Blind Estimation of Spreading Sequence and Data Bits in Direct-Sequence Spread Spectrum Communication Systems

HOESANG CHOI AND HICHAN MOON, (Member, IEEE)  
Department of Electrical Engineering, Hanyang University, Seoul 04763, South Korea  
Corresponding author: Hichan Moon (hcmoon@hanyang.ac.kr)  
This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korean government (KNPA), LTE-based accurate positioning technique for emergency rescue, under Grant 2019001291.

ABSTRACT At the transmitter of a Direct Sequence Spread Spectrum (DSSS) communication system, data bits are spread using a spreading sequence. To demodulate the transmitted data bits, the received signal should be de-spread using the same spreading sequence. However, there are many non-cooperative communication systems, where a receiver does not have the information on the spreading sequence. In such systems, to demodulate the transmitted data bits across a wide area, it is necessary to blindly estimate the spreading sequence at low Signal-to-Noise Ratio (SNR). Furthermore, for real-time processing, the computational complexity for the blind estimation should be minimized. In this paper, a novel algorithm is proposed to blindly estimate a spreading sequence and data bits based on turbo processing for both synchronous and asynchronous cases. The proposed algorithm can significantly improve the error rates of estimated data bits and the chips in the spreading sequence compared to an Eigen-Value Decomposition (EVD) based algorithm. Furthermore, compared to the EVD-based algorithm, the proposed algorithm drastically reduces the number of required multiplications.

INDEX TERMS Blind estimation, direct sequence spread spectrum, turbo processing.

I. INTRODUCTION  
Direct Sequence Spread Spectrum (DSSS) has been widely used in both military and commercial communication systems [1]–[6]. At the transmitter of a DSSS communication system, to generate a wideband signal, data bits are multiplied by a spreading sequence, which comprises multiple chips [2]. The transmitted data bits are demodulated after the received signal is despread with the same sequence. The ratio of the chip rate to the data bit rate is defined as processing gain. Due to the processing gain, the received signal has low probability of interception and can be demodulated even below the level of Additive White Gaussian Noise (AWGN) power [2], [5]–[7].

There are several non-cooperative communication systems, where the receiver does not have information on a transmitted DSSS signal. The applications of these systems include eavesdropping, spectrum surveillance and source localization [8]–[11]. In non-cooperative communication systems, it is common for the receiver to possess information on the modulation parameters, including data rate and a spreading sequence used in the transmitter. Furthermore, there are cases where the receiver does not even have information of the presence of a DSSS signal.

The receiver should blindly estimate several modulation parameters and a spreading sequence to recover transmitted data bits. Without the information on a DSSS signal, the usual procedure for blind estimation is implemented as follows [12]: The receiver detects whether there is a DSSS signal. If there exists a DSSS signal, the receiver estimates the modulation parameters such as the data bit rate and the length of a spreading sequence from the received signal. After parameter estimation, a spreading sequence is estimated to demodulate the transmitted signal.

To detect the presence of a DSSS signal, several studies have been performed based on energy radiometry [13] and the autocorrelation of received signals [7], [14]. In addition, several studies have investigated the blind estimation of modulation parameters [15]–[18].

To recover transmitted data bits in non-cooperative communication systems, critical challenge is the estimation of a spreading sequence from a received DSSS signal.
This estimation problem is extremely difficult even with the information of the data bit rate and the length of a spreading sequence. Several studies have examined the blind estimation of a spreading sequence [9], [10], [19]–[25]. In [9], a spreading sequence was estimated in a multi-path environment based on a multi-channel identification technique [26], [27]. Neural network algorithms were applied for the blind estimation of a spreading sequence in [22]–[24]. In [10] and [25], an Eigen-Value Decomposition (EVD)-based algorithm was proposed for the blind estimation of a spreading sequence. This algorithm estimates a spreading sequence based on the EVD of the autocorrelation matrix of the received signal. The EVD-based algorithm is one of the best performing algorithms for blind estimation of a spreading sequence. However, the EVD-based algorithm requires high computational complexity. Furthermore, there exists a partial-encoding problem in an asynchronous case [25]. To solve the partial-encoding problem, several modified algorithms have been proposed based on the EVD-based algorithm in [15], [28]–[30]. However, these algorithms result in increased computation complexity and performance degradation.

In non-cooperative communication systems, it is essential to blindly estimate a spreading sequence and data bits from DSSS signals at low SNR. Furthermore, for real-time processing, computational complexity should be minimized for the blind estimation.

Therefore, in this paper, a novel algorithm is proposed to blindly estimate a spreading sequence and data bits based on turbo processing. With the proposed algorithm, a receiver calculates extrinsic information based on received signals and updates Log-Likelihood Ratios (LLRs), iteratively [31], [32]. Turbo processing has been applied to decoding problems, equalization, coded modulation, and joint source and channel estimation [31]–[33]. However, there has been no previous research on the blind estimation of a spreading sequence and data bits based on turbo processing.

