EUSIPCO 2013 1569743715

IMPROVING THE DECODING OF M-SEQUENCES BY EXPLOITING THEIR DECIMATION PROPERTY

Mathieu des Noes and *Valentin Savin*

CEA, LETI, Minatec campusB
38054 Grenoble cedex9, France
Email: mathieu.desnoes@cea.fr

Laurent Ros and *Jean Marc Brossier*

GIPSA-Lab
BP46, 38402 Saint-Martin d'Hères, France

ABSTRACT

M-sequences are widely used in communications and positioning systems for synchronization purposes. In these systems, the receiver does not know which sequence is used by the transmitter, this needs to be detected. This paper establishes first a link between the conventional detection theory and the recently developed detection technique relying on iterative message-passing algorithm. Then, a novel decoding strategy is proposed. It exploits the decimation property of m-sequences to improve significantly the detection performance compared to the existing decoding strategy.

Index Terms— m-sequence, decoding, belief-propagation

1. INTRODUCTION

A maximal length sequence (M-sequence) is a binary sequence with excellent auto-correlation properties [1]. Hence, they are widely used for synchronization purpose in wireless communications and positioning systems. They are for instance used for the cell search procedure in the WCDMA system or for the acquisition of GPS's satellites [2][3].
The conventional method to synchronize with a m-sequence is to correlate the received signal with a replica of the searched m-sequence [4]. If a correlation peak is observed and is above a given threshold, the synchronization is declared. This correlation can be implemented either with a standard FIR filter whose coefficients are equal to the chips of the sequence, or with a FFT [5].
An alternative method consists in performing synchronization through a decoding of the received sequence. In fact, a m-sequence generator can be regarded as a linear code generator. It is thus possible to detect a transmitted sequence with a suitable decoder. This solution was originally proposed in cryptography for fast correlation attacks on stream ciphers [6][7]. This has been applied more recently in wireless communications and localization [8][9]. Exploiting the unique properties of m-sequence, an iterative message-passing algorithm can be implemented to decode the received signal [10]. The main drawback of this decoding procedure is its sensitivity to the weight of the generator polynomial of the m-sequence. The weight is given by the number of non zero coefficients of the polynomial.
In this paper, the link between conventional detection theory and sequence decoding is first established. Then, a novel decoding strategy is proposed. It exploits the decimation property of m-sequences to decode any m-sequence of polynomial degree r, with the generator polynomial of the m-sequence having the smallest weight. This ensures to improve the decoding performances.
The paper is organized as follows. Section 2 establishes the relationship between the Generalized Likelihood Ratio Test technique (GLRT) and iterative decoding for the detection of m-sequences. Section 3 describes the main properties of m-sequences that will be exploited in this paper. Section 4 details the conventional message-passing algorithm used for decoding m-sequence. Section 5 presents the novel algorithm exploiting the decimation property of m-sequences. Section 6 presents simulation results and Section 7 concludes this paper.
Notation: a sequence will be written in uppercase letters in its anti-modal representation ($S(k) \in \{-1, +1\}$) and in lowercase in its binary representation ($s(k) \in \{0, 1\}$). $\|\mathbf{S}\|^2$ is the Euclidian norm of vector $\mathbf{S}$.

2. RELATIONSHIP BETWEEN GLRT AND DECODING

Many synchronization problems involving m-sequences can be modeled as a binary hypothesis testing problem [11]. At each sampling time, the receiver wants to decide for one of the 2 hypothesis :

$$\begin{aligned} H_1 &: \mathbf{Y} = \mathbf{S} + \mathbf{n} \\ H_0 &: \mathbf{Y} = \mathbf{n} \end{aligned} \tag{1}$$

When the receiver is synchronized, hypothesis H_1 is the correcte one, while H_0 is valid otherwise. However, the receiver can be synchronized with the transmitted sequence but without knowing which sequence was sent. This happens for instance in the cell search of UMTS and CDMA200 systems or for the acquisition of C/A code in the GPS system. The re-

ceiver has to detect the transmitted sequence and doing so it is synchronized with this sequence.

$\mathbf{Y} = (Y(0), \ldots, Y(N-1))^T$ is the vector containing the N received samples. $\mathbf{n}$ is a vector of white Gaussian noise with covariance matrix $\sigma^2 \mathbf{I}_N$. $\mathbf{S} = (S(0), \ldots, S(N-1))^T$ is the m-sequence to be detected. If sequence $\mathbf{S}$ is known by the receiver, a conventional Log Likelihood Ratio Test (Log-LRT) can be applied.

