# Low-Complexity Detection for Faster-than-Nyquist Signaling based on Probabilistic Data Association

Michel Kulhandjian, *Member, IEEE,* Ebrahim Bedeer, *Member, IEEE,* Hovannes Kulhandjian, *Member, IEEE,* Claude D'Amours, *Member, IEEE,* and Halim Yanikomeroglu, *Fellow, IEEE*

*Abstract*—In this paper, we investigate the sequence estimation problem of faster-than-Nyquist (FTN) signaling as a promising approach for increasing spectral efficiency (SE) in future communication systems. In doing so, we exploit the concept of Gaussian separability and propose two probabilistic data association (PDA) algorithms with polynomial time complexity to detect binary phase-shift keying (BPSK) FTN signaling. Simulation results show that the proposed PDA algorithm outperforms the recently proposed SSSSE and SSSgb$K$SE algorithms for all SE values with a modest increase in complexity. The PDA algorithm approaches the performance of the semidefinite relaxation (SDRSE) algorithm for SE values of $0.96$ bits/sec/Hz, and it is within the $0.5$ dB signal-to-noise ratio (SNR) penalty at SE values of $1.10$ bits/sec/Hz for the fixed values of $\beta = 0.3$.

*Index Terms*—Faster-than-Nyquist (FTN) signaling, intersymbol interference (ISI), Mazo limit, sequence estimation.

## I. INTRODUCTION

IMPROVING spectral efficiency (SE) is one of the key requirements of future communication systems. Faster-than-Nyquist (FTN) signaling is a promising physical layer transmission technique for increasing SE of future communication systems compared to conventional Nyquist signaling [1]. The basic idea of FTN signaling is to transmit the time domain pulses with a rate that exceeds the Nyquist limit, which intentionally creates inter-symbol interference (ISI) at the receiver side. Interestingly, J. Mazo in 1975 showed that transmitting time domain pulses every $0.802\,T$, where $T$ is the symbol duration, results in the same minimum Euclidean distance (and, hence, the same error probability) as conventional Nyquist signaling [2]. This is equivalent to a $25\%$ increase in SE for the same energy per symbol and transmission bandwidth. The concept of FTN signaling has been extended to different domains, e.g., frequency-domain [3], non-binary modulation [4], other pulse shapes [5], to name a few. Recent research results extended the concept of FTN signaling to different domains and suggested that it can be a promising candidates for high capacity peer-to-peer links.

Given that optimal detection of FTN signaling is shown to be NP-hard [6], several linear/nonlinear detection techniques have been developed to remove ISI with different degrees of computational complexity [1]. For high spectral efficiency

M. Kulhandjian and C. D'Amours are with the School of Electrical Engineering, & Computer Science, University of Ottawa, Ottawa, Canada, e-mail: mkk6@buffalo.edu, cdamours@uottawa.ca.

E. Bedeer is with the Department of Electrical & Computer Engineering, University of Saskatchewan, Saskatoon, Canada, e-mail: e.bedeer@usask.ca.

H. Kulhandjian is with the Department of Electrical & Computer Engineering, California State University, Fresno, U.S.A., e-mail: hkulhandjian@csufresno.edu.

H. Yanikomeroglu is with the Department of Systems & Computer Engineering, Carleton University, Ottawa, Canada, e-mail: halim@sce.carleton.ca.

systems, i.e., severe ISI, M-Bahl, Cocke, Jelinek, and Raviv (M-BCJR) algorithm [7] and the truncated Viterbi algorithm [8] are examples of nonlinear detectors that have shown SE gains; however, this has included exponential complexity in ISI length. It is worth mentioning that iterative detection of FTN signaling was shown to be effective in approaching the zero-ISI performance at high signal-to-noise ratio (SNR) values in [1], [9]. To achieve moderate SE gains with reasonable detection complexity, nonlinear FTN signaling detection algorithms based on semi-definite relaxation were proposed in [10] for high-order modulations with polynomial time complexity. A low-complexity nonlinear frequency domain equalizer was proposed in [11]; however, it is found that the insertion of the cyclic prefix could result in performance loss for short block transmissions. The authors in [12] proposed a linear algorithm that exploits the structure of ISI at the receiver and detects FTN signaling on a symbol-by-symbol basis. It was shown in [12] that the proposed algorithm is suitable for low ISI scenarios.

In this paper, we exploit the concept of Gaussian separability to propose relatively low-complexity probabilistic data association (PDA) algorithms to detect binary phase-shift keying (BPSK)[1] FTN signaling at polynomial time complexity. The proposed algorithms are iterative, i.e., at each iteration a given FTN signaling symbol is detected to improve the overall error performance. Simulation results show the effectiveness of the proposed algorithms and their merit over other competing schemes in the literature. The proposed algorithms outperform the successive symbol-by-symbol sequence estimation (SSSSE) and successive symbol-by-symbol with go-back-$K$ sequence estimation (SSSgb$K$SE) algorithms in [12] for all SE values. Additionally, the proposed algorithms approach the performance of the semidefinite relaxation (SDRSE) algorithm proposed in [10] for SE values of 0.96 bits/sec/Hz, and they are within 0.5 dB SNR penalty at SE values of 1.10 bits/sec/Hz.