The contributions of this paper are as follows:

- A novel algorithm is proposed to simultaneously blindly estimate a spreading sequence and data bits simultaneously. This is the first approach based on turbo processing in estimating a spreading sequence for non-cooperative communication systems.
- The proposed algorithm can be implemented for both synchronous and asynchronous cases. In other words, even when the receiver does not have prior information on time synchronization, the proposed algorithm provides a method to jointly estimate a spreading sequence and data bits and to acquire time synchronization.
- The performance of the proposed algorithm is evaluated via simulations. From the simulation, it is shown that the proposed algorithm can achieve better estimation performance than that of an EVD-based algorithm. Especially, a performance gain of approximately 3.0 dB performance gain is observed in an asynchronous case.
- The computational complexity is analyzed for the proposed algorithm. As the proposed algorithm drastically reduces the number of multiplications compared with the EVD-based algorithm, it is less computationally complex.
- With the proposed algorithm, a spreading sequence is estimated without a partial-encoding problem.

The remainder of this paper is organized as follows. In Section II, the system model is presented for this paper. In Section III, a novel algorithm is proposed to blindly estimate a spreading sequence and data bits. Then, simulation results are shown and discussed in Section IV. Finally, the conclusions of this paper are presented in Section V.

II. SYSTEM MODEL

In this section, a system model is presented for the proposed blind estimation algorithm. In the system, a transmitter sends a baseband DSSS signal, which is the product of data bits and a spreading sequence $c = \{c_1, c_2, \ldots, c_L\}$, where $L$ is the length of a spreading sequence. The spreading sequence is a random binary sequence of length $L$ and a data bit is modulated by Binary Phase Shift Keying (BPSK). It is assumed that the length of the spreading sequence is equal to spreading gain and that the spreading sequence is repeatedly used to transmit multiple data bits.

The transmitted baseband signal $s(t)$ is expressed as

$$s(t) = \sum_{k=1}^{\infty} a_k \sum_{l=1}^{L} c_l p(t - lT_c - kT_s),$$  \hspace{1cm} (1)

where $a_k$ is the $k$th data bit ($a_k \in \{-1, 1\}$), $c_l$ is the $l$th chip in a spreading sequence ($c_l \in \{-1, 1\}$), $T_c$ is the chip duration of a spreading sequence ($T_c = T_s/L$), $T_s$ is the period of a data bit, and $p(t)$ is a pulse-shaping filter.

In non-cooperative communication systems [8]–[11], a receiver does not have the information of the spreading sequence $c$. Therefore, in this paper, the receiver simultaneously estimates a spreading sequence $c$ and data bits. It is assumed that the data bit period $T_c$ and the length $L$ of a spreading sequence are known to the receiver before the blind estimation of a spreading sequence as in previous studies [9], [10], [19].

At a receiver, a received signal $r(t)$ is expressed as

$$r(t) = s(t - \tau) + w(t), \hspace{1cm} t \geq 0,$$  \hspace{1cm} (2)

where $w(t)$ is the complex AWGN with zero mean and variance $N_0$ and $\tau$ is the unknown desynchronization value between the receiver and a transmitter. In this paper, without loss of generality, it is assumed that $\tau = 0$, but the receiver does not have the information of the $\tau$ value. In addition, $N_0$ is assumed to be known to the receiver and channel estimation is not required since the transmitted signal is a baseband signal. The received signal is sampled at a rate $T_c$. The sampled signal $z[i]$ can be expressed as

$$z[i] = r(iT_c + \zeta), \hspace{0.5cm} i = 0, 1, 2, \ldots,$$  \hspace{1cm} (3)

where $\zeta$ is an arbitrary value in the range of $[0, T_c)$.

1In non-cooperative systems, since a receiver does not have information of the desynchronization value, the boundary of a data bit is not known at the receiver.
Let us define two different types of blind estimation of a spreading sequence and data bits. One is a synchronous case, where a receiver has the information of the desynchronization value \( \tau \). The other is an asynchronous case, where the receiver does not have the information of the desynchronization value \( \tau \).

With the proposed algorithm, a spreading sequence and \( M \) spread data bits are simultaneously estimated based on an input matrix, which is constructed by sampling the received signal. The block diagram of the system is shown in Fig. 1.

Let us define \( y_{m,\Delta} \) as follows:

\[
y_{m,\Delta} = \begin{bmatrix} z[(m-1)L + \Delta] \\ z[(m-1)L + \Delta + 1] \\ \vdots \\ z[mL + \Delta] \end{bmatrix}^T,
\]

where \((\cdot)^T\) is a transpose operator and \( \Delta \) denotes a time offset. Here, the time offset \( \Delta \) is the number of shifted samples to construct an input matrix. Subsequently, with the time offset \( \Delta \), an input matrix \( Y_\Delta \) is expressed as follows:

\[
Y_\Delta = \begin{bmatrix} y_{1,\Delta} \\ y_{2,\Delta} \\ \vdots \\ y_{M,\Delta} \end{bmatrix}.
\]

The element in the \( k \)th row and the \( l \)th column of \( Y_\Delta \) is denoted by \( y_{k,l,\Delta} \). For a time offset \( \Delta \), let us define a time synchronization error \( \epsilon_\Delta \) as the time difference between the beginning of a data bit and the first sample in an input matrix, where \( \epsilon_\Delta \) is a value in the range of \([0, T_a] \).