Let $L(\mathbf{Y}) = \frac{p(\mathbf{Y}|H_1)}{p(\mathbf{Y}|H_0)}$ be the likelihood ratio. $p(\mathbf{Y}|H_1)$ is the pdf of $\mathbf{Y}$ under hypothesis H_1, and $p(\mathbf{Y}|H_1)$ the pdf of $\mathbf{Y}$ given under hypothesis H_0. Since the noise is white with covariance matrix $\sigma^2 \mathbf{I}_N$, these two pdf are defined by :

$$\begin{aligned} p(\mathbf{Y}|H_1) &= N(\mathbf{S}, \sigma^2 \mathbf{I}_N) \\ p(\mathbf{Y}|H_0) &= N(0, \sigma^2 \mathbf{I}_N) \end{aligned}$$

Since $S(k) \pm 1$, we have $\|\mathbf{S}\|^2 = N$. The Log-LRT becomes [12]:

$$T(\mathbf{Y}) = \Re(\mathbf{Y}^H.\mathbf{S}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{2}$$

$\Re(z)$ is the real part of the complex variable z.

γ is the detection threshold which is set according to the desired missed detection and false alarm probabilities:

$$\begin{aligned} P_{FA} &= P(H_1|H_0) \\ P_D &= P(H_1|H_1) \end{aligned} \tag{3}$$

where $P(H_i|H_j)$ means the probability to decide for H_i while H_j was true. The Log-LRT results in the conventional detection scheme, by correlation.

Unfortunately, the received sequence is not always known by the receiver. This happens for instance in the GPS system. There are 32 satellites in the constellation, each of them transmitting its own sequence. As a result, the receiver does not know which sequence is received.

In order to solve this issue, the receiver may implement a GLRT strategy. This empirical approach has been defined especially when some parameters are unknown to the receiver. The basic principle is to estimate an unknown parameter θ according to a Maximum Likelihood criteria (ML), and then to apply the LRT with this estimate [12]:

$$\begin{aligned} &L(Y) = \frac{p(\mathbf{Y}|\hat{\theta}_1, H_1)}{p(\mathbf{Y}|\hat{\theta}_0, H_0)} \\ &\text{where } \hat{\theta}_i = \text{argmax } p(\mathbf{Y}|\hat{\theta}_i, H_i) \; i = 0, 1 \end{aligned} \tag{4}$$

If parameter θ does not appear in the model corresponding to hypothesis H_0, the GLRT becomes:

$$\begin{aligned} L(Y) &= \max_{\theta} L(Y, \theta) \\ L(Y, \theta) &= \frac{p(\mathbf{Y}|\theta, H_1)}{p(\mathbf{Y}|H_0)} \end{aligned} \tag{5}$$

In our context, the unkbnown parameter θ is the transmitted sequence $\mathbf{S}$. Applying the GLRT to the sequence estimation problem with $\theta = \mathbf{S}$ corresponds to the Maximum Likelihood Sequence Estimation algorithm (MLSE)[13]:

$$\hat{\mathbf{S}} = \max_{\mathbf{S} \in \mathbf{A}} \Re(\mathbf{Y}^H.\mathbf{S}) \tag{6}$$

$\mathbf{A} = \{S_0, \cdots, S_{N_s-1}\}$ is the search space, constituted of N_s possibles m-sequences. The algorithm selects the sequence which gives the largest correlation peak, and then compare it to the pre-defined detection threshold γ (2). If N_s is not too large, it is feasible to implement this kind of processing. This is the strategy adopted for the GPS receivers ($N_s = 32$). On the other hand, if N_s is too large, this solution is not feasible. One could implement for instance a Viterbi decoder that will estimate the ML received sequence. As the decoder complexity increases exponentially with the number of states, it can be implemented only if this number is not too large. This is generally not the case for the detection of m-sequences in practical situations [7][9][14].

An alternative solution is to perform a Maximum A Posteriori (MAP) symbol decoding instead of MLSE:

$$\hat{s}(i) = \max_{s(i)} p(s(i)|\mathbf{Y}, H_1) \tag{7}$$

This approach is very well adapted for the decoding of m-sequences, as explained in section 4. If the decoder finds a valid m-sequence, it may either be the transmitted sequence (correct decoding) or a delayed version of this sequence (wrong decoding).