The remainder of this paper is organized as follows. Section II presents the FTN signaling system model and the linear separability concept. The concept of Gaussian separability and the proposed PDA algorithms are presented in Section III, while the simulation results are discussed in Section IV. Finally, the paper is concluded in Section V.

The following notations are used in this paper. All boldface lower case letters indicate column vectors, and upper case letters indicate matrices. Additionally, $(\cdot)^H$ denotes Hermitian operation, $\mathsf{sgn}(\cdot)$ denotes the sign function, $|\cdot|$ is the scalar magnitude, $\mathsf{var}[\cdot]$ is variance, $\mathsf{cov}[\cdot]$ denotes covariance, $\mathsf{diag}[\cdot]$ denotes a matrix with diagonal elements, and $\mathbb{E}\{\cdot\}$

---

[1]The proposed algorithms can be straightforwardly extended to quadrature-phase-shift keying (QPSK).

denotes the expected value.

## II. FTN SIGNALING SYSTEM MODEL

At the FTN signaling transmitter, the BPSK data symbols are shaped by a unit-energy pulse $p(t)$. We transmit a total of $N$ symbols at every $\tau T$, where $\tau \in (0,1]$ is the time packing parameter. The received signal, after being affected by an additive white Gaussian noise (AWGN) channel, passes through a filter matching the transmit pulse and is given as

$$y(t) = \sqrt{\tau E_s} \sum_{n=1}^{N} a_n g(t - n\tau T) + q(t), \qquad (1)$$

where $a_n, n = 1, \ldots, N$, is the BPSK data symbol, $E_s$ is the data symbol energy, $g(t) = \int p(x)p(x-t)dx$, $q(t) = \int n(x)p(x-t)dx$, where $n(t)$ is the AWGN with zero mean and variance $\sigma^2$, and $1/(\tau T)$ is the signaling rate. The output signal of the matched-filter is sampled at every $\tau T$ and can be expressed as

$$
\begin{aligned}
y_k &= \sqrt{\tau E_s} \sum_{n=1}^{N} a_n g(k\tau T - n\tau T) + q(k\tau T), \\
&= \underbrace{\sqrt{\tau E_s}\, a_k\, g(0)}_{\text{desired symbol}} \qquad\qquad\qquad (2) \\
&\quad + \underbrace{\sqrt{\tau E_s} \sum_{n=1,\, n\neq k}^{N} a_n\, g((k-n)\tau T)}_{\text{ISI}} + q(k\tau T).
\end{aligned}
$$

We rewrite (2) in a vector form as

$$\mathbf{y}_c = \sqrt{\tau E_s}\, \mathbf{a} * \mathbf{g} + \mathbf{q}_c, \qquad (3)$$

where $\mathbf{a}$, $\mathbf{g}$, $\mathbf{q}_c$, and $*$ are the transmit data symbol vector, ISI vector, colored noise vector, and convolution operator, respectively. Since the noise samples are correlated, we design a whitening matched filter by applying spectral factorization to the z-transform of $\mathbf{g}$ to obtain the approximate whitening filter coefficients. For more details on the design of such filters for FTN receivers, we refer the reader to [8], [10]. Hence, after passing (3) through the whitening filter, we have

$$\mathbf{y} = \sqrt{\tau E_s}\, \mathbf{a} * \mathbf{v} + \mathbf{n}_w, \qquad (4)$$

where $\mathbf{n}_w$ is AWGN $\sim \mathcal{N}(0, \sigma^2)$ and $\mathbf{v}$ is the causal ISI vector and is constructed from $\mathbf{g}$ as $\mathbf{v}[n] * \mathbf{v}[-n] = \mathbf{g}$. Equation (4) can be rewritten as

$$\mathbf{y} = \mathbf{G}\mathbf{a} + \mathbf{n}_w, \qquad (5)$$

where $\mathbf{G}$ is an $N \times N$ Toeplitz Gram matrix that represents the ISI. The maximum likelihood sequence estimation (MLSE) problem for detecting the FTN signaling in (5) is formally expressed as

$$\widehat{a} = \underset{\mathbf{a} \in \{\pm 1\}^{N \times 1}}{\arg\min}\ \mathbf{a}^H \mathbf{G} \mathbf{a} - 2\mathbf{y}^H \mathbf{a}. \qquad (6)$$

It is known that obtaining the MLSE solution is generally NP-hard [6], which may be prohibitively complex for use in practical detectors. That being said, finding low-complexity optimal detectors of FTN signaling is important for improving SE in future communication systems.