In a synchronous case, it is sufficient to consider only an input matrix \( Y_\Delta \) with \( \Delta = 0 \). Then, the input matrix \( Y_0 \) is expressed as

\[
Y_0 = \begin{bmatrix} a_1c_1 & a_1c_2 & \cdots & a_1c_L \\ a_2c_1 & a_2c_2 & \cdots & a_2c_L \\ \vdots & \vdots & \ddots & \vdots \\ a_Mc_1 & a_Mc_2 & \cdots & a_Mc_L \end{bmatrix} + W_0,
\]

where \( W_0 \) is the noise part of the input matrix \( Y_0 \) and is expressed as

\[
W_0 = \begin{bmatrix} w[0] & w[1] & \cdots & w[L-1] \\ w[L] & w[L+1] & \cdots & w[2L-1] \\ \vdots & \vdots & \ddots & \vdots \\ w[ML-L] & w[ML-L+1] & \cdots & w[ML-1] \end{bmatrix}.
\]

In the \( k \)th row of \( Y_\Delta \), \( a_k \) is multiplied by the last \( L - \Delta \) chips in the spreading sequence and \( a_{k+1} \) is multiplied by the first \( \Delta \) chips in the spreading sequence due to the synchronization error.

III. PROPOSED ALGORITHM

In this section, the proposed algorithm is presented for blind estimation of a spreading sequence and data bits simultaneously. This proposed algorithm is based on the turbo processing principle for a maximum \textit{a posteriori} probability decision. In the proposed algorithm, the channel LLRs of chips in a spreading sequence and those of data bits are iteratively calculated to obtain \textit{a posteriori} LLRs.
At each iteration, after the channel LLRs of chips in a spreading sequence are first calculated, \emph{a priori} probabilities of chips in a spreading sequence are updated with the calculated channel LLRs. Subsequently, the channel LLRs of data bits are calculated with the updated \emph{a priori} probabilities of chips in a spreading sequence. Moreover, \emph{a priori} probabilities of data bits are updated with the calculated channel LLRs.\footnote{It is possible to change the order of the LLR calculation between chips in a spreading sequence and data bits. For convenience, in this paper, the LLR calculation of chips in a spreading sequence is followed by the LLR calculation of data bits.}

The extrinsic information of the calculated channel LLRs is used in the next iteration.

**A. LOG-LIKELIHOOD RATIO**

Let us consider a synchronous case. In this case, the input matrix $\mathbf{Y}_\Delta$ is constructed with a time offset $\Delta = 0$.

With a time offset $\Delta$, \emph{a posteriori} LLR $L^\Delta(c_l)$ of $c_l$ is defined as

$$L^\Delta(c_l) \triangleq \log \log \frac{\Pr(y_{(k,l)} | c_l = 1)}{\Pr(y_{(k,l)} | c_l = -1)} $$

(13)

where $y_{(k,l)}$ is the $l$th column of $\mathbf{Y}_\Delta$, $L_{ch}(y_{(k,l)} | c_l)$ is the channel LLR of $y_{(k,l)}$ for $c_l$ and $L_{pr}(c_l)$ is \emph{a priori} LLR of $c_l$. Here,

$$L_{ch}(y_{(k,l)} | c_l) = \log \left[ \frac{\Pr(y_{(k,l)} | c_l = 1)}{\Pr(y_{(k,l)} | c_l = -1)} \right],$$

(14)

and

$$L_{pr}(c_l) = \log \left[ \frac{\Pr(c_l = 1)}{\Pr(c_l = -1)} \right].$$

(15)

With a time offset $\Delta$, \emph{a posteriori} LLR $L(a_k)$ of $a_k$ is defined as

$$L^\Delta(a_k) \triangleq L(a_k | u_{(k,\Delta)})
\begin{align*}
&= L_{ch}(a_k | u_{(k,\Delta)}) + L_{pr}(a_k) \\
&= \sum_{l=1}^{L} L_{ch}(y_{(k,l)} | a_k) + L_{pr}(a_k),
\end{align*}$$

(16)

where $u_{(k,\Delta)}$ is the $k$th row of $\mathbf{Y}_\Delta$, $L_{ch}(y_{(k,l)} | a_k)$ is the channel LLR of $y_{(k,l)}$ for $a_k$ and $L_{pr}(a_k)$ is \emph{a priori} LLR of $a_k$. Here,

$$L_{ch}(y_{(k,l)} | a_k) = \log \left[ \frac{\Pr(y_{(k,l)} | a_k = 1)}{\Pr(y_{(k,l)} | a_k = -1)} \right],$$

(17)

(18)

(19)

As $\max(x_1, x_2)$ is the same as $(x_1 + x_2 + |x_1 - x_2|)/2$, the channel LLR of $y_{(k,l)}$ for $c_l$ is expressed as

$$L_{ch}(y_{(k,l)} | c_l) \approx \frac{4\text{Re}[y_{(k,l)}]}{N_0} + L_{pr}(a_k) - \frac{4\text{Re}[y_{(k,l)}]}{N_0} + L_{pr}(a_k),$$

(20)

where $\text{sgn}(x)$ is a sign function [31].

**FIGURE 3. LLR calculation from an input matrix $Y_\Delta$.**

where $u_{(k,\Delta)}$ is the $k$th row of $\mathbf{Y}_\Delta$, $L_{ch}(y_{(k,l)} | a_k)$ is the channel LLR of $y_{(k,l)}$ for $a_k$ and $L_{pr}(a_k)$ is \emph{a priori} LLR of $a_k$. Here,

$$L_{ch}(y_{(k,l)} | a_k) = \log \left[ \frac{\Pr(y_{(k,l)} | a_k = 1)}{\Pr(y_{(k,l)} | a_k = -1)} \right],$$

(16)

and

$$L_{pr}(a_k) = \log \left[ \frac{\Pr(a_k = 1)}{\Pr(a_k = -1)} \right].$$

(17)

From (12) and (15), it is observed that \emph{a posteriori} LLR of $c_l$ can be calculated with the $l$th column of $\mathbf{Y}_\Delta$ and \emph{a priori} LLR of $a_k$ can be calculated with the $k$th row of $\mathbf{Y}_\Delta$. Fig. 3 shows how to calculate \emph{a posteriori} LLRs from an input matrix.

