Blind Estimation of Long and Short Pseudo-random Codes in Multi-rate LSC-DS-CDMA Signals

Fangfang Qiang, Zhijin Zhao, Xiaowei Gu and Xianyang Jiang

***Abstract*—**Aiming at the problem of blind estimation of pseudo-noise codes in multi-rate long and short codes direct sequence code division multiple access signal, in this paper, a novel codes estimation method based on Fast-ICA algorithm and sequences properties is proposed in this paper. The received signal is firstly segmented twice according to its maximum long scrambling code period and minimum spreading code period. Each user's composite code fragments consisting of long and short codes are separated by Fast-ICA algorithm. Then the user's long and short codes are estimated in descending order of the data rate, and the specific steps are as follows. Firstly, the separated composite code fragments make up the fuzzy sequences, and the double delay-and-multiply method is used to eliminate the order fuzzy and spread code interference. Secondly, combined with cyclotomic cosets and triple correlation properties, the method of feature information matching is used to estimate the long scrambling codes of all users with the same data rate. Meanwhile, the short spread codes are estimated by correlation operation. Lastly, all the estimated users' composite code fragments are deleted by using similarity matrix, fuzzy sequences are reconstructed, and three steps above are repeated until all users' long and short codes are estimated. Simulation results show the effectiveness of the proposed method.

***Keywords*—**Multi-rate long and short codes direct sequence code division multiple access signal, blind source separation, feature information matching, triple correlation

I. Introduction

DUE to the strong abilities of anti-interference and concealment as well as the excellent capacity for code division multiple access (CDMA), direct sequence spread spectrum (DSSS) signals are widely used in civil and military communication systems. With the rapid development of spread spectrum communication, the traditional DSSS technology has limited its application in high-speed data transmission due to the conflict between processing gain and data rate. In order to provide better service and application, multi-rate DSSS technology has introduced in spread spectrum communication [1-2]. DS-CDMA system is based on DSSS and CDMA technology, in non-cooperative communication, it is significant to blind estimate the Pseudo-Random (PN) codes in long and short codes direct sequence spread spectrum code division multiple access (LSC-DS-CDMA) signals.

Typical PN code estimation methods include eigenvalue decomposition [3-6], blind source separation [7-8], subspace based methods [9-10], EM methods [11] and so on. These methods have a certain effect on simple DSSS signals, but they can't be directly applied to multi-rate DS-CDMA signals. Reference [12] combined Frobenius square norm and eigenvalue decomposition method to estimate the spreading sequence of asynchronous multi-rate DS-CDMA signal. Reference [13] used multi-rate sampling method to estimate multi-rate CDMA signals' code chip rate. Aiming at the different characteristics of asynchronous and synchronous multi-rate DS-CDMA signals, different cyclic segmentation methods to reduce dimension are proposed in [14]. Based on [14], [15] proposed a parallel PN code estimation method combined with JADE algorithm for multi-rate DS-CDMA signal.

However, the methods above mainly aim at the short code direct sequence spread spectrum code division multiple access (SC-DS-CDMA) signals, and are not suitable for the multi-rate LSC-DS-CDMA signals with complex signal structure. At present, the research of PN code mainly focuses on short code or long code DSSS signal with a single rate, and there is less published research output regarding the blind estimation technique for estimating the PN codes of multi-rate LSC-DS-CDMA signals.

Due to solve this problem, in this paper, a new method based on Fast-ICA algorithm and properties of m-sequence to blind estimate the PN codes in multi-rate LSC-DS-CDMA signal is proposed. The received signal is firstly segmented twice according to the maximum long scrambling code period and the minimum spreading code period. Each user in signal is independent, then its composite code segments can be separated by Fast-ICA algorithm. Combined with the triple correlation theory, the feature information matching method is used to estimate each user's long scrambling code. Finally, according to

This work was supported by the National Natural Science Foundation of China(Grant No.61571172).

Fangfang Qiang is with the School of Electronic and Information, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China (E-mail: qiangffmy@163.com).

Zhijin Zhao is with the School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China.

Xiaowei Gu is with the School of Computing and Communications, Lancaster University, Lancaster, UK.

Xianyang Jiang is with the School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China.

the composite code and estimated long scrambling code to obtain the short spreading code.

II. Signal Model

Consider a multi-rate LSC-DS-CDMA signal with U users. Total data rates number is V, and the data rates are expressed as $R_1, R_2, \cdots R_V$, the corresponding spreading codes are $Lb_1, Lb_2, \cdots Lb_V$ ($Lb_i = 1/R_i$), meantime, $Lb_V = 2Lb_{V-1} =, \cdots, = 2^{V-1}Lb_1$. U_i is the user number with rate R_i, $U_1 + U_2 + \cdots + U_V = U$. After chip rate sampling, the received signal can be expressed as $r(n)$ [15]:

$$\begin{aligned} r(n) &= \sum_{i=1}^{V}\sum_{j=1}^{U_i} A_{i,j} d_{i,j}(n) s_{i,j}(n) + w(n) \\ &= \sum_{i=1}^{V}\sum_{j=1}^{U_i} A_{i,j} d_{i,j}(n) b_{i,j}(n) c_{i.j}(n) + w(n) \end{aligned} \tag{1}$$

where n is sampling instance and $1 \le n \le N$, N is the signal length, $A_{i.j}$, $d_{i,j}(n)$, $b_{i,j}(n)$ and $c_{i,j}(n)$ is respectively the signal amplitude, data code, spreading code and scrambling code of *j-th* user with rate R_i, $s_{i.j}(n) = b_{i,j}(n)c_{i,j}(n)$ is the composite code consisting of spreading code and scrambling code, $w(n)$ is additive white Gaussian noise, $w(n) \sim N(0,\sigma^2)$.