If the probability of wrong decoding and the probability of false alarm are negligible, the decoding step not only gives the initial state of the transmitted sequence but also the synchronization with the beginning of the sequence. In this case, the verification step with the LRT of Eq. 2 is not needed, the probability of detection is very close to the probability of correct decoding. If these probabilities are not negligible, the verification phase is needed. As a consequence, the complexity of the proposed approach will depend on the performance of the iterative message-passing algorithm that will be selected to approximate the MAP decoder.

3. M-SEQUENCES

A m-sequence with r shift registers is built with a Linear Feedback Shift Register (LFSR) generator whose polynomial $g(D) = \sum_{i=0}^{r} g_i D^i$ is primitive [1]. Fig. 1 shows a LFSR sequence generator according to the Fibonacci representation [15]. A m-sequence is a periodic sequence of maximal period $N = 2^r - 1$. It has many interesting mathematical properties which are detailed in [1]. In this paper, we will exploit the decimation property. Any m-sequence of length N, $s(k)$ can be found by a suitable decimation of any other m-sequence of the same length N, $x(k)$. In other words, there exists an integer d such that ([1, theorem 10.2]):

$$s(k) = x(dk \mod N)$$

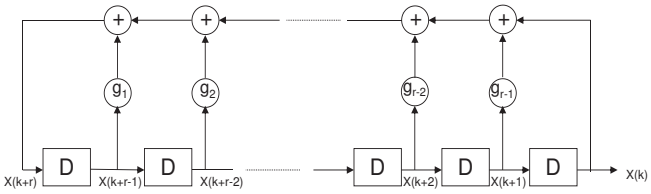

Fig. 1. LFSR sequence with the Fibonacci representation

This property will be exploited to improve the performance of the sequence decoder, as detailed in section 5

4. DECODING OF A M-SEQUENCE

The objective of the decoder is to find the initial state of the shift registers. Once it has been found, it is possible to generate the m-sequence with the architecture presented in Fig. 1. A m-sequence $x(k)$ satisfies the following parity check equation ($g_0 = g_r = 1$), for all $k \geq 0$:

$$\bigoplus_{i=0}^{r} g_{r-i} x(k+i) = 0$$

It is thus a cyclic linear code with coding rate $\frac{r}{N}$. A codeword of length N is generated by specifying an initial state of the shift registers. The parity check matrix of this code depends on the sequence's primitive polynomial $g(D)$:

$$H = \begin{bmatrix} g_r & \cdots & g_0 & 0 & \cdots & \cdots & 0 \\ 0 & g_r & \cdots & g_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & g_r & \cdots & g_0 & 0 \\ 0 & \cdots & \cdots & 0 & g_r & \cdots & g_0 \end{bmatrix} \quad (8)$$

Once the parity check matrix has been defined, it is possible to decode the received vector $(Y(0), \cdots, Y(N-1))$ with a standard iterative message passing algorithm [10]. In addition, as it was proposed in [16], the use of Redundant Graphical Model (RGM) improves significantly the decoder performance. If $g(D)$ is the sequence polynomial in GF(2), it satisfies:

$$g(D^{2^n}) = g(D)^{2^n}$$

This property is exploited to create additional parity check equations. Polynomial $g_n(D) = g(D^{2^n})$ also generates a parity check matrix H_n similar to H. Theses matrices can be concatenated to create a larger parity check matrix H_{RGM} [16]:

$$H_{\text{RGM}} = \begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_{n_{\text{RGM}}-1} \end{bmatrix} \quad (9)$$

where n_{RGM} is the number of RGMs used for decoding. These RGMs increase the column weight of the parity check matrix, while keeping constant the row weight. The column weight corresponds to the degree of the variable node in the bipartite graph [10]. Having a large degree improves the probability to correct an error on that variable because it receives more informations from its neighboring nodes. This explains the performance gain observed with the RGMs.

If the decoding is successful (all parity check equations are satisfied), the soft decision output of the decoder is converted into a binary representation with a hard decision rule. Then, according to the Fibonacci representation (Fig. 1), the first r bits of the codeword represents the content of the shift registers at initialization.

5. EXPLOITATION OF THE DECIMATION PROPERTY

It is well established in the Low Density Parity Check (LDPC) codes litterature that decoding performances degrade when the weight of the parity-check equations increases [17]. If the number of variable nodes linked to a check node is large, the information provided by each variable is diluted and this reduces its impact on the decoding procedure. Even worst, cycles of length 4 may appear, which are known to degrades seriously the decoding performances. It is thus a good design strategy to decrease the weight of the check nodes. In our context, this weight is equal to the weight of the m-sequence generator's polynomial (i.e. the number of non-zero coefficients). The proposed algorithm gives a method to decode any m-sequence of polynomial degree r, with the generator polynomial of the m-sequence having the smallest weight. This ensures to improve the decoding performances.