In a synchronous case, as $y_{(k,l)}$ is expressed as $y_{(k,l)} = a_k c_l + \alpha_{(k,l)}$, the channel LLR of $y_{(k,l)}$ for $c_l$ is calculated as (18), shown at the bottom of the page where $y(\gamma) = -x^2 + \log \Pr(y), \Re[x]$ is a real part of $x$ and the approximation \footnote{It is possible to change the order of the LLR calculation between chips in a spreading sequence and data bits. For convenience, in this paper, the LLR calculation of chips in a spreading sequence is followed by the LLR calculation of data bits.} is obtained from [34]

$$\log[\exp(x_1) + \exp(x_2)] \approx \max(x_1, x_2).$$

(19)
Similarly, $L_{ch}(\gamma_{(k,l)}^\Delta | a_k)$, the channel LLR of $\gamma_{(k,l)}^\Delta$ for $a_k$, can be expressed as

$$L_{ch}(\gamma_{(k,l)}^\Delta | a_k) = \log \left[ \frac{\sum_{c_l \in \{-1,1\}} \Pr(\gamma_{(k,l)}^\Delta | a_k = 1, c_l) \Pr(c_l)}{\sum_{c_l \in \{-1,1\}} \Pr(\gamma_{(k,l)}^\Delta | a_k = -1, c_l) \Pr(c_l)} \right],$$

which is approximated by

$$\approx \left[ \frac{4\Re[\gamma_{(k,l)}^\Delta]}{N_0} + L_{pr}(c_l) \right] - \left[ \frac{-4\Re[\gamma_{(k,l)}^\Delta]}{N_0} + L_{pr}(c_l) \right],$$

$$= \left\{ \begin{array}{ll}
\sgn \left( \Re[\gamma_{(k,l)}^\Delta] \right) 2L_{pr}(c_l), & \text{if } \left| \Re[\gamma_{(k,l)}^\Delta] \right| > \frac{N_0}{4}L_{pr}(c_l), \\
\frac{8\Re[\gamma_{(k,l)}^\Delta]}{N_0}, & \text{otherwise.}
\end{array} \right. \tag{21}$$

For the turbo processing, the extrinsic information for $c_l$ and $a_k$ are denoted as $\lambda_{c_l}^\Delta$ and $\lambda_{a_k}^\Delta$, respectively. The extrinsic information is expressed as follows

$$\lambda_{c_l}^\Delta = \frac{1}{2n_1}L_{ch}(v_l, \lambda_l(c_l)), \tag{22}$$

$$= \frac{1}{2n_1} \sum_{k=1}^M L_{ch}(\gamma_{(k,l)}^\Delta | c_l),$$

$$\lambda_{a_k}^\Delta = \frac{1}{2n_2}L_{ch}(a_k, \lambda_k | a_k), \tag{23}$$

$$= \frac{1}{2n_2} \sum_{l=1}^L L_{ch}(\gamma_{(k,l)}^\Delta | a_k),$$

where $n_1$ and $n_2$ are the smallest integer greater than or equal to $\log_2 M$ and $\log_2 L$, respectively. The extrinsic information is used to update a priori probability.

### B. SYNCHRONOUS CASE

Let us consider a synchronous case, where the receiver has the information of a desynchronization value $\tau$. Then, the a posteriori LLRs of chips in the spreading sequence $L_0^0(c_l)$ and those of data bits $L_0^0(a_k)$ are calculated from (12) and (15) for the input matrix $Y_0$, respectively. At each iteration, $L_0^0(c_l)$ and $L_0^0(a_k)$ are calculated for $l = 1, 2, \ldots, L$ and $k = 1, 2, \ldots, M$. To calculate $L_0^0(c_l)$ and $L_0^0(a_k)$, the a priori probabilities $\Pr(c_l = 1)$ and $\Pr(a_k = 1)$ are updated based on the extrinsic information obtained from the previous iteration.

Before the first iteration, the a priori probabilities $\Pr(c_l = 1)$ and $\Pr(a_k = 1)$ are initialized for all $l$ and $k$ values. After the initialization, $L_0^0(c_l)$ is first calculated from (12). Subsequently, the extrinsic information $\lambda_{a_k}^0$ for $c_l$ is fed back to update the a priori probability of $\Pr(c_l = 1)$ as follows:

$$\Pr(c_l = 1) \leftarrow f_{atv} \left( \lambda_{c_l}^0 \right). \tag{24}$$

where $f_{atv}(x)$ is an activation function. In this paper, as in [35]–[37], a logistic function is used for the activation function. This function is expressed as [38]:

$$f_{atv}(x) = \frac{1}{1 + e^{-\mu x}}, \tag{25}$$

where $\mu$ is a positive coefficient. Fig. 4 shows plots of the activation function (25) for different coefficient $\mu$ values.