For each user used i *-th* rate, OVSF code with Lb_i chips is used as spreading code, m-sequence is used as scrambling code.

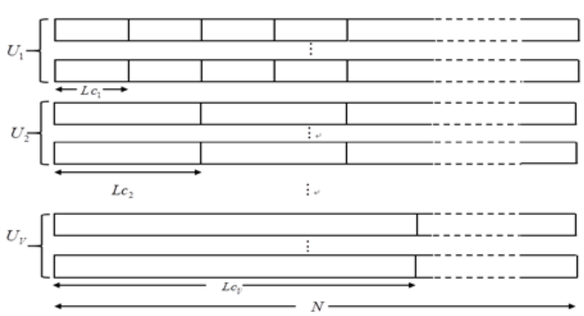

Fig. 1 Signal structure diagram

Since the period of m-sequence is odd, the first chip of m-sequence can be added to its end and then an obtained new period is Lc_i. Users' short spreading code and long scrambling code are different, but their periods of short code and long code are the same if they have the same data rate. In addition, $Lc_V = 2Lc_{V-1} \cdots ,= 2^{V-1}Lc_1$. Signal structure is shown in Fig.1.

Other assumptions in this paper are as follows:

(1) The spreading code period and scrambling code period are both known or have been obtained according to the period estimation method proposed in [15]. The user number is known.

(2) All the date sequences are statistically independent.

(3) Each PN sequence in the signal is statistically independent from others and has zero mean.

(4) Z ($Z = N / Lc_V$) is greater than user number U.

Since the received signal is a multi-rate LSC-DC-CDMA signal contained Z completed periods of long scrambling code, the received signal can be divided by the maximal period length of scrambling code Lc_V and converted into array signal consisting of Z signal segments. Every segment can be viewed as a signal received by an array element. Then formula (1) can be expressed as matrix form as follows:

$$\boldsymbol{r}(n) = \boldsymbol{A}(n)\boldsymbol{S}(n) + \boldsymbol{W}(n) \tag{2}$$

Where, $1 \le n \le Lc_V$, $\boldsymbol{r}(n) = [r_1(n)\ \ r_2(n) \cdots r_Z(n)]^{\mathrm{T}}$, $r_z(n) = r((z-1)\cdot Lc_V + n)$, $\boldsymbol{S}(n) = [s_{1,1}(n) \cdots s_{i,j}(n)\ \ \cdots s_{V,U_V}(n)]^{\mathrm{T}}$, $1 \le j \le U_i$, $\boldsymbol{W}(n) = [w_1(n)\ \ w_2(n) \cdots w_Z(n)]^{\mathrm{T}}$, $w_z(n) = w((z-1)\cdot Lc_V + n)$.

$$\boldsymbol{A}(n) = \begin{bmatrix} A_{1,1}d_{1,1}(e_{1,1}(n)) \cdots A_{i,j}d_{i,j}(e_{i,1}(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,1}(n)) \\ A_{1,1}d_{1,1}(e_{1,2}(n)) \cdots A_{i,j}d_{i,j}(e_{i,2}(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,2}(n)) \\ \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\ A_{1,1}d_{1,1}(e_{1,Z}(n)) \cdots A_{i,j}d_{i,j}(e_{i,Z}(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,Z}(n)) \end{bmatrix} \tag{3}$$

Where $e_{i,z}(n) = \left\lceil \dfrac{(z-1)\cdot Lc_V + n}{Lb_i} \right\rceil$. Because there are multiple data symbols in every signal segment, the mixing matrix $\boldsymbol{A}(n)$ is time-varying when $1 \le n \le Lc_V$. Fast-ICA algorithm can't be directly used to estimate the composite code.

According to the minimum spreading period Lb_1, each signal segment is divided into K fragments, then the k *-th* fragment can be expressed as:

$$\boldsymbol{r}^k(n) = \boldsymbol{A}^k(n)\boldsymbol{S}^k(n) + \boldsymbol{W}^k(n) \tag{4}$$

where $1 \le n \le Lb_1$, $1 \le k \le K$, $\boldsymbol{r}^k(n) = [r_1^k(n)\ \ r_2^k(n) \cdots\ r_Z^k(n)]^{\mathrm{T}}$, $r_z^k(n) = r_z((k-1)\cdot Lb_1 + n)$, $s_{i,j}^k(n) = s_{i,j}((k-1)\cdot Lb_1 + n)$, $\boldsymbol{W}^k(n) = [w_1^k(n)\ \ w_2^k(n) \cdots\ w_Z^k(n)]^{\mathrm{T}}$, $w_z^k(n) = w_z((k-1)\cdot Lb_1 + n)$, $\boldsymbol{A}^k(n)$ is a mixing matrix, which is