Let $x(k)$ be the m-sequence of degree r whose generator's polynomial $g_x(D)$ has the smallest weight. As stated in Section 3, any m-sequence $s(k)$ of degree r can be obtained by a suitable decimation of $x(k)$. If a specific initial state is chosen as a reference for each sequence $x(k)$ and $s(k)$, there exist 2 integers d and h such that [1]:

$$\begin{aligned} \phi(k) &= (dk + h) \bmod N \\ s(k) &= x(\phi(k)) \end{aligned} \quad (10)$$

h depends on the initial states chosen as a reference.

The basic decoding principles are to build from (1) a vector representing an observation of sequence $x(k)$, to decode $x(k)$ with a message-passing algorithm and, eventually, to compute sequence $s(k)$ according to (10).

The decoding steps are the followings :

- Step 1 : transform the observation of sequence $s(k)$ in (1) into the observation of sequence $x(k)$:

$$Y_d(k) = Y(\phi^{-1}(k)) = X(k) + n_d(k) \quad (11)$$

$n_d(k) = n(\phi^{-1}(k))$ is a simple permutation of the noise samples. Hence it has the same statistical properties.

- Step 2 : decode sequence $x(k)$ with a message-passing algorithm. The decoding algorithm described in section 4 is applied with polynomial $g_x(D)$ and input vector $Y_d(k)$ for $k = 0, \ldots, N-1$.
- Step 3 : compute the initial state of sequence $s(k)$:

$$s(k) = x(\phi(k)) \text{ for } k = 0, \ldots, r-1$$

6. SIMULATION RESULTS

The performance of the algorithm are measured by the probabilities of correct detection P_D, wrong detection P_{WD}, false alarm P_{FA} and missed detection P_M, defined as follows:

$$\begin{aligned} P_D &= P(Ic = 1 \text{ and } \hat{Q}_s = Q_s | H_1) \\ P_{WD} &= P(Ic = 1 \text{ and } \hat{Q}_s \neq Q_s | H_1) \\ P_{FA} &= P(Ic = 1 | H_0) \\ P_M &= 1 - P_D - P_{WD} \end{aligned}$$

Ic is the indication function of the decoder:

$$Ic = \begin{cases} 1 \text{ if all parity check equations are satisfied} \\ 0 \text{ otherwise} \end{cases}$$

$\hat{Q}_s$ is the estimated initial state vector of sequence $s(k)$, given by the decoder output.

Simulations have been performed for m-sequences of length $N = 1023$, with $r = 10$ registers. In this case, the sequence $x(k)$ generated by the polynomial $g_x(D) = D^{10} + D^3 + 1$ has the smallest weight. It is assumed that the initial state of each sequence ($x(k)$ or $s(k)$) is the all '1' configuration. Table 1 lists the sequences that have been used to evaluate the performance of the decoding procedure. One sequence has been selected for each possible weight ($5, 7$ or 9). For each sequence, the generator polynomial (taken from [18]) and the decimation parameters (d, h) are also given.

Performances are measured with the following simulation assumptions:

- When measuring P_M and P_{WD}, the receiver is synchronized with the beginning of the frame (i.e. hypothesis H_1 is satisfied), while its is not synchronized when P_{FA} is evaluated. 10^4 trials are used to measure P_M, and 10^6 for P_{WD}.
- the input noise is AWGN with variance σ^2. The signal to noise ratio is defined by : SNR $= 1/\sigma^2$
- The decoder implements either a Min-Sum (MS) or a Self-Corrected Min-Sum (SCMS) message-passing algorithm [19][20]. The SCMS performs very close to the "optimal" Belief propagation (Sum-Product) decoding with a reduced implementation complexity. In addition, the decoder does not need to have any knowledge about the noise variance as it is required for the Sum-Product algorithm. The decoder stops when either all the parity-check equations are satisfied or the maximum number of iteration $N_{iter} = 60$ is reached. The number of RGMs is $n_{\text{RGM}} = 7$.

s	weight	g(D)	(d,h)
s_1	5	$g_{s_2}(D) = D^{10} + D^8 + D^4 + D^3 + 1$	(43, 36980)
s_2	7	$g_{s_4}(D) = D^{10} + D^9 + D^8 + D^6 + D^3 + D^2 + 1$	(65, 27300)
s_3	9	$g_{s_5}(D) = D^{10} + D^9 + D^7 + D^6 + D^4 + D^3 + D^2 + D + 1$	(173,9342)

Table 1. Sequences

Two algorithms have been evaluated : the "conventional" one for which the parity check matrix is built with the generator polynomial of the considered sequence $g_s(D)$ (see section 4) and the novel algorithm described in the previous section. The performance of the former algorithm will be denoted for instance "Seq S1", and the latter "Seq S1 decim".