After $\Pr(c_l = 1)$ is updated, $L_0^0(a_k)$ is calculated from (15). Then, the extrinsic information $\lambda_{a_k}^0$ for $a_k$ is also fed back to update the a priori probability of $\Pr(a_k = 1)$ as follows:

$$\Pr(a_k = 1) \leftarrow f_{atv} \left( \lambda_{a_k}^0 \right). \tag{26}$$

The updated a priori probabilities are used at the next iteration to calculate $L_0^0(c_l)$ and $L_0^0(a_k)$. Fig. 5 summarizes the overall iterative procedure for the proposed algorithm.

After the end of the last iteration, the estimated chips in a spreading sequence and data bits are obtained based on the sign of $L_0^0(c_l)$ and $L_0^0(a_k)$ as follows:

$$\tilde{c}_l = \sgn \left( L_0^0(c_l) \right), \quad l = 1, 2, \ldots, L, \tag{27}$$

$$\tilde{a}_k = \sgn \left( L_0^0(a_k) \right), \quad k = 1, 2, \ldots, M. \tag{28}$$

### C. ASYNCHRONOUS CASE

Let us consider an asynchronous case, where the receiver does not have the information of the desynchronization value $\tau$. In this case, the receiver estimates a spreading sequence and data bits for all possible desynchronization values. Therefore, $L$ input matrices are generated from the consecutive $(M + 1)L - 1$ samples with $L$ time offsets $\Delta$. 
Subsequently, among the \( L \) estimations of chips in a spreading sequence and data bits, the receiver selects the most reliable one.

For each time offset \( \Delta \), the \textit{a posteriori} LLRs of chips in a spreading sequence and those of data bits are calculated from (12) and (15). After the end of the last iteration for each \( \Delta \) value, tentative estimations of a spreading sequence and data bits are obtained as

\[
\hat{c}_i(\Delta) = \text{sgn}(L^\Delta(c_i)), \quad l = 1, 2, \cdots, L,
\]

\[
\hat{a}_k(\Delta) = \text{sgn}(L^\Delta(a_k)), \quad k = 1, 2, \cdots, M.
\]

Then, from the tentative estimation, an input matrix \( \hat{Y}_\Delta \) is regenerated as

\[
\hat{Y}_\Delta = \text{sgn} \left[ \begin{bmatrix} \hat{c}_1\Delta \\ \hat{c}_2\Delta \\ \vdots \\ \hat{c}_L\Delta \end{bmatrix} \cdot [\hat{a}_1\Delta, \hat{a}_2\Delta, \cdots, \hat{a}_M\Delta] \right].
\]

The receiver computes the Euclidean distance between an input matrix \( Y_\Delta \) and a regenerated input matrix \( \hat{Y}_\Delta \). The computed Euclidean distance is used as a reliability measure for a tentative estimation. Let us define \( \Theta \) as the time offset \( \Delta \), with which the minimum Euclidean distance is obtained. Then, \( \Theta \) is expressed as

\[
\Theta = \arg \min_{\Delta} \left( \left| Y_\Delta - \hat{Y}_\Delta \right|^2 \right)
\]

\[
= \arg \min_{\Delta} \left( \sum_{k=1}^{M} \sum_{l=1}^{L} \epsilon^2[(k-1)L + l - 1 + \Delta] - 2\hat{Y}_{(k,l)}\Delta^\Delta[(k-1)L + l - 1 + \Delta] + (\hat{Y}_{(k,l)}\Delta^\Delta)^2 \right)
\]

\[
\equiv \arg \min_{\Delta} \left( \sum_{k=1}^{L-1} \sum_{l=1}^{L} \epsilon^2[(k-1)L + l - 1 + \Delta] \right)
\]

where \( \hat{Y}_{(k,l)}\Delta^\Delta \) is the \( k \)th row and the \( l \)th column element of \( \hat{Y}_\Delta \). (a) is derived as follows:

(i) In the second line of (32), with \( i = L, L+1, \cdots, ML-1 \), the first term \( \epsilon^2[(k-1)L + l - 1 + \Delta] \) is included for all \( \Delta \). Therefore, this term can be simplified as the first and the second terms of the third line in (32).

(ii) As the value of \( \hat{Y}_{(k,l)}\Delta^\Delta \) is 1 or \(-1\), \( \hat{Y}_{(k,l)}\Delta^\Delta^2 \) is always one for all \( k \) and \( l \).

Therefore, the final estimation of a spreading sequence and data bits can be obtained as follows:

\[
\hat{c}_i = \hat{c}_i(\Theta), \quad l = 1, 2, \cdots, L,
\]

\[
\hat{a}_k = \hat{a}_k(\Theta), \quad k = 1, 2, \cdots, M.
\]

Algorithm 1 presents the proposed algorithm as a pseudo code for an asynchronous case.