$$\boldsymbol{A}^k(n) = \begin{bmatrix} A_{1,1}d_{1,1}(e_{1,1}^k(n)) \cdots A_{i,j}d_{i,j}(e_{i,1}^k(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,1}^k(n)) \\ A_{1,1}d_{1,1}(e_{1,2}^k(n)) \cdots A_{i,j}d_{i,j}(e_{i,2}^k(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,2}^k(n)) \\ \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\ A_{1,1}d_{1,1}(e_{1,Z}^k(n)) \cdots A_{i,j}d_{i,j}(e_{i,Z}^k(n)) \cdots A_{V,U_V}d_{V,U_V}(e_{V,Z}^k(n)) \end{bmatrix} \tag{5}$$

where $e_{i,z}^k(n) = e_{i,z}((k-1)\cdot Lb_1 + n)$. So $\boldsymbol{A}^k(n)$ is time-invariant. Formula (4) is a typical instantaneous liner mixed signal model. In order to estimate each user's composite code fragments, the Fast-ICA algorithm can be used to blindly separate the received signal fragment $\boldsymbol{r}^k(n)$ and then to gain the $\boldsymbol{S}^k(n)$.

III. Blind Estimation of Long And Short PN Codes

A. Blind separation of composite code fragments based on Fast-ICA algorithm

We calculate the covariance matrix of the observed signal then apply eigenvalue decomposition [16].

$$\begin{aligned} \boldsymbol{R}^k &= E[\boldsymbol{r}^k \cdot (\boldsymbol{r}^k)^{\mathrm{H}}] \\ &= [\boldsymbol{\psi}_s \boldsymbol{\psi}_w] \begin{bmatrix} \boldsymbol{\Lambda}_s & 0 \\ 0 & \boldsymbol{\Lambda}_w \end{bmatrix} [\boldsymbol{\psi}_s \boldsymbol{\psi}_w]^{\mathrm{H}} \end{aligned} \tag{6}$$

where $\boldsymbol{\psi}_s$ and $\boldsymbol{\psi}_w$ is respectively the signal subspace and the noise subspace, $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_w$ is respectively the corresponding eigenvalue diagonal matrices. The signal fragment $\boldsymbol{r}^k$ is whitened, and the whitened matrix $\boldsymbol{\Gamma}^k$ is as follows:

$$\boldsymbol{\Gamma}^k = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{\psi}_s^{\mathrm{H}}\boldsymbol{r}^k \tag{7}$$

The dimension of $\boldsymbol{\Gamma}^k$ is $U \times Lb_1$. Then the separation matrix $\boldsymbol{Y}^k = [\boldsymbol{y}_{1,1}^k \ \cdots \ \boldsymbol{y}_{i,j}^k \cdots \ \boldsymbol{y}_{V,U_V}^k]^{\mathrm{T}}$ ($1 \le i \le V$, $1 \le j \le U_i$) is estimated by setting an intimal value for every $\boldsymbol{y}_{i,j}^k$ and iterating (8) till converging,

$$\begin{gathered}\boldsymbol{y}_{i,j}^k = \boldsymbol{y}_{i,j}^k - \frac{\mu\{E[\boldsymbol{\Gamma}^k g((\boldsymbol{y}_{i,j}^k)^T\boldsymbol{\Gamma}^k)] - \varepsilon \cdot \boldsymbol{y}_{i,j}^k\}}{\{E[g'((\boldsymbol{y}_{i,j}^k)^T\boldsymbol{\Gamma}^k)] - \varepsilon\}} \\ \boldsymbol{y}_{i,j}^k = \frac{\boldsymbol{y}_{i,j}^k}{\left\|\boldsymbol{y}_{i,j}^k\right\|}\end{gathered} \tag{8}$$

where, $\varepsilon = E[(\boldsymbol{y}_{i,j}^k)^{\mathrm{T}}\boldsymbol{\Gamma}^k g((\boldsymbol{y}_{i,j}^k)^T\boldsymbol{\Gamma}^k)]$, $g'(\cdot)$ is the derivative of nonlinear function $g(\cdot)$, $g(x) = x^3$ is selected. Then $\hat{\boldsymbol{s}}_{i,j}^k$ obtained through the Fast-ICA method is expressed as:

$$\hat{\boldsymbol{s}}_{i,j}^k = (\boldsymbol{y}_{i,j}^k)^{\mathrm{T}}\boldsymbol{\Gamma}^k = [\hat{s}_{i,j}^k(1) \ \ \hat{s}_{i,j}^k(2) \ \cdots \ \hat{s}_{i,j}^k(Lb_1)] \tag{9}$$

As we can see in (9), the k *-th* fragment of the j *-th* user's mixing PN sequence corresponding to rate R_i is separated from the multi-rate signal. In order to separate U different fragments, which are corresponding to U users, we can set U different initial values $\boldsymbol{y}_{i,j}^k$. To make sure that each extracted fragment is unique and not obtained before, the extracted component must be removed before the next extraction, and it can be achieved by Schmidt normalization, which is

$$\begin{aligned}\boldsymbol{y}_{i,j}^k &= \boldsymbol{y}_{i,j}^k - \sum_{t_1=1}^{i}\sum_{t_2=1}^{j-1}\left\langle \boldsymbol{y}_{i,j}^k, \boldsymbol{y}_{t_1,t_2}^k \right\rangle \cdot \boldsymbol{y}_{t_1,t_2}^k \\ \boldsymbol{y}_{i,j}^k &= \boldsymbol{y}_{i,j}^k / \left\|\boldsymbol{y}_{i,j}^k\right\|\end{aligned} \tag{10}$$

where $1 \le i \le V$, $2 \le j \le U_i$. After U times separation all users' composite code fragments $\hat{\boldsymbol{s}}_{i,j}^k$ can be obtained, which include all users' spreading codes and scrambling codes.