Fig. 2 compares the probability of missed detection P_M obtained with the MS and SCMS decoding algorithms. It is observed that the novel algorithm outperforms the conventional one by 9 to 13 dB, which is very significant. It can also be observed that the SCMS algorithm outperforms by almost 3dB the MS when the weight of the generator polynomial is large (9 for S_3). With the novel algorithm, the generator polynomial used for decoding is $g_x(D)$ which weight equals 3. In this case, the difference between the MS and SCMS is around 1dB.

Fig. 3 shows the probability of wrong decoding with the novel algorithm for the MS or SCMS decoding procedures. The SCMS gives a high probability of wrong detection, which makes it unsuitable for detection purposes. Simulations not reported in this paper also show that the probability of false alarm is high with the SCMS (10^{-3}) while it remains below 10^{-5} with the MS. Hence, decoding a m-sequence with the MS algorithm is the best choice.

Fig. 4 presents the probability of missed detection P_M for the 3 selected sequences defined in Table 1, and a MS decoding algorithm. It shows that while P_M depends on the weight of each sequence's polynomial for the conventional method, it is independant with the novel algorithm. In addition, the performance improvement is in the range of 6 to 12 dB. There is however a degradation of 2 dB with respect to the detection by correlation. For this simulation, the detection threshold was set to keep $P_{FA} < 10^{-5}$.

7. CONCLUSION

The relationship between the conventional GLRT detection method and the detection by iterative message-passing decod-