**Algorithm 1 Blind Estimation of a Spreading Sequence and Data Bits**

**Input**: \((M + 1)L - 1\) consecutive samples

**Output**: Estimated chips in a spreading sequence and data bits

1. **for** \( \Delta = 0 \) to \( L - 1 \) **do**
2. \[ Initialization \text{ of} \Pr(c_l) \text{ and} \Pr(a_k) \]
3. \[ for \quad n = 1 \text{ to} \quad N_{\text{iter}} \text{ do} \]
4. \[ \text{calculate} \phantom{L^\Delta} L_{\text{ch}}(y_{(k,l)}^\Delta | c_l), L^\Delta(c_l) , \lambda_{c_l}^\Delta ; \]
5. \[ \Pr(c_l) \leftarrow f_{\text{avg}}(\lambda_{c_l}^\Delta) ; \]
6. \[ \hat{c}_l(\Delta) = \text{sgn}(L^\Delta(c_l)) ; \]
7. \[ end \]
8. \[ \text{for} \quad k = 1 \text{ to} \quad M \text{ do} \]
9. \[ \text{calculate} \phantom{L^\Delta} L_{\text{ch}}(y_{(k,l)}^\Delta | a_k) , L^\Delta(a_k) , \lambda_{a_k}^\Delta ; \]
10. \[ \Pr(a_k) \leftarrow f_{\text{avg}}(\lambda_{a_k}^\Delta) ; \]
11. \[ \hat{a}_k(\Delta) = \text{sgn}(L^\Delta(a_k)) ; \]
12. \[ end \]
13. \[ end \]
14. \[ \Theta = \arg \min_{\Delta} \left( \left| Y_\Delta - \hat{Y}_\Delta \right|^2 \right) ; \]
15. \[ \hat{c}_i = \hat{c}_i(\Theta), \quad l = 1, 2, \cdots, L ; \]
16. \[ \hat{a}_k = \hat{a}_k(\Theta), \quad k = 1, 2, \cdots, M ; \]

**D. Computational Complexity**

In this subsection, the computational complexity is analyzed for the proposed algorithm. In addition, the computational complexity is compared with that of the conventional EVD-based algorithm.

Let us first consider a synchronous case. The computational complexity is summarized as follows:

- Normalization of an \( M \times L \) input matrix by \( N_0 \) (signal-to-noise ratio normalization):
  - \( ML \) multiplications
- Calculation of channel LLRs as in (20) and (21):
  - \( 2MLN_{\text{iter}} \) additions$^4$
- Calculation of extrinsic information based on the calculated channel LLRs as in (22) and (23):
  - \( (2ML - M - L)N_{\text{iter}} \) additions
  - \( (M + L)N_{\text{iter}} \) bit shift operations
- Update of \textit{a priori} probabilities using a logistic function as in (24) and (26), that can be implemented with a look-up table [39]:
  - \( (M + L)N_{\text{iter}} \) memory accesses for a one-dimensional look-up table of a logistic function
- Calculation of \textit{a priori} LLRs similar to that in (14) and (17), which can be implemented with a look-up table:

$^4$The calculation of a channel LLR comprises of a magnitude comparison of an input value and multiplications by a power of two. The complexity of magnitude comparison is the same as that of an addition. In addition, multiplication by a power of two can be implemented with a shift operation, which has negligible computational complexity compared with addition.
Therefore, the proposed algorithm requires approximately $ML$ multiplications and $4N_{\text{iter}}ML$ additions for the synchronous case, as the shift operation and memory access for a look-up table are negligible compared with an addition or a multiplication.

On the other hand, in the asynchronous case, the proposed algorithm requires $L$ input matrices of size $M \times L$ obtained from $(M + 1)L - 1$ samples. In addition, to calculate the Euclidean distance between an input matrix $Y_\Delta$ and its regen-

erated matrix $\tilde{Y}_\Delta$, $2L - 1$ multiplications and $L(ML - 1)$ additions are required. Therefore, the proposed algorithm requires $(M + 1)L + 2L - 2$ multiplications and $4N_{\text{iter}}ML^2 + L(ML - 1)$ additions. Thus, for the asynchronous case, computational complexity is approximately $L$ times that of the synchronous case.

For the conventional EVD-based algorithm [10], [20], [25], it is necessary to compute the average of $M$ covariance matrices of size $L \times L$. Subsequently, after the calculation of eigen-decomposition of the averaged covariance matrix, the spreading sequence is estimated based on two largest eigenvalues and the corresponding eigenvectors. Therefore, $ML^2 + L^3$ additions and $ML^2 + L^3$ multiplications are required for this algorithm [21], [40], [41]. With the EVD-based algorithm, almost the same computational complexity is required for both synchronous and asynchronous cases.

Table 1 summarizes the computational complexity of the EVD-based algorithm and proposed algorithm in synchronous and asynchronous cases.

### IV. NUMERICAL RESULT

In this section, simulation results of the proposed algorithm are presented. A simulator was built to obtain the results for the proposed algorithm in an AWGN channel. The performance of the proposed algorithm was compared with that of an EVD-based algorithm.

The estimation performances were obtained from 10,000 runs of Monte Carlo simulations for an SNR value between $-20.0$ dB and $-10.0$ dB. For the simulation, 100 data bits were spread with a spreading sequence of length $31$ and the coefficient $\mu$ for an activation function is fixed to 3.0. The results were obtained after $N_{\text{iter}}$ iterations, where $N_{\text{iter}}$ is an integer in the range $[1, 25]$. In the asynchronous cases, the desynchronization value was set to $iT_c$, where $i$ is an integer, uniformly distributed between 0 and $L - 1$. Table 2 summarizes the parameters used for the simulation.