B. Blind estimation of long and short PN codes based on triple correlation theory

The blind separation of the composite code fragments of U users is completed in Section *A*, but the problems of uncertainty in order and amplitude are not solved. Moreover, the spreading code and long scrambling code have not been separated. In this section, the triple correlation theory is used to solve these problems.

Shift superposition characteristics and periodicity are important characteristics of m-sequences. We use these characteristics to estimate the m-sequence from the composite code. Triple correlation function (TCF) is defined as[17-18]:

$$\begin{aligned}C(p,q) &= \mathrm{E}[m(n)m(n+p)m(n+q)] \\ &= \begin{cases} 1 & , \ m(i+p)m(i+q) = m(i) \\ -1/L & , \ m(i+p)m(i+q) \ne m(i)\end{cases}\end{aligned} \tag{11}$$

where $1 \le p,q \le L$, L is the period length of m-sequence. The peak coordinates of the triple correlation function are unique for the m-sequence obtained by different primitive polynomials with the same order. Thus, the peak coordinates can be regarded as feature information of the m-sequence and used to estimate the m-sequence.

When assembling $\hat{\boldsymbol{s}}_{i,j}^k$ in order, there are U^K fuzzy sequences with the length of Lc_V , and then the complexity of estimating all users' long and short PN codes directly from the whole fuzzy sequences will be very high. We adopt the method of estimation according to different rates in order. The PN codes of U_1 users with R_1 rate are estimated firstly. Specific steps are as follows.

Step1: TCF estimation of delayed sequences

A sequence with the length of Lc_V contains 2^{V-1} complete sequences with period Lc_1 . We can select $K_1 = Lc_1 / Lb_1$ fragments in order for assembling to estimate the m-sequences with period Lc_1. Fuzzy sequences $\boldsymbol{\alpha}_h$ can be expressed as

$$\boldsymbol{\alpha}_h = [\hat{\boldsymbol{s}}_{i,j}^1 \ \ \hat{\boldsymbol{s}}_{i,j}^2 \ \cdots \ \hat{\boldsymbol{s}}_{i,j}^k \cdots \ \hat{\boldsymbol{s}}_{i,j}^{K_1}] \tag{12}$$

where $1 \le h \le U^{K_1}$, $1 \le i \le V$, $1 \le j \le U_i$, $1 \le k \le K_1$. Only U_1 sequences of the $\boldsymbol{\alpha}_h$ are the real mixing PN sequences of the users with rate R_1 . Here we denote these special $\boldsymbol{\alpha}_h$ as $\boldsymbol{\alpha}_{h'}^*$ ($1 \le h' \le U_1$). Although the spreading codes and scrambling codes are included in $\boldsymbol{\alpha}_{h'}^*$, estimating the scrambling codes via their triple correlation properties is still impossible. As a result, the operation of double delay-and-multiply processing is used to eliminating the influence of order and spreading code, which is:

$$\begin{aligned}\boldsymbol{\beta}_{h'}^*(n) &= \boldsymbol{\alpha}_{h'}^*(n) \cdot \boldsymbol{\alpha}_{h'}^*((n+1) \bmod Lc_1) \\ \boldsymbol{\gamma}_{h'}^*(n) &= \boldsymbol{\beta}_{h'}^*(n) \cdot \boldsymbol{\beta}_{h'}^*((n+Lb_1) \bmod Lc_1)\end{aligned} \tag{13}$$

where $1 \le n \le Lc_1$. The TCF of $\boldsymbol{\gamma}_{h'}^*(n)$ is calculated as follows:

$$C_h(p,q) = \frac{1}{Lc_1}\sum_{i=1}^{Lc_1}\boldsymbol{\gamma}_{h'}^*(n)\boldsymbol{\gamma}_{h'}^*(n+p)\boldsymbol{\gamma}_{h'}^*(n+q) \tag{14}$$

All the feature information of scrambling codes is kept in $C_h(p,q)$. As the scrambling code in multi-rate signal is m-sequence, there must be some obvious peaks in $C_h(p,q)$. For other $U^{K_1} - U_1$ fuzzy sequences $\boldsymbol{\alpha}_h$, the triple correlation properties are destroyed after double delay-and-multiply processing, there must be no obvious peaks in $C_h(p,q)$.