In this work, we aim to identify what properties the ISI matrix $\mathbf{G}$ needs to have such that the sequence estimation can achieve asymptotically zero probability of error detection. In [12], the authors exploit the idea of linear separability to

identify an operating region where perfect reconstruction of FTN signaling is guaranteed for noise-free transmission. It was found that such a linear separability operating region depends on the raised cosine pulse shape, its roll-off factor $\beta$, and the FTN signaling time acceleration parameter $\tau$. The perfect reconstruction conditions of the BPSK FTN signaling in [12] can be summarized in the following lemma.

*Lemma 1:* [12] We assume perfect estimation conditions for BPSK FTN signaling over the noise-free transmission. Regardless of the value of the current data symbol $a_k$, the upcoming $L - 1$ data symbols $a_{k+1}, \ldots, a_{k+L-1}$, and the value of $L$, the following inequality holds for a certain range of $\tau$ and $\beta$:

$$|G_{1,1} a_k| > |G_{1,2} a_{k+1} + \ldots + G_{1,L} a_{k+L-1}|, \qquad (7)$$

where $G_{i,j}$ is the $i$-th row and $j$-th column value of the ISI matrix $\mathbf{G}$. Following (7), in this work, we define the *conditionally linear margin* as

$$\delta_k = |G_{1,1} a_k| - |G_{1,2} a_{k+1} + \ldots + G_{1,L} a_{k+L-1}|, \qquad (8)$$

which can be seen that $\delta_k > 0$ is equivalent to *Lemma* 1 and it is the sufficient and necessary condition for linear separability. Therefore, the linear margin, $\delta_k$, represents a distance measure for the separability condition of *Lemma* 1 to hold. Obviously, the greater the linear margin $\delta_k$, the better the error performance.

## III. PROPOSED DETECTION ALGORITHMS

In this section, low-complexity iterative detectors are proposed to detect FTN signaling on the basis of Gaussian separability [13]. We show that Gaussian separability has a greater classification margin compared to the linear separability in [12], which thereby improves the detection error performance of FTN signaling. To illustrate the idea, let us define a subset $\mathcal{F} \subset \mathcal{U} = \{1, 2, \ldots, N\}$ of the symbols that are already detected, i.e., $\{a_l = a_l^*, l \in \mathcal{F}\}$. To detect the $k$-th symbol $k \in \mathcal{U} - \mathcal{F}$, we should be able to separate the set of undetected symbols in $\mathcal{U} - \mathcal{F}$ in two groups corresponding to $a_k = 1$ and $a_k = -1$, respectively. Following this notation, the observation vector $\mathbf{y}$ in (5) is rewritten as

$$\mathbf{y} = \mathbf{g}_k a_k + \boldsymbol{\mu}_{\mathcal{F}}^* + \boldsymbol{\mu}_{\mathcal{F}_k} + \mathbf{n}_w, \qquad (9)$$

where $\mathbf{g}_k$ is the $k$-th column of ISI matrix $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_N]$, $\boldsymbol{\mu}_{\mathcal{F}}^* = \sum_{l \in \mathcal{F}} \mathbf{g}_l a_l^*$ and $\boldsymbol{\mu}_{\mathcal{F}_k} = \sum_{j \in \mathcal{U} - \mathcal{F} - k} \mathbf{g}_j a_j$. Linear detectors can be viewed as hyperplanes on the subspace spanned by $\mathbf{g}_k$ for $1 \leq k \leq N$ columns of matrix $\mathbf{G}$. A linear detector that achieves asymptotic negligible error probability as $\sigma^2 \to 0$ is said to be asymptotically efficient. Mathematically, it can be expressed as, $\forall a_j \in \{\pm 1\}$, $j \in \mathcal{U} - \mathcal{F} - k$,

$$\mathbf{c}^H \mathbf{y} \geq 0 \quad \text{if } a_k = 1$$
$$\mathbf{c}^H \mathbf{y} < 0 \quad \text{if } a_k = -1.$$

From this definition of linear separability, the following theorem can be proved.

*Theorem 3.1:* [14] The $k$-th symbol is asymptotically linearly separable conditionally on $\mathcal{F}$ if for $\sigma^2 \to 0$, an $N$-dimensional linear filter $\mathbf{c}$,

$$|\mathbf{c}^H \mathbf{g}_k| > \sum_{j \in \mathcal{U} - \mathcal{F} - k} |\mathbf{c}^H \mathbf{g}_j|. \qquad (10)$$

It can be shown that Lemma 1 is a special case of Theorem 3.1 when the linear filter $\mathbf{c} = \mathbf{e}_k$ in (10), where $\mathbf{e}_k$ is a vector with value 1 at position $k$ and 0 in all other positions under the constraint that data symbols $a_k \in \{\pm 1\}$, $\forall k$.

