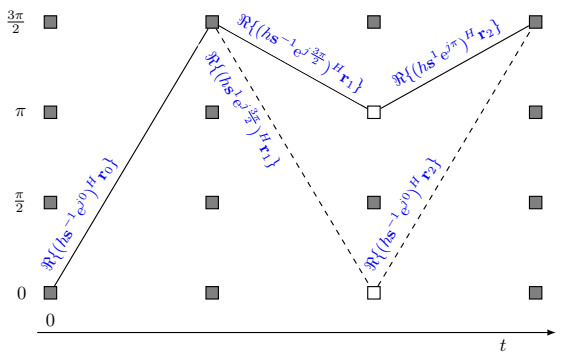

Figure 4.2: Potential winning path #1.

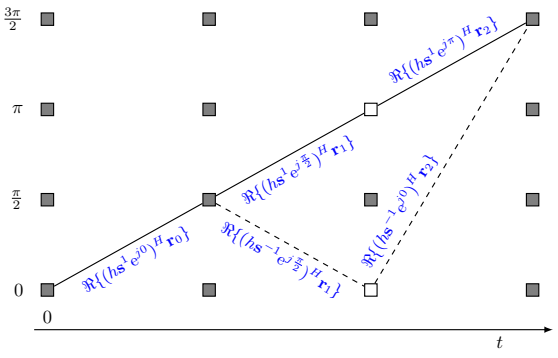

Figure 4.3: Potential winning path #2.

them. Based on the knowledge available there are four potential paths from which we must choose. Path $\{1, 6, 11, 16\}$, path $\{1, 6, 9, 16\}$, path $\{1, 8, 11, 16\}$ and path $\{1, 8, 9, 16\}$. The metrics of all four potential paths have the factors f_1, g_2 in common and hence, the factors f_1, g_2 do not affect the decision we want to make. Ultimately, we conclude that in order to optimally coherently decide for symbol n, we can either perform sequence detection, or decide based on the received vectors $\mathbf{r}_{n-1}, \mathbf{r}_n, \mathbf{r}_{n+1}$ of the time window $n-1, n, n+1$. The proof is concluded. □

4.2 Proof of Lemma 2

Proof. Using Fig. 4.1 as a reference, we arbitrary consider the path of nodes $1, 6, 11, 16$ as a portion of the winning path of the sequence as a whole. By this knowledge we know that $f_1 + a_{n-1} > f_2 - b_{n-1}$. The last inequality is equivalent to $f_1 + b_{n-1} > f_2 - a_{n-1}$, or in others words, the most likely path ending up to node 8, is also passing through node 1. Ultimately, we infer that at any given moment the two most likely paths originate from the same node. The proof is concluded. □

4.3 Proof of Lemma 3

Proof. The paths (solid lines) in Fig. 4.2 and Fig. 4.3 represent two potential winning paths. If we desire to find the ϕ's for which the decision changes on the node which is filled with white, based on Fig. 4.2, then we may calculate the 2 ϕ's according to (2.25) as

$$\begin{aligned}(\phi_1, \phi_2) &= \pm\frac{\pi}{2} + \angle\underline{(\mathbf{s}^{-1}\mathrm{e}^{j\frac{3\pi}{2}})^H\mathbf{r}_1 + (\mathbf{s}^{1}\mathrm{e}^{j\pi})^H\mathbf{r}_2 - (\mathbf{s}^{1}\mathrm{e}^{j\frac{3\pi}{2}})^H\mathbf{r}_1 - (\mathbf{s}^{-1}\mathrm{e}^{j0})^H\mathbf{r}_2} \\ &= \pm\frac{\pi}{2} + \angle\underline{j(\mathbf{s}^{-1})^H\mathbf{r}_1 - (\mathbf{s}^{1})^H\mathbf{r}_2 - j(\mathbf{s}^{1})^H\mathbf{r}_1 - (\mathbf{s}^{-1})^H\mathbf{r}_2} \\ &= \pm\frac{\pi}{2} + \angle\underline{j\left(\mathbf{s}^{-1} - \mathbf{s}^{1}\right)^H\mathbf{r}_1 - \left(\mathbf{s}^{1} + \mathbf{s}^{-1}\right)^H\mathbf{r}_2.}\end{aligned} \tag{4.2}$$