Step2: scrambling code estimation based on feature information matching

Suppose there are J finite sets in the cyclotomic cosets of scrambling code. All of them are found out and the obtained coset heads are defined as the set $\{\eta_j \,|\, 1 \le j \le J\}$. The TCF values of m-sequences corresponding to all primitive polynomials (the number is F) with this order are calculated when the abscissa is range of $\{\eta_j \,|\, 1 \le j \le J\}$. The TCF peak points are found out, which constitute the m-sequence's feature information matrix $\boldsymbol{\Phi}$. The TCF values of $\gamma_h(n)$ at the peak points of F sequences are computed via (14). Then the average value is computed by the following formula:

$$\bar{C}_f = \frac{\sum_{j=1}^{J} C_h(\eta_j, q_{f,j})}{J} \tag{15}$$

where, f is the serial number of m-sequence and $1 \le f \le F$, $q_{f,j}$ represents the ordinate of peak point coordinate of the m-sequence corresponding to f -*th* primitive polynomial when the abscissa is η_j. Finding out the maximum of $\bar{C}_f$, the serial number f is recorded into set ξ. And U^{K_1} elements of the set ξ can be obtained by the U^{K_1} sequences. And then sorting the U^{K_1} elements by decreasing number of occurrences, the first U_1 digits are the sequence numbers in the feature information matrix, which are corresponding to U_1 users' scrambling codes.

Step 3: spreading code estimation

According to composite code $\boldsymbol{\alpha}^*_{h'}(n)$ and estimated scrambling code $\boldsymbol{c}_{h'}(n)$, spreading code can be obtained by the following operation.

$$\boldsymbol{b}_{h'}(n)=\boldsymbol{\alpha}^*_{h'}(n)\cdot \boldsymbol{c}_{h'}(n) \tag{16}$$

where $1\le n\le Lb_1$, $1\le h'\le U_1$. In this case, the obtained sequence $\boldsymbol{b}_{h'}(n)$ or its reversed sequence is the short spreading code. It can be judged according to the characteristics of the OVSF code, and then the short spreading code sequence is gotten.

All the long scrambling codes and short spreading codes of U_1 users with R_1 rate have been estimated then.

Step 4: clear estimated composite code

By using similarity coefficient matrix, all the fragments of U_1 users are removed from the composite code fragments, which are obtained by blind separation based on Fast-ICA method. The similarity coefficient is defined as

$$\delta^k_{\lambda,j'}=\frac{\sum_{n=1}^{Lb_1}\left|\hat{s}^k_{i,j}(n)+s'^{\,k}_{1,j'}(n)\right|}{2\cdot Lb_1} \tag{17}$$

where $\lambda=\sum_{t=1}^{i-1}U_t+j$, $1\le i\le V$, $1\le j\le U_i$, $1\le j'\le U_1$, $\hat{s}^k_{i,j}(n)$ is the k -*th* composite code fragment of j -*th* user with i -*th* rate obtained by blind separation. $\hat{s}'^{\,k}_{1,j'}(n)$ is the k -*th* composite code fragment of j' -*th* user with rate R_1 obtained from estimated sequences $\boldsymbol{\alpha}^*_{h'}(n)$. Then the similarity coefficient matrix corresponding to k -*th* fragment is

$$\boldsymbol{\delta}^k=\begin{bmatrix}\delta^k_{1,1} & \delta^k_{1,2} & \cdots & \delta^k_{1,U_1}\\ \delta^k_{2,1} & \delta^k_{2,2} & \cdots & \delta^k_{2,U_1}\\ \vdots & \vdots & \ddots & \vdots\\ \delta^k_{U,1} & \delta^k_{U,2} & \cdots & \delta^k_{U,U_1}\end{bmatrix} \tag{18}$$

The fragments $\hat{s}^k_{i,j}(n)$ and $\hat{s}'^{\,k}_{1,j'}(n)$ are the same when $\delta^k_{\lambda,j'}=1$. Due to the special property of m-sequence, the minimum value of $\delta^k_{\lambda,j'}$ is approximately equal to 0.5. Theoretically, every row in the matrix $\boldsymbol{\delta}^k$ has only one value equaling to 1, and the remaining values are approximately equal to 0.5.

Finding out the λ corresponding to the maximum of $\delta^k_{\lambda,j'}$ in j' -*th* column ($1\le j'\le U_1$), the λ -*th* row is removed from $\hat{\boldsymbol{S}}^k$ and the updated $\hat{\boldsymbol{S}}^k$ is obtained. After that, $\hat{\boldsymbol{S}}^k$ no longer contains composite code fragments belonging to users with rate R_1.

Restructuring $(U-\sum_{t=1}^{i-1}U_t)^{K_i}$ ($2\le i\le V, K_i=Lc_i/Lb_1$) possible fuzzy sequences $\boldsymbol{\alpha}_h(n)$ $(1\le h\le (U-\sum_{t=1}^{i-1}U_t)^{K_i})$, repeating the above steps, replacing U_1 with U_i, we can estimate the long and short PN codes of U_i users with rate R_i. Finally, all the PN codes can be obtained.