TABLE I: Symbol-by-symbol Separability ($\beta = 0.3$)

| $\tau$ | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| $\delta_{max}, \delta_{ave}$ | $0.97, 0.11$ | $0.97, 0.20$ | $0.97, 0.38$ | $0.97, 0.53$ |

TABLE II: Gaussian Separability ($\beta = 0.3$)

| SNR (dB) | $\tau = 0.6$ | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| 0 | 0.00, 0.00 | 1.16, 0.49 | 1.08, 0.58 | 1.02, 0.73 |
| 2 | 2.01, 0.68 | 1.84, 0.82 | 1.71, 0.96 | 1.62, 1.21 |
| 4 | 3.19, 1.12 | 2.92, 1.35 | 2.71, 1.57 | 2.56, 2.00 |
| 6 | 5.06, 1.73 | 4.62, 2.16 | 4.29, 2.55 | 4.06, 3.30 |
| 8 | 8.02, 2.71 | 7.32, 3.45 | 6.80, 4.13 | 6.43, 5.42 |

Recall that if the condition densities for all the symbols $\mathbf{f}(\mathbf{y}|a_j)$ are Gaussian then the optimal filter for user-$k$ is a scaled version of e.g., $\mathbf{c}_k \triangleq \mathbf{R}_{\mathcal{F}_k}^{-1} \mathbf{g}_k$. The solution of the optimal separating hyperplane in a linear detector (classifier), despite the simple formulation, may be very difficult when the classes do not follow a Gaussian distribution. If we assume that the ISI is Gaussian, then the linear classifier becomes the optimal maximum likelihood (ML) decision with hyperplanes dividing the $N$-dimensional space.

*Theorem 3.2:* [14] The $k$-th symbol is a Gaussian separable conditionally on $\mathcal{F}$ if for $\sigma^2 \to 0$,

$$\mathbf{g}_k^H \mathbf{R}_{\mathcal{F}_k}^{-1} \mathbf{g}_k > \sum_{j \in \mathcal{U} - \mathcal{F} - k} |\mathbf{g}_k^H \mathbf{R}_{\mathcal{F}_k}^{-1} \mathbf{g}_j|, \quad (11)$$

where

$$\mathbf{R}_{\mathcal{F}_k} = \sum_{j \in \mathcal{U} - \mathcal{F} - k} \mathbf{g}_j \mathbf{g}_j^H + \sigma^2 \mathbf{I}_N, \quad (12)$$

and the parameter

$$\Delta_{\mathcal{F}_k} = \mathbf{g}_k^H \mathbf{R}_{\mathcal{F}_k}^{-1} \mathbf{g}_k - \sum_{j \in \mathcal{U} - \mathcal{F} - k} |\mathbf{g}_k^H \mathbf{R}_{\mathcal{F}_k}^{-1} \mathbf{g}_j| \quad (13)$$

is called *conditionally Gaussian margin*. Unlike the linear margin $\delta_k$, the Gaussian margin $\Delta_{\mathcal{F}_k}$ value depends on the noise variance $\sigma^2$. We measured conditionally margins, $\delta_k$ and $\Delta_{\mathcal{F}_k}$, in the operating region, where ISI matrix $\mathbf{G}$ is formed with roll-off factor $\beta = 0.3$ and $\tau \in \{0.6, 0.7, 0.8, 0.9\}$. At each pair of $\beta, \tau$, we generate matrix $\mathbf{G}$ and then compute $\delta_k$ and $\Delta_{\mathcal{F}_k}$ for $1 \le k \le N$ using equations (8) and (13)[2], respectively. We evaluate their maximum and average values as $\delta_{max} = \text{argmax}_k \delta_k$, $\delta_{ave} = \frac{1}{N} \sum_{k=1}^{N} \delta_k$, $\Delta_{max} = \text{argmax}_k \Delta_{\mathcal{F}_k}$, $\Delta_{ave} = \frac{1}{N} \sum_{k=1}^{N} \Delta_{\mathcal{F}_k}$.

Since the linear margin $\delta_k$ does not depend on the noise variance, $\sigma^2$, there is only one row with values ($\delta_{max}$, $\delta_{ave}$) as shown in Table I. On the other hand, Table II shows the ($\Delta_{max}$, $\Delta_{ave}$) for different SNR values at each row, where we vary the noise variance, $\sigma^2$, and keep the matrix $\mathbf{G}$ unchanged. We observe from Table I and Table II that the

[2]We note that $\Delta_{\mathcal{F}_k}$ is computed using the following set $\mathcal{F} = \{i_1, i_2, \ldots, i_{k-1}\}$ for $1 \le k \le N$. The order of $\{i_1, i_2, \ldots, i_{k-1}\}$ depends on the ISI matrix $\mathbf{G}$ and $i_t$'s are precisely computed in our PDA algorithm at line 6.

Gaussian classification margins have greater maximum and average values than their linear counterparts. Therefore, we can theorize that using a Gaussian classification margin can potentially improve the performance in terms of bit error rate (BER).