If on the other hand, we desire to find the ϕ's for which the decision changes on the node which is filled with white, based on Fig. 4.3, then we may calculate the 2 ϕ's according to (2.25) as

$$\begin{aligned}(\phi_1, \phi_2) &= \pm\frac{\pi}{2} + \angle\underline{(\mathbf{s}^{1}\mathrm{e}^{j\frac{\pi}{2}})^H\mathbf{r}_1 + (\mathbf{s}^{1}\mathrm{e}^{j\pi})^H\mathbf{r}_2 - (\mathbf{s}^{-1}\mathrm{e}^{j\frac{\pi}{2}})^H\mathbf{r}_1 - (\mathbf{s}^{-1}\mathrm{e}^{j0})^H\mathbf{r}_2} \\ &= \pm\frac{\pi}{2} + \angle\underline{-j(\mathbf{s}^{1})^H\mathbf{r}_1 - (\mathbf{s}^{1})^H\mathbf{r}_2 + j(\mathbf{s}^{-1})^H\mathbf{r}_1 - (\mathbf{s}^{-1})^H\mathbf{r}_2} \\ &= \pm\frac{\pi}{2} + \angle\underline{j\left(\mathbf{s}^{-1} - \mathbf{s}^{1}\right)^H\mathbf{r}_1 - \left(\mathbf{s}^{1} + \mathbf{s}^{-1}\right)^H\mathbf{r}_2.}\end{aligned} \tag{4.3}$$

We then observe that the 2 ϕ's produced by (4.2), are identical to those produced by (4.3). In a similar manner it can be shown that, no matter which is the winning path passing through a node, the ϕ's for which the decision changes on that node will always be the same. The proof is concluded.

□

Bibliography

[1] P. Kimuli, “Introduction to GSM and GSM mobile RF transceiver derivation," RF Design, Jun. 2003.

[2] W. Osborne and M. Luntz, “Coherent and noncoherent detection of CPFSK," *IEEE Trans. Commun.*, vol. 22, no. 8, pp. 1023–1036, Aug. 1974.

[3] H. Leib and S. Pasupathy, “Block coded noncoherent MSK," in *Proc. IEEE SUPERCOMM/ICC 1990*, Atlanta, GA, Apr. 1990, vol. 2, pp. 448–452.

[4] M. K. Simon and D. Divsalar, “Maximum-likelihood block detection of noncoherent continuous phase modulation," *IEEE Trans. Commun.*, vol. 41, no. 1, pp. 90–98, Jan. 1993.

[5] G. Colavolpe and R. Raheli, “Noncoherent sequence detection of continuous phase modulations," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1303–1307, Sep. 1999.

[6] M. Morelli, D. Benfatto, D. Lucano, M. Luise, and U. Mengali, “Simple non-coherent detectors for CPM signals transmitted over Rayleigh flat-fading channels," in *Proc. IEEE SPAWC 2003*, Rome, Italy, Jun. 2003, pp. 492–496.

[7] D. Warrier and U. Madhow, “Spectrally efficient noncoherent communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651–668, Mar. 2002.

[8] 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, no. 10, pp. 1676–1689, Oct. 2003.

[9] 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.

[10] 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.

[11] D. Divsalar and M. K. Simon, “Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[12] H. Leib and S. Pasupathy, “Noncoherent block demodulation of PSK," in *Proc. IEEE VTC 1990*, Orlando, FL, May 1990, pp. 407–411.

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

[14] K. Mackenthun, “A fast algorithm for maximum likelihood detection of QPSK or $\pi/4$-QPSK sequences with unknown phase," in *Proc. IEEE PIMRC 1992*, Boston, MA, Oct. 1992, pp. 240–244.

[15] K. M. Mackenthun, “A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1471–1474, Feb./Mar./Apr. 1994.

[16] G. N. Karystinos and D. A. Pados, “Rank-2-optimal adaptive design of binary spreading codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3075–3080, Sep. 2007.

[17] D. S. Papailiopoulos, G. A. Elkheir, and G. N. Karystinos, “Maximum-likelihood noncoherent PAM detection," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1152–1159, Mar. 2013.

[18] P. N. Alevizos, Y. Fountzoulas, G. N. Karystinos, and A. Bletsas, “Log-linear-complexity GLRT-optimal noncoherent sequence detection for orthogonal and RFID-oriented modulations," *IEEE Trans. Commun.*, to appear.

[19] I. Motedayen-Aval, A. Krishnamoorthy, and A. Anastasopoulos, “Optimal joint detection/estimation in fading channels with polynomial complexity," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 209–223, Jan. 2007.

[20] 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, no. 2, pp. 197–207, Feb. 2003.

[21] D. J. Ryan, I. B. Collings, and I. V. L. Clarkson, “GLRT-optimal noncoherent lattice decoding," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3773–3786, Jul. 2007.

[22] R. G. McKilliam, D. J. Ryan, I. Vaughan L. Clarkson, and I. B. Collings, “An improved algorithm for optimal noncoherent QAM detection," in *Proc. AusCTW 2008*, Christchurch, New Zealand, Feb. 2008, pp. 64–68.

[23] J. G. Proakis and M. Salehi, *Digital Communications*, 5th Ed. Upper Saddle River, NJ: Prentice-Hall, Nov. 2007.

[24] A. Svensson and C. E. Sundberg, “Optimum MSK-type receivers for CPM on Gaussian and Rayleigh fading channels," *IEE Proc., Pt. F, Commun., Radar, Sig. Process.*, vol. 131, no. 5, pp. 480–490, Aug. 1984.