IV. Algorithm Simulation and Performance Analysis

Simulations are given to verify the effectiveness of the proposed method for the estimation of long and short PN codes of multi-rate LSC-DS-CDMA signal. Assuming that the LSC-DS-CDMA signal is synchronized, and each user has the same amplitude. Data codes are randomly generated from {±1}. The signal to noise ratio is defined as SNR = $10\lg(A^2/\sigma^2)$. Users are divided into three groups according to rate. For each user, we use the OVSF code as the spreading code (short code) and m-sequence as scrambling code (long code), the period of short code and long code is *Lb* and *Lc*, respectively. The rate, *Lb*, *Lc* and primitive polynomial of each user are shown in Table 1. The signal's user combination situations for algorithm simulations are shown in Table 2, which means that signal 1 contains 3 users with 1 rate (users 1 to 3 in Table 1), signal 2 contains 3 users with 2 rates (users 1 to 2 and user 4 in Table 1), and so on.

Table 1 User parameters

user	rate	*Lb*	*Lc*	primitive polynomial
1	R_1	64	255	$x^8+x^5+x^3+x^2+1$
2	R_1	64	255	$x^8+x^7+x^3+x^2+1$
3	R_1	64	255	$x^8+x^7+x^2+x+1$
4	R_2	128	511	x^9+x^4+1
5	R_2	128	511	$x^9+x^8+x^4+x+1$
6	R_3	256	1023	$x^{10}+x^3+1$
7	R_3	256	1023	$x^{10}+x^7+x^3+x^2+1$

Table 2 Signal's user combination

No.	rate	user combination
Signal 1	R_1	User 1、2、3
Signal 2	R_1、R_2	User 1、2、4
Signal 3	R_1、R_2	User 1、4、5
Signal 4	R_1、R_2、R_3	User 1、4、6
Signal 5	R_1、R_2	User 1、2、4、5
Signal 6	R_1、R_3	User 1、2、6、7
Signal 7	R_2、R_3	User 4、5、6、7

In this section, the influence of the number of rates in each signal on the estimation performance of composite codes, the influence of the number of rates in each signal and rate value of each user on long and short codes estimation performance, and the comparison of long and short codes estimation performance with references [7] and [14] is analyzed.

The bit error rate is used as the composite code estimation performance evaluation index, which is defined as the ratio of the number of error bits to the total number of the composite code bits. The correct estimation probability of long code primitive polynomial and short code sequence is used as the long and short codes estimation performance index, and assuming that the estimation of short code is correct when the bit error rate of short code is less than 3%. The curves in the following experiments are the average results of 100 Monte Carlo simulations.

A. The influence of rate number on composite code estimation performance

Signals 1 to 4 in Table 2 are selected as the multi-rate LSC-DS-CDAM signals, and 3 users are contained in each signal. Assuming that the signal length is $N = 100Lc_3$, the curves of average bit error rate of composite code under different SNR are shown in Fig.2.

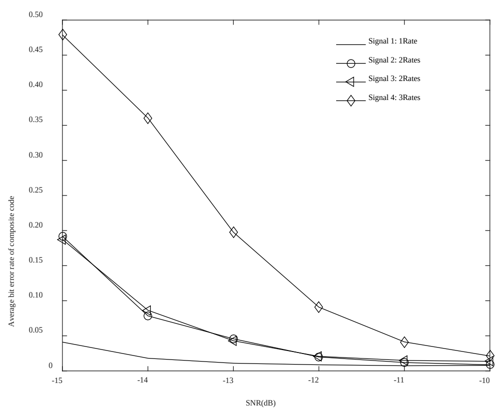

Fig.2 Influence of rate number on composite code estimation performance

Fig.2 shows that, the average bit error rate of the composite code increases with the increase of the number of signal rates. The more the number of rates, the worse the estimation performance of the composite code, which makes the performance of the long code and short code estimation worse eventually. This is because under the certain condition of the received signal length, the more the number of rates, the more complex the signal structure. As a result, the separation effect of composite code fragments becomes worse. In addition, when the total number of users in LC-DS-CDMA signal is the same and the number of rates is the same (signal 2 and signal 3), the estimation performance of composite codes has little difference.

B. The influence of rate number on long and short codes estimation performance

The selected signals to be simulated is the same as *Experiment A*, the curves of the correct estimation probability of long and short PN codes are shown in Fig.3.

As is shown in Fig.3, the more the number of rates in received signal, the worse the estimation performance of long and short PN codes. This is because according to *Experiment A*, the more the number of rates, the worse the separation performance of composite code, which result in the worse estimation performance of the long and short PN codes. For signals 2 and 3, both contains 3 users with two kinds of user rate, the estimation performance of composite codes is not significant, thus, the estimation performance of long and short codes has little difference.

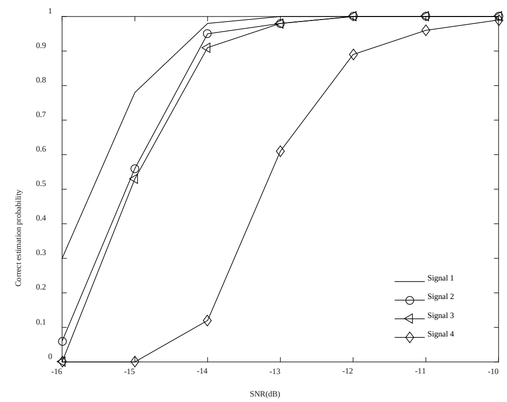

Fig.3 Influence of rate number on long and short codes estimation performance

C. The influence of rate value on long and short codes estimation performance

Signals 5 to 7 in Table 2 are selected as the multi-rate LSC-DS-CDAM signals. The signal length is $N = 100Lc_3$. The curves of the correct estimation probability of long and short PN codes are shown in Fig.4.