The main idea of the PDA detection algorithm is based on conditionally Gaussian separability criterion and the fact that the Gaussian margin $\Delta_{\mathcal{F}_k}$ is greater than the linear margin $\delta_k$, as shown in Tables I and II. One can re-express (9) as

$$\mathbf{y} = \mathbf{g}_k a_k + \mathbf{w}_k, \quad (14)$$

where $\mathbf{w}_k = \mathbf{G}_k \mathbf{a}_k + \mathbf{n}_w$, $\mathbf{G}_k$ is the $N \times N - 1$ resulting matrix after excluding the $k$-th column of $\mathbf{G}$, and $\mathbf{a}_k$ is $N - 1 \times 1$ vector resulting after excluding $k$-th element of $\mathbf{a}$.

We associate the probabilities $P_a(k)$ and $1 - P_a(k)$ with the estimate of the symbols $a_k = 1$ and $a_k = -1$, respectively. Now, for an arbitrary symbol $a_k$, we treat the other symbols $a_j$, $\forall j \ne k$ as binary random variables and $\mathbf{w}_k$ as a Gaussian noise vector. Then the likelihood ratio for the $a_k$ symbol can be written as

$$\begin{aligned}
\Lambda(a_k) &= \frac{P\{a_k = 1|\mathbf{y}, \{P_a(j)\}_{j \ne k}\}}{P\{a_k = -1|\mathbf{y}, \{P_a(j)\}_{j \ne k}\}}, \\
&= \frac{f_\mathbf{y}(\mathbf{y}|a_k = 1|, \{P_a(j)\}_{j \ne k})}{f_\mathbf{y}(\mathbf{y}|a_k = -1, \{P_a(j)\}_{j \ne k})} \\
&= \frac{\mathcal{N}(\mathbf{y}|\mathbf{g}_k + \boldsymbol{\mu}_k, \mathbf{C}_k)}{\mathcal{N}(\mathbf{y}|\mathbf{g}_k - \boldsymbol{\mu}_k, \mathbf{C}_k)} \\
&= \exp\left[2(\mathbf{y} - \boldsymbol{\mu}_k)^H \mathbf{C}_k^{-1} \mathbf{g}_k\right], \quad (16)
\end{aligned}$$
(15)

where $P_a(k) = P\{a_k = 1|\mathbf{y}, \{P_a(j)\}_{j \ne k}\}$ and $f_\mathbf{y}$ is the multivariate Gaussian distribution conditioned on $\{\mathbf{y}, \{P_a(j)\}_{j \ne k}\}$. The covariance can be calculated as

$$\begin{aligned}
\mathbf{C}_k &= \text{cov}[\mathbf{w}_k] \\
&= \mathbf{G}_k \mathbb{E}[\mathbf{a}_k \mathbf{a}_k^H] \mathbf{G}_k^H - \mathbf{G}_k \mathbb{E}[\mathbf{a}_k] \mathbb{E}[\mathbf{a}_k^H] \mathbf{G}_k^H + \mathbf{R}_n \\
&= \mathbf{G}_k \mathbf{C}_{\mathbf{a}_k} \mathbf{G}_k^H + \mathbf{R}_n, \quad (17)
\end{aligned}$$

where $\text{var}[a_k] = 4P_a(k)(1 - P_a(k))$, $\mathbf{C}_{\mathbf{a}_k} = \text{cov}[\mathbf{a}_k] = \text{diag}(\text{var}[a_1], \cdots, \text{var}[a_{k-1}], \text{var}[a_{k+1}], \cdots, \text{var}[a_n])$ and $\mathbf{R}_n = \mathbb{E}[\mathbf{n}_w \mathbf{n}_w^H]$, are the covariance matrix of $\mathbf{a}_k$ and the autocorrelation of noise vector, respectively. The knowledge about the $a_k$ symbol is carried over in each iteration through a dynamic update of means and covariances as follows

$$\begin{aligned}
\boldsymbol{\mu}_k &= \mathbb{E}[\mathbf{w}_k] = \mathbb{E}[\mathbf{G}_k \mathbf{a}_k + \mathbf{n}_w] = \mathbf{G}_k \mathbb{E}[\mathbf{a}_k] \\
&= \mathbf{G}_k \begin{bmatrix} 2P_a(1) - 1 \\ \vdots \\ 2P_a(k-1) - 1 \\ 2P_a(k+1) - 1 \\ \vdots \\ 2P_a(N) - 1 \end{bmatrix} \\
&= \mathbf{G}_k \mathbf{p}_{a,k}, \quad (18)
\end{aligned}$$

where $\mathbf{p}_{a,k}$ is the estimated $\mathbf{a}$ signal vector without the $k$-th signal. As we can see, the likelihood ratio in (15) is in the form of $P_a(k)/(1 - P_a(k)) = \exp(x)$. Hence, the posteriori probability $P_a(k)$ can be calculated as

$$P_a(k) = 1/\left(1 + \exp\left[-2(\mathbf{y} - \boldsymbol{\mu}_k)^H \mathbf{C}_k^{-1} \mathbf{g}_k\right]\right). \quad (19)$$