### TABLE 1. Computational complexity.

|                          | Proposed Algorithm | EVD-based Algorithm |
|--------------------------|--------------------|---------------------|
|                          | Sync.              | Async.              |                    |
| Multiplication           | $ML$               | $\approx (M + 1)L$  | $ML^2 + L^3$       |
| Addition                 | $4N_{\text{iter}}ML$ | $\approx 4N_{\text{iter}}ML^2$ | $ML^2 + L^3$       |
| Lookup table             | $2N_{\text{iter}}(M + L)$ | $2N_{\text{iter}}L(M + L)$ | 0                  |

- $(M + L)N_{\text{iter}}$ memory accesses for an one-dimensional look-up table of a logarithm function
- Calculation of $a$ posteriori LLRs as in (12) and (15):
  - $(M + L)N_{\text{iter}}$ additions

Let us define $P_e(x)$, which is an ideal BER (bit error rate) performance of BPSK for SNR of $x$ after despreading. $P_e(x)$ is expressed as

$$P_e(x) = \frac{1}{2} \text{erfc} \left( \sqrt{Gx} \right),$$

where $\text{erfc}(x)$ is an error function and $G$ is the processing gain [42]. As the proposed estimation algorithm cannot out-

perform the ideal BER $P_e(x)$, $P_e(x)$ is used as a lower bound for estimation performance. For the estimation of a data bit, as a data bit is spread by a spreading sequence of length $L$, $G$ becomes $L$. On the other hand, for the estimation of a chip in a spreading sequence, as it can be regarded that a chip in a spreading sequence is spread by a sequence of length $M$, $G$ becomes $M$.

Fig. 6 shows the error rate versus SNR for a chip in a spreading sequence and a data bit. The solid and dashed lines represent the performances of the algorithms in synchronous and asynchronous cases, respectively. The results were obtained after 10 iterations with the proposed algorithm.

Fig. 6.(a) shows the plots of the error rate of a chip in a spreading sequence for the proposed and EVD-based algorithms. In the synchronous case, for the error rate of $10^{-2}$, $-13.9$ dB and $-12.6$ dB are required with the proposed and EVD-based algorithms, respectively. Furthermore, in the asynchronous case, for the error rate of $10^{-2}$, $-13.4$ dB and $-10.0$ dB are required with the proposed and EVD-based algorithms, respectively. Compared to the EVD-based algorithm, the proposed algorithm achieves better performance with respect to chips in a spreading sequence. Particularly, in the asynchronous case, the proposed algorithm has performance gain of approximately $3.0$ dB compared with the EVD-based algorithm.

Fig. 6.(b) shows plots of the error rate of a data bit for a proposed and EVD-based algorithms. In the synchronous case, for the error rate of $10^{-1}$, $-14.7$ dB and $-14.5$ dB are required with the proposed and EVD-based algorithms, respectively. Furthermore, in the asynchronous case, for the error rate of $10^{-1}$, $-14.4$ dB and $-13.6$ dB are required with the proposed and EVD-based algorithms, respectively. Compared to the EVD-based algorithm, the proposed algo-

rithm achieves better performance with respect to data bit error rate.

From the Fig. 6.(a) and Fig. 6.(b), it is evident that the estimation performance of a chip in a spreading sequence is
better than that of a data bit for a fixed SNR. This is because the size of column $M$ is smaller than that of row $L$ with an $M \times L$ input matrix.

The proposed algorithm is based on the iterative calculations of a posteriori LLRs. Therefore, the estimation performance is dependent on the number $N_{\text{iter}}$ of iterations. To investigate the influence of $N_{\text{iter}}$ on the estimation performance, Fig. 7 shows the plots of the error rate for different $N_{\text{iter}}$. As $N_{\text{iter}}$ increases, the error rate of a chip in a spreading sequence converges to approximately $2.1 \times 10^{-2}$ or $4.2 \times 10^{-2}$ in a synchronous or an asynchronous case, respectively. In addition, the error rate of a data bit converges to approximately $9.8 \times 10^{-2}$ or $12.5 \times 10^{-2}$ in synchronous or asynchronous case, respectively. Fig. 7 show that the performances saturate after 10 and 20 iterations for an asynchronous and a synchronous cases, respectively. In the asynchronous case, as estimation performance degrades due to the time synchronization error, the error rate converges to a higher value compared with that in the synchronous case.

Therefore, in the asynchronous case, the error rate saturates after fewer iterations than in the synchronous case.

V. CONCLUSION

In this paper, a novel algorithm is proposed to blindly estimate chips in a spreading sequence and data bits based on turbo processing principle in non-cooperative baseband communication systems. With the proposed algorithm, the a posteriori LLRs of chips in a spreading sequence and data bits are calculated and a priori probabilities are updated based on channel LLRs. The proposed algorithm can be implemented for both synchronous and asynchronous cases. Therefore, regardless of the information of a desynchronization value, chips in a spreading sequence and data bits can be estimated simultaneously. The simulation results show that the proposed algorithm can achieve better estimation performance than an EVD-based algorithm. Especially, in the asynchronous case, a performance gain of approximately $3.0 \text{ dB}$ was achieved compared to the EVD-based algorithm to estimate the chips in a spreading sequence.

In addition, this paper analyzed the computation complexity of the proposed algorithm. The complexity of the proposed algorithm is compared with that of an EVD-based algorithm. As the proposed algorithm considerably reduced the number of multiplications required, it is less computationally complex than the EVD-based algorithm.

REFERENCES

[1] R. L. Peterson, D. E. Borth, and R. E. Ziemer, An Introduction to Spread-Spectrum Communications, 1st ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.

[2] A. Viterbi, CDMA: Principles of Spread Spectrum Communication (Addison-Wesley Wireless Communications Series). Reading, MA, USA: Addison-Wesley, 1995.