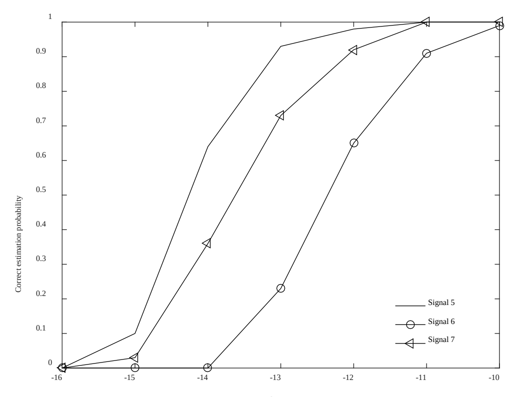

Fig.4 Influence of rate number on long and short codes estimation performance

As is shown in Fig.4, the estimation performance of long and short PN codes of signal 5 is the best when the received signal length is the same. This is because the minimum user rate R_2 in signal 5 is larger than the minimum user rate R_3 in signals 6 and 7, so the maximum long code period Lc_2 in signal 5 is smaller than the maximum long code period Lc_3 in signal 6 and signal 7, then the number of complete cycles in signal 5 is larger than that in signals 6 and 7, the performance of blind estimation based on Fast-ICA algorithm is much better. Both signal 6 and signal 7 contain two users with rate R_3, and other two users' rate is R_1 in signal 6 and R_2 in signal 7. The number of composite code cycles Z ($Z = N / Lc_3$) in signal 6 and signal 7 is the same, but the number of short code cycles in signal 6 is twice of that in signal 7. According to the signal segmentation method proposed in this paper, the number of composite code fragments obtained after segmenting for signal 6 is twice of that for signal 7. Therefore, the bit error rate increases after assembling the fragments obtained by blind separation, which eventually makes the estimation performance of long and short PN codes of signal 6 worse.

D. The estimation performance comparison

In order to further verify the performance of the proposed algorithm, we compared it with the method in [7] and [14]. Signal 7 is selected as the multi-rate LSC-DS-CDMA signal, which contains 4 active users. When the length of received signal is $N = 100Lc_3$ and $N = 200Lc_3$, respectively, the curves of the correct estimation probability of long and short PN codes

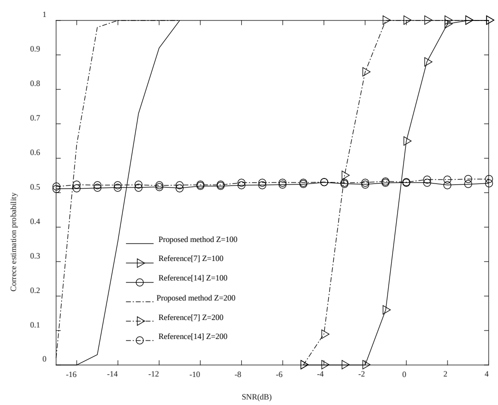

Fig.5 Algorithm performance comparison

are shown in Fig.5.

As is shown in Fig.5, we can get the conclusions that, (1) compared with the reference [7], the proposed algorithm improves about 13.3dB and 13.6dB, respectively, when the correct estimation probability is 90% at the signal length of $N = 100Lc_3$ and $N = 200Lc_3$. This is because matrix filling error exists in [7]'s method. In addition, TCF feature information matching method reduces the error probability. (2) when $N = 100Lc_3$ and the SNR is above -13.5dB, or $N = 200Lc_3$ and the SNR is above -16.2dB, the proposed method exhibits higher correct estimation probability than comparative method in [14]. Otherwise, the comparative method has the better estimation performance. This is because noise has a large effect on the third-order statistics at low SNR. As the chip assembling operation is used in the comparative method, there apparently will be more errors in the estimated sequences. Thus, the performance of comparative method is improved negligibly with higher SNR. Therefore, we can conclude that the proposed method in this paper has a better performance of PN code estimation.

V. CONCLUSION

In order to estimate the PN codes of multi-rate LSC-DS-CDMA signal, a new method combined with Fast-ICA algorithm, triple correlation theory and feature information matching method is proposed in this paper. The Fast-ICA algorithm is used to separate the composite code fragments of different users, which can reduce the interference between users when estimating each user's PN codes. The triple correlation function of m-sequence is insensitive to noise, so based on the triple correlation theory and feature information matching method, the PN codes estimation can still be effective at low SNR. The simulation results prove the proposed method can effectively estimate the PN codes of multi-rate LSC-DS-CDMA signal. In addition, the superiority of the proposed method is verified by comparing with the existing methods. However, this paper only considers the situation of signal synchronization, and using the feature matching method needs to construct the feature information matrix in advance, the calculation is complicated. The solution to these problems is the next research focus.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundations of China (Grant No.61571172) and Zhejiang Provincial Key Lab of Data Storage and Transmission Technology, Hangzhou Dianzi University.