---

**Proposed PDA Algorithm**

---

**Input:** $\mathbf{y}$, $\mathcal{F} \leftarrow \{1, 2, \ldots, N\}$, $M$
1: Initialize : $P_a(k) = 1/2$, $k = 1, \ldots, N$
2: Compute $D_k$, $k = 1, \cdots, N$
3: **for** $i = 1$ to $M$
4:     $\mathcal{U} \leftarrow \mathcal{F}$
5:     **while** ($\mathcal{U} \neq \emptyset$)
6:         Update $P_a(k_m)$, $k_m = \text{argmax}_{k \in \mathcal{U}}(D_k)$
7:         Subtract $k_m$ from $\mathcal{U}$
8:         Recompute $D_k$, $\forall k \in \mathcal{U}$
**Output:** $\hat{\mathbf{a}} = \text{sgn}(\mathbf{p}_a)$

---

That being said, the proposed PDA algorithm for detecting FTN signaling can be summarized in the next page, where $M$ is the number of iterations and $D_k = \mathbf{g}_k^H \mathbf{C}_k^{-1} \mathbf{g}_k$.

*A. Modified PDA algorithm*

The complexity of the PDA algorithm can be further reduced if the detected data symbols whose probabilities lie within a confidence interval are removed in each iteration. In other words, if the probability of the symbol $k$, $P_a(k)$ lies in confidence region $\epsilon$, if $|P_a(k) - d| < \epsilon$, where $d = \{0, 1\}$. All the symbols that satisfy the confidence interval define the set $C$. We assign $P_a(k) = \{d : |P_a(k) - d| < \epsilon\}$ for $\forall k \in C$. Next, we remove symbols $C$ from $\mathcal{F}$, $\mathcal{F} = \mathcal{F} - C$. The modified PDA algorithm is presented next.

---

**Modified PDA Algorithm**

---

**Input:** $\mathbf{y} = \mathbf{Ga} + \mathbf{n}$, $\mathcal{F} \leftarrow \{1, 2, \ldots, N\}$, $M$
1: Initialize $P_a(k) = 1/2$, $k = 1, \cdots, N$
2: Compute $D_k$, $k = 1, \ldots, N$
3: **for** $i = 1$ to $M$
4:     Find $k$ that satisfy $|P_a(k) - d| < \epsilon$
5:     **if** $C \neq \emptyset$
6:         $P_a(k) = d$, $\forall k \in C$
7:         $\mathcal{F} \leftarrow \mathcal{F} - C$
8:     $\mathcal{U} \leftarrow \mathcal{F}$
9:     **while** ($\mathcal{U} \neq \emptyset$)
10:         Update $P_a(k_m)$, $k_m = \text{argmax}_{k \in \mathcal{U}}(D_k)$
11:         Subtract $k_m$ from $\mathcal{U}$
12:         Recompute $D_k$, $\forall k \in \mathcal{U}$
**Output:** $\hat{\mathbf{a}} = \text{sgn}(\mathbf{p}_a)$

---

*B. Complexity analysis*

The main complexity of the proposed PDA algorithm lies in updating $\mathbf{C}_k^{-1}$ in (19). However, efficient numerical techniques such as the Durbin-Levinson algorithm can exploit the Toeplitz structure of $C_k$ and compute its inverse, i.e., $\mathbf{C}_k^{-1}$, with a complexity of $O(N^2)$ [15]. However, since $C_k$ is not always guaranteed to be Toeplitz, the calculations of $C_k^{-1}$ can be in the order of $O(N^3)$. Since evaluating (19) is expected to occur at most $N$ times and the algorithm has deterministic execution steps, the total average and worst case computational complexity of the proposed PDA algorithm can be in the order of $O(N^4)$. The PDA algorithm has the same polynomial complexity order (i.e., both complexities scale in polynomial time with the number of transmit symbols) as the SDRSE algorithm in [10] (please note that complexity measure based
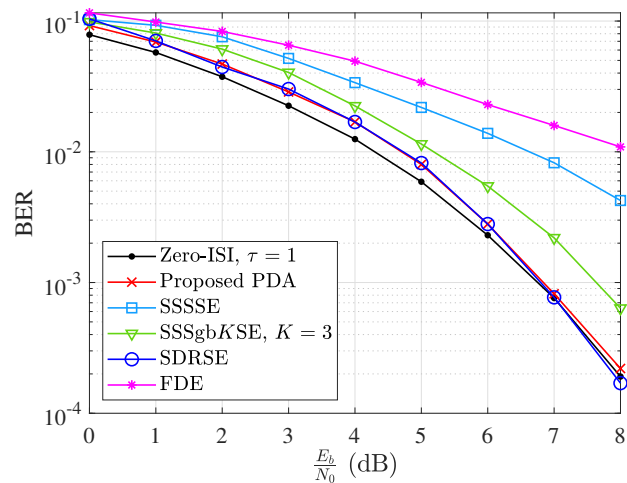


Fig. 1: BER performance of BPSK FTN signaling as a function of $\frac{E_b}{N_o}$ using the proposed PDA algorithm, SSSSE [12], SSSgb$K$SE [12], and the SDRSE [10], at $\beta = 0.3$ and $\tau = 0.8$, SE = 0.96 bits/sec/Hz.

on the number of operations depends mainly on hardware implementations and is outside the scope of this work), while it has higher complexity compared to the work in [12].