[3] D. T. Magill, F. D. Natali, and G. P. Edwards, “Spread-spectrum technology for commercial applications,” Proc. IEEE, vol. 82, no. 4, pp. 572–584, Apr. 1994.

[4] A. Tayebi, S. Berber, and A. Swain, “Security enhancement of fix chaotic-DSSS in WSNs,” IEEE Commun. Lett., vol. 22, no. 4, pp. 816–819, Apr. 2018.

[5] Y. Xing, A. Hu, J. Zhang, L. Peng, and G. Li, “On radio frequency fingerprint identification for DSSS systems in low SNR scenarios,” IEEE Commun. Lett., vol. 22, no. 11, pp. 2326–2329, Nov. 2018.
[6] A. Masmoudi, F. Bellili, S. Affes, and A. Ghnayem, “Maximum likelihood time delay estimation from single-and multi-carrier DSSS multipath MIMO transmissions for future 5G networks,” IEEE Trans. Wireless Commun., vol. 16, no. 8, pp. 4851–4865, Aug. 2017.

[7] G. Burel, C. Boudier, and O. Berder, “Detection of direct sequence spread spectrum transmissions without prior knowledge,” in Proc. IEEE GLOBE- COM, vol. 1, Nov. 2001, pp. 236–239.

[8] H. Zhang, P. Wei, and Q. Mou, “A semidefinite relaxation approach to blind despreading of long-code DS-SS signal with carrier frequency offset,” IEEE Signal Process. Lett., vol. 20, no. 7, pp. 705–708, Jul. 2013.

[9] M. K. Tsatsanis and G. B. Giannakis, “Blind estimation of direct sequence spread spectrum signals in multipath,” IEEE Trans. Signal Process., vol. 45, no. 5, pp. 1241–1252, May 1997.

[10] G. Burel and C. Boudier, “Blind estimation of the pseudo-random sequence of a direct sequence spread spectrum signal,” in Proc. MILCOM, vol. 2, Oct. 2000, pp. 967–970.

[11] C. French and W. Gardner, “Spread-spectrum despreading without the code,” IEEE Trans. Commun., vol. 34, no. 4, pp. 404–407, Apr. 1986.

[12] L. Chang, F. Wang, and Z. Wang, “Detection of DSSS signal in non-cooperative communications,” in Proc. Int. Conf. Commun. Technol., Nov. 2006, pp. 1–4.

[13] S. Davidioci and E. G. Kanterakis, “Radiometric detection of direct-sequence spread-spectrum signals using interference excision,” IEEE J. Sel. Areas Commun., vol. 7, no. 4, pp. 576–589, May 1989.

[14] A. Polydoros and C. Weber, “Detection performance considerations for direct-sequence and time-hopping LPI waveforms,” IEEE J. Sel. Areas Commun., vol. 3, no. 5, pp. 727–744, Sep. 1985.

[15] C. Boudier, S. Azou, and G. Burel, “A robust synchronization procedure for blind estimation of the symbol period and the timing offset in spread spectrum transmissions,” in Proc. IEEE 7th Int. Symp. Spread Spectr. Techn. Appl., vol. 1, Sep. 2002, pp. 238–241.

[16] H. Chang, F. Wang, and Z. Wang, “A blind estimation algorithm of parameters in PN sequence estimation of DSSS signal in non-cooperative communication system,” J. Syst. Eng. Electron., vol. 18, no. 1, pp. 14–21, Mar. 2007.

[17] Z. Qiu, H. Peng, and T. Li, “A blind despreading and demodulation method for QPSK-DSSS signal with unknown carrier offset based on matrix subspace analysis,” IEEE Access, vol. 7, pp. 125700–125710, Sep. 2019.

[18] P. Y. Qi, Z. T. Huang, W. L. Jiang, and C. Zhang, “Improved blind-spreading sequence estimation algorithm for direct sequence spread spectrum signals,” IET Signal Process., vol. 2, no. 2, pp. 139–146, Jun. 2008.

[19] S. Mehboodi, A. Jamshidi, and M. Farhang, “A low-complexity near-optimal algorithm for blind estimation of pseudo-noise sequences in DSSS communication systems,” in Proc. 8th Int. Symp. Telecommun. (IST), Sep. 2016, pp. 218–221.

[20] C. Liang, W. Fuping, and W. Zanji, “Low complexity method for spreading sequence estimation of DSSS signal in non-cooperative communication systems,” J. Syst. Eng. Electron., vol. 20, no. 1, pp. 41–49, 2009.

[21] B. Sklar, Digital Communications: Fundamentals and Applications. Englewood Cliffs, NJ, USA: Prentice-Hall, 2016.

HOESANG CHOI received the B.S. and M.S. degrees in electronics engineering from Hanyang University, Seoul, South Korea, in 2014 and 2016, respectively, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering. His research interests include wireless communication systems and communication theory.

HICHAN MOON (Member, IEEE) received the B.S. (summa cum laude) and M.S. degrees in electronics engineering from Seoul National University, Seoul, South Korea, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2004. He was with Samsung Electronics Company, South Korea, from 1994 to February 2011. In Samsung, he was involved in development of mobile station modems and standardization of wireless cellular systems including 3GPP W-CDMA, LTE, and LTE-Advanced. Since March 2011, he has been an Associate Professor with the Department of Electronic Engineering, Hanyang University, Seoul. His research interests include wireless communication systems, resource allocation, and communication theory. He holds 50 issued U.S. patents, issued 80 Korean patents and several pending.