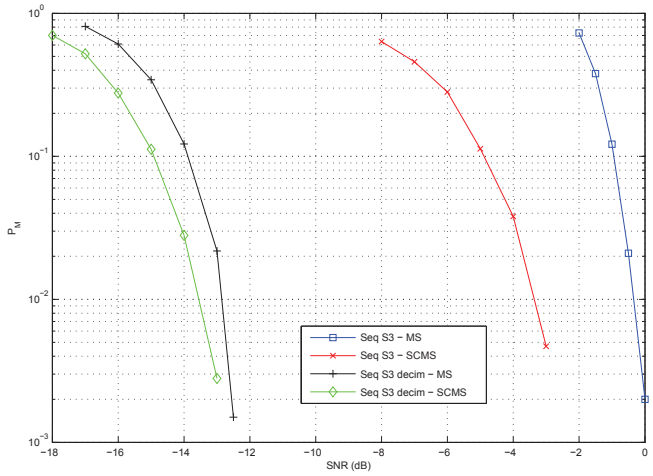

Fig. 2. P_M : MS versus SCMS - Sequence s_3

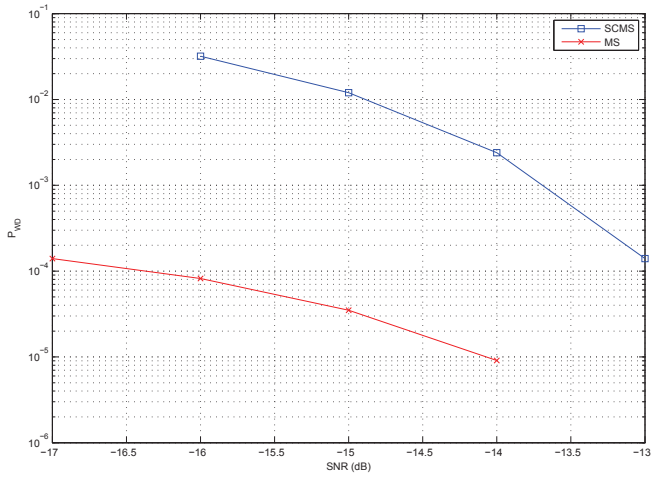

Fig. 3. P_{WD}: MS versus SCMS - Sequence s_3

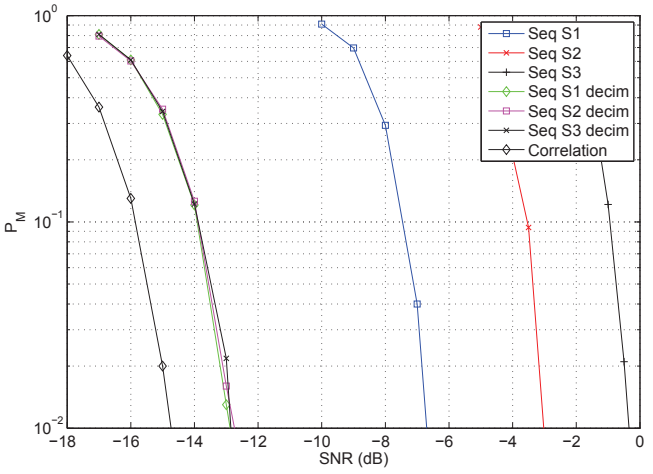

Fig. 4. Probability of missed detection - MS

ing has been established for the acquisition of m-sequences. Based on this framework, a novel detection algorithm has been proposed and evaluated. It exploits the decimation property between m-sequences of the same length, to improve the probability of detection for all the m-sequence which generator polynomial does not have the smallest weight. Simulation results show that this algorithm improves significantly the probability of detection. It was also shown that the selection of the iterative message-passing algorithm has a decisive impact on the probability of wrong decoding and false alarm.

8. REFERENCES

[1] R.J. McEliece, *Finite fields for computer scientists and engineers*, Springer, 1987.

[2] *TS25.213 v.4.4.0 Spreading and modulation (FDD)*, 3GPP, 2004.

[3] E.D. Kaplan and C.J. Hegarty, *Understanding GPS: principles and applications*, Artech House Publishers, 2006.

[4] A. Polydoros and C. Weber, "A unified approach to serial search spread-spectrum code acquisition–part I: general theory," *IEEE Trans. on Communications*, vol. 32, no. 5, pp. 542–549, 1984.

[5] D. Akopian, "Fast FFT based GPS satellite acquisition methods," in *IEE Proceedings on Radar, Sonar and Navigation*. IET, 2005, vol. 152, pp. 277–286.

[6] W. Meier and O. Staffelbach, "Fast correlation attacks on certain stream ciphers," *Journal of Cryptology*, vol. 1, no. 3, pp. 159–176, 1989.

[7] A. Canteaut and M. Trabbia, "Improved fast correlation attacks using parity-check equations of weight 4 and 5," in *Advances in Cryptology - EUROCRYPT 2000*. Springer, 2000, pp. 573–588.

[8] K.M. Chugg and M. Zhu, "A new approach to rapid PN code acquisition using iterative message passing techniques," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 5, pp. 884–897, 2005.

[9] F. Principe, K.M. Chugg, and M. Luise, "Performance evaluation of message-passing-based algorithms for fast acquisition of spreading codes with application to satellite positioning," in *NAVITEC, Noordwijk, The Netherlands*, December 2006.

[10] F.R. Kschischang, B.J. Frey, and H.A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[11] M. des Noes, V. Savin, Ros L., and Brossier J.M., "Blind identification of the uplink scrambling code index of a WCDMA transmission and application to femtocell networks," in *IEEE International Conference on Communications, Budapest, Hungary*, 2013.

[12] SM Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2, Prentice Hall, 1998.

[13] J.G. Proakis, *Digital communications*, vol. 1221, McGraw-hill, 1987.

[14] M. des Noes, V. Savin, Ros L., and Brossier J.M., "Blind identification of the scrambling code of a reverse link CDMA 2000 transmission," in *IEEE International Conference on Communications Budapest, Hungary*, 2013.

[15] R.L. Peterson, R.E. Ziemer, and D.E. Borth, *Introduction to spread-spectrum communications*, Prentice Hall, 1995.

[16] O.W. Yeung and K.M. Chugg, "An iterative algorithm and low complexity hardware architecture for fast acquisition of long PN codes in UWB systems," *The Journal of VLSI Signal Processing*, vol. 43, no. 1, pp. 25–42, 2006.

[17] Nandakishore Santhi and Alexander Vardy, "On the effect of parity-check weights in iterative decoding," in *Proceedings of the International Symposium on Information Theory (ISIT)*. IEEE, 2004.

[18] W Wesley Peterson and Edward J Weldon, *Error-correcting codes, Revised*, MIT press, 1972.

[19] N. Wiberg, *Codes and decoding on general graphs*, Ph.D. dissertation, Linköping University, Sweden, 1996.

[20] V. Savin, "Self-corrected min-sum decoding of LDPC codes," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, 2008, pp. 146–150.