## IV. SIMULATION RESULTS

In this section, we assess the performance of our proposed algorithms for detecting BPSK FTN signaling with a root-raised cosine (rRC) pulse and roll-off factor of $\beta = 0.3$. Simulations were performed over the AWGN channel, both with and without channel encoding. The iteration number for proposed PDA is 8 for all cases and the number of iterations we averaged the simulations over are 50,000 and 5,000 for uncoded and coded cases, respectively.

Figure 1 depicts the BER of BPSK FTN signaling as a function of $\frac{E_b}{N_o}$ for the proposed PDA algorithm, comparing it with the SSSSE and SSSgb$K$SE ($K = 3$) algorithms presented in [12], the SDRSE algorithm in [10], and the frequency domain equalizer (FDE) in [11] for $\beta = 0.3$ and $\tau = 0.8$, i.e., a SE of 0.96 bits/sec/Hz. The SE is calculated as $\frac{\log_2 M}{(1+\beta)\tau}$, where $M$ is the constellation size and the SE loss due to adding the cyclic prefix in the FDE in [11] is not considered. As we can see, the proposed PDA algorithm approaches the performance of the SDRSE algorithm in [10] and outperforms the works in [11], [12] for SE value of 0.96 bits/sec/Hz.

Figure 2 plots the BER of the BPSK FTN signaling as a function of $\frac{E_b}{N_o}$ for the proposed PDA algorithm, the SSSSE and SSSgb$K$SE ($K = 3$) algorithms in [12], and the SDRSE algorithm in [10] for $\beta = 0.3$ and $\tau = 0.7$, i.e., a SE of 1.10 bits/sec/Hz. As we can see in Fig. 2, the proposed PDA algorithm's performance is within almost 0.5 dB of the SDRSE performance in [10], and it extensively outperforms the works in [12] that fail at such value of SE. We observe in Fig. 3 of our simulation for the confidence region $\epsilon = 0.4$, the average number of iteration steps reduces by around 30%. At higher SNR values, we observe higher percentage of reduction in the iteration steps.

We also performed the simulation using a low-density parity-check (LDPC) and turbo codes. For implementation, we used a custom parity based matrix generator for LDPC and the long-term evolution (LTE) turbo channel codes described in [16]. For both cases, we used 1/3 code rate channel coding
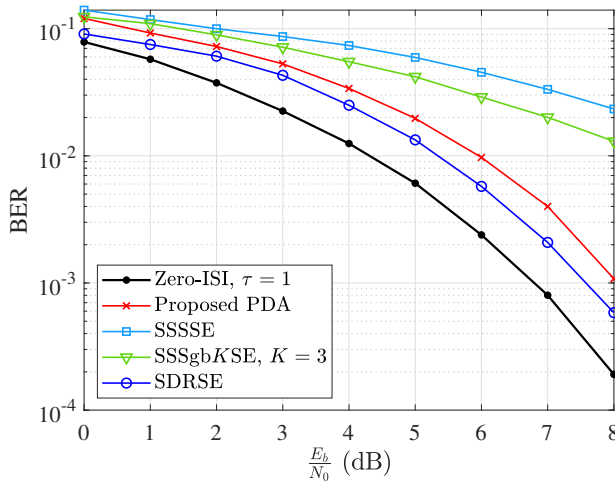
Fig. 2: BER performance of BPSK FTN signaling as a function of $\frac{E_b}{N_o}$ using the proposed PDA algorithm, SSSSE [12], SSSgb$K$SE [12], and the SDRSE [10], at $\beta = 0.3$ and $\tau = 0.7$, SE = 1.10 bits/sec/Hz.
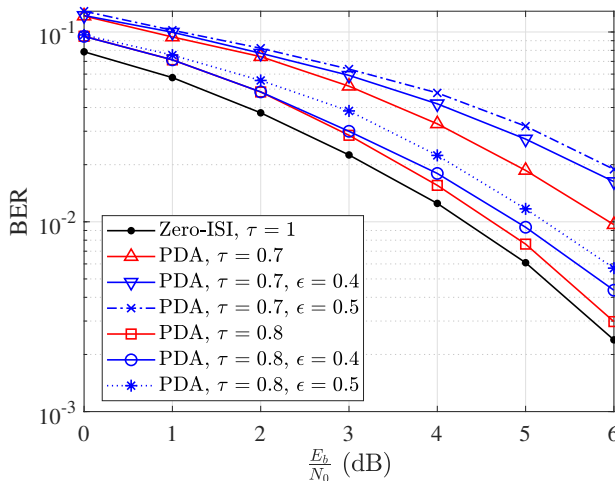


Fig. 3: BER performance for confidence region.

with an input message length of 320 and an output code length of 972 for both LDPC and LTE turbo, respectively. The construction of the LTE interleaver was based on the quadratic permutation polynomial (QPP) scheme [16]. Figure 4 shows the BER performance for LDPC and turbo encoded with $\tau = 0.7$ and $\tau = 0.8$. The BER performance of turbo coded is better compared to LDPC in both cases. The SSSSE and SSSgb$K$SE are not included in the simulation results since they do not produce soft outputs required by the LDPC and turbo soft decoders.