REFERENCES

[1] S Buzzi, M Lops, and A Pauciullo. "Iterative cyclic subspace tracking for blind adaptive multiuser detection in multirate CDMA systems," IEEE Transactions on Vehicular Technology, vol.52, no.6, pp.1463-1475, 2003.

[2] C N Nzeza, R Gautier, G Burel. "Blind multiuser identification in multirate CDMA transmissions: a new approach," in Proceedings of the IEEE Fortieth Asilomar Conference on Signals, Systems and Computers, pp.2162-2166, 2006.

[3] Z Luo, L Zhu. "A charrelation matrix-based blind adaptive detector for DS-CDMA systems," Sensors, vol.15, no.8, pp. 20152-20168, 2015.

[4] G Burel, C Bouder. "Blind estimation of the pseudo-random sequence of a direct sequence spread spectrum signal," in Proceedings of the IEEE Century Military Communications, pp.967-970, 2000.

[5] T Q Zhang, Q Yang, Y L Song, and M Xiong. "Similarity based blind despread method of DS-SS signal," System Engineering and Electronics, vol.39, no.7, pp.1451-1456, 2017.

[6] T Q Zhang, S S Dai, W Zhang, and X Y Gao. "Blind estimation of the PN sequence in lower SNR DS-SS signals with residual carrier," Digital Signal Processing, vol.22, no.1,pp.106-113, 2012.

[7] Z J Zhao, M Li, J N Shang. "Blind estimation of LSC-DS-CDMA signal based on matrix completion and triple correlation," Journal of Electronics & Information Technology, vol.38, no.7, pp. 1788-1793, 2016.

[8] Z J Zhao, X W Gu, F F Qiang, and L Shen. "Blind estimation of PN codes in multi-user LSC-DSSS signals," Journal of Communications, vol.12, no.1, pp.55-61, 2017.

[9] R L Hou, T Q Zhang, T Pang, and X Y Gao. "Spread spectrum code estimation for DS-CDMA signal based on subspace," Application Research of Computers, vol.28, no.2, pp. 673-676, 2011.

[10] L Q Zhou, H Leung, P Xu, et al. "The Kalman filtering blind adaptive multi-user detector based on tracking algorithm of signal subspace," Information, vol.6, no.1,pp.3-13, 2015.

[11] Y W Yao, H V Poor. "EM-based blind demodulation of synchronous CDMA," in Proceedings of the IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications, pp. 738-742, 2000.

[12] C N Nzeza, R Gautier, and G Burel. "Parallel blind multiuser synchronization and sequences estimation in multirate CDMA transmissions," in Proceedings of the IEEE Fortieth Asilomar Conference on Signals, Systems and Computers, pp. 2157-2161, 2006.

[13] S Ghavami and B Abolhassani. "Blind chip rate estimation in multirate CDMA transmissions using multirate sampling at slow flat fading channels," in Proceedings of the IEEE Singapore International Conference on Communication Systems, pp.1344-1348, 2009.

[14] X T Ren, H Xu, Z T Huang, F H Wang, and F B Lu. "Blind source separation based blind estimation of the spreading sequences for multi-rate DS/CDMA signals," Acta Aeronautica Et Astronautica Sinica, vol.33, no.8, pp.1455-1465, 2012.

[15] W J Wu, T Q Zhang, S Shi, and Y J Zhang. “Parallel blind estimation of spreading sequences in multi-rate DS/CDMA signals,” Information and Control, vol.44, no.2, pp. 171-178, 2015.
[16] X T Ren, H Xu, Z T Huang, F H Wang, and F B Lu. “Fast-ICA based blind estimation of the spreading sequences for down-link multi-rate DS/CDMA signals,” in Proceedings of the IEEE Fifth International Conference on Intelligent Computation Technology and Automation, pp.501-504, 2012.
[17] Z J Zhao, F F Qiang, X W Gu, and X Y Jiang. “Blind estimation of pseudo-random codes in period long code spread spectrum signals by using triple correlation feature information,” Signal Processing, vol.32, no.6, pp.739-745, 2016.
[18] X W Gu, Z J Zhao, L Shen. “Blind estimation of pseudo-random codes in periodic long code direct sequence spread spectrum signals,” IET Communications, vol.10, no.11, pp. 1273-1281, 2016.

Fangfang Qiang was born in 1992. She will receive the PhD degree in electronic science and technology from Hangzhou Dianzi University, China. Currently, she is a doctoral student and her major research interests are signal processing and blind source separation.

Zhijin Zhao was born in 1959. She received the PhD degree in Xidian University in 2009.Currently, she is a professor and doctoral supervisor in the school of communications, Hangzhou Dianzi University, China. Her major research interests are cognitive radio and signal processing.

Xiaowei Gu was born in 1990. He received the MD degree in Hangzhou Dianzi University, China. Currently, he is a doctoral student in school of computing and communications, Lancaster University, UK. His major research interest is big data analysis.

Xianyang Jiang was born in 1971. He received the PhD degree in Tsinghua University. Currently, he is a associate professor in the school of communications, Hangzhou Dianzi University, China. His major research interest is broadband wireless communications.