## V. CONCLUSION

Faster-than-Nyquist (FTN) signaling is a promising candidate for improving spectral efficiency (SE) in future communication systems. In this paper, we showed that Gaussian separability has a greater margin compared to linear separability, which has been used recently in the detection of FTN. In exploiting Gaussian separability, we proposed two probabilistic data association (PDA) algorithms of polynomial time computational complexity to detect binary phase-shift keying (BPSK) FTN signaling. Simulation results showed that the proposed algorithms outperform the successive symbol-by-symbol sequence estimation (SSSSE) and successive symbol-by-symbol with go-back-$K$ sequence estimation (SSSgb$K$SE)
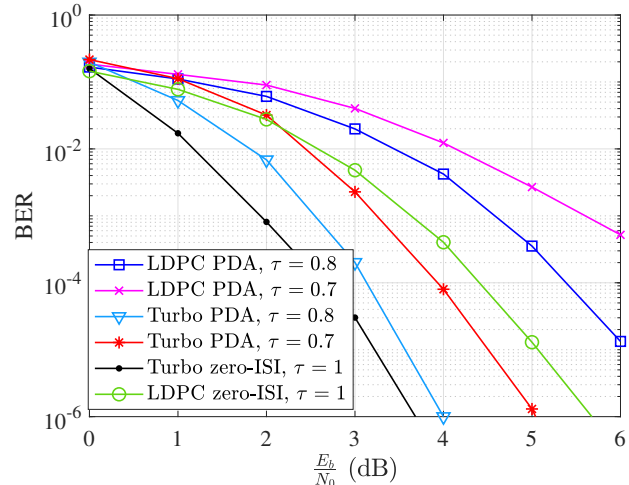


Fig. 4: BER performance for LDPC and turbo encoding.

algorithms in [12] for all SE values. Additionally, they approach the performance the semidefinite relaxation (SDRSE) algorithm in [10] for SE values of 0.96 bits/sec/Hz and they are within the 0.5 dB signal-to-noise ratio (SNR) penalty at SE values of 1.10 bits/sec/Hz.

## REFERENCES

[1] J. B. Anderson, F. Rusek, and V. Öwall, "Faster-than-Nyquist signaling," *Proc. IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug. 2013.
[2] J. Mazo, "Faster-than-Nyquist signaling," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451–1462, Oct. 1975.
[3] F. Rusek and J. B. Anderson, "The two dimensional Mazo limit," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Adelaide, Australia, Sep. 2005, pp. 970–974.
[4] ——, "Non binary and precoded faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 808–817, May 2008.
[5] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 51, no. 9, pp. 1502–1511, Sep. 2003.
[6] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.
[7] J. B. Anderson, A. Prlja, and F. Rusek, "New reduced state space BCJR algorithms for the ISI channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, Korea, Jun. 2009, pp. 889–893.
[8] A. Prlja, J. B. Anderson, and F. Rusek, "Receivers for faster-than-Nyquist signaling with and without turbo equalization," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Toronto, Canada, Jul. 2008, pp. 464–468.
[9] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2951–2959, Oct. 2009.
[10] E. Bedeer, M. H. Ahmed, and H. Yanikomeroglu, "Low-complexity detection of high-order QAM faster-than-Nyquist signaling," *IEEE Access*, vol. 5, pp. 14 579–14 588, Jul. 2017.
[11] S. Sugiura, "Frequency-domain equalization of faster-than-Nyquist signaling," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, Oct. 2013.
[12] E. Bedeer, M. H. Ahmed, and H. Yanikomeroglu, "A very low complexity successive symbol-by-symbol sequence estimator for faster-than-Nyquist signaling," *IEEE Access*, vol. 5, no. 1, pp. 2169–3536, Mar. 2017.
[13] J. Luo, K. R. Pattipati, and P. K. Willett, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361–366, Sept. 2001.
[14] G. Romano, F. Palmieri, P. K. Willett, and D. Mattera, "Separability and gain control for overloaded CDMA," in *Proc. of the Conf. on Inf. Sci. and Sys. (CISS)*, Princeton, NJ, USA, Mar. 2005, pp. 1–6.
[15] T. Wysocki, H. Razavi, and B. Honary, *Digital signal processing for communication systems*. Springer Science & Business Media, 2013.
[16] H. Zarrinkoub, *Understanding LTE with MATLAB: From Mathematical Modeling to Simulation and Prototyping*, 1st ed. Wiley Publishing, 2014.