2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Associative Memory Learns Markov Patterns Better than I.I.D. Patterns in Asynchronously Overlapping Networks

Yutaka Jitsumatsu, Shota Inoue and Tohru Kohda

6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
Email: {jitumatu, inoue, kohda}@kairo.csce.kyushu-u.ac.jp

Abstract—The analogy between the mathematical models of neural networks and code division multiple access (CDMA) systems have been discussed by many papers. Most of them assumed *synchronous* CDMA. We propose an associative memory described by *asynchronous* CDMA communications and analyze its performance. Interestingly, input patterns generated by a Markov chain can be memorized in our associative memory model more than independent and identically distributed (i.i.d.) patterns.

1. Introduction

The conventional direct sequence/code division multiple access (DS/CDMA) systems employ a matched filter receiver, where multiple-access interference (MAI) is treated as Gaussian noise. Such a receiver only utilizes the intended user's spread spectrum (SS) codes, therefore it is called single user detection. On the other hand, MAI can be significantly suppressed by multiuser detection (MUD) techniques, which exploit all users' SS codes. Since the computational complexity required for the optimal MUD receiver increases exponentially with the number of users, many suboptimal receivers have been proposed. Reduction of the complexity and derivation of the capacity bound of CDMA channels are central problems in the analysis of MUD techniques. Recently, analysis tools developed in the field of statistical mechanics, such as Ising model, spin glass, mean field approximation and replica method, have been applied to analyze the performance of MUD receivers [1, 2, 3]. Such an approach has attracted considerable attention.

MUD is often viewed as a base station model of cellular phone systems. Since mobile stations are randomly dispersed in the cell, asynchronous systems are favorable for uplink channels. There is an interesting fact in asynchronous DS/CDMA systems that Markovian SS codes can outperform linear feedback shift register (LFSR) codes as well as i.i.d. codes in terms of bit error rate (BER) [4, 5]. It is reported that Markov codes show superiority in a multiuser environment as well [6].

Motivated by such researches, we consider an associative memory model for asynchronous DS/CDMA systems[1]. Three types of connection weight matrices, i.e, correlation type, projection type and their hybrid type, are examined, which correspond to single user correlators, decorrelators and MMSE receivers, respectively. Numerical simulation shows an interesting result that input patterns with Markovity can be acquired to this associative memory more than i.i.d. patterns.

[1]Here, we use the term "asynchronous" not for the time evolution in dynamics of neural networks but for the overlapping patterns.

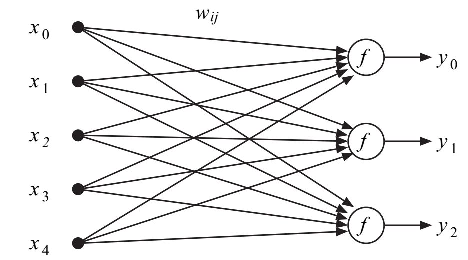

Figure 1: Hetero-associative network with five input variables and three output variables. We assume nonlinear functions f to be signum functions.

2. Associative memory

Consider a hetero-associative network illustrated in Fig. 1. The i-th element has N input variables $\{x_j\}_{j=0}^{N-1}$ with weighted connection w_{ij} so that the total input to this element is $\sum_{j=0}^{N-1} w_{ij}x_j$. Output of the element is

$$y_i = \mathrm{sgn}(\textstyle\sum_{j=0}^{N-1} w_{ij}x_j), \tag{1}$$

where $\mathrm{sgn}(u) = +1$ for $u \geq 0$ and $\mathrm{sgn}(u) = -1$ for $u < 0$. Let $(\boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)})$ be a pair of input and output vectors to be memorized ($k = 1, 2, \ldots K$), where

$$\boldsymbol{x}^{(k)} = (x_0^{(k)}, x_1^{(k)}, \ldots x_{N-1}^{(k)})^T, \tag{2}$$

$$\boldsymbol{y}^{(k)} = (y_0^{(k)}, y_1^{(k)}, \ldots y_{L-1}^{(k)})^T, \tag{3}$$

where superscript T denotes transposition of a vector or a matrix. We assume that $x_i^{(k)}$ and $y_i^{(k)}$ take values on $\{-1, +1\}$. A simple way to construct $W = \{w_{ij}\}$ is

$$W_c = YX^T = \sum_{k=1}^{K} \boldsymbol{y}^{(k)}(\boldsymbol{x}^{(k)})^T, \tag{4}$$

where $X = (\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(K)})$ and $Y = (\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)}, \ldots, \boldsymbol{y}^{(K)})$. Using a generalized inverse of the matrix X, denoted by $X^{\dagger}$, is another way to construct a connection weight matrix, which is explained in the next section. The memorized patterns are retrieved as follows.

$$\widehat{y}_i^{(k)} = \text{sgn}((W\boldsymbol{x}^{(k)})_i), \quad (5)$$
$$= \text{sgn}\left(Ny_i^{(k)} + \textstyle\sum_{j\neq k} y_i^{(j)}(\boldsymbol{x}^{(j)})^T\boldsymbol{x}^{(k)}\right) \quad (6)$$

Many researchers have pointed out an analogy among associative memory, statistical mechanics and CDMA communication systems. We observe Eq. (6) is identical to a single user receiver for synchronous CDMA systems when $y_i^{(k)}$ represents k-th user's data symbol of i-th period and $\boldsymbol{x}^{(k)}$ spread spectrum (SS) codes. In the textbook of CDMA communication, Eq.(6) is often written as

$$\widehat{\boldsymbol{y}_i} = \text{sgn}(\boldsymbol{y}_i R), \quad (7)$$

where $\boldsymbol{y}_i = (y_i^{(1)}, \ldots, y_i^{(K)})$, signum function is performed for each element and $R = (r_{jk})$ is called *correlation matrix* given by

$$r_{jk} = (\boldsymbol{x}^{(j)})^T\boldsymbol{x}^{(k)}.$$

Hence, $\boldsymbol{y}_i$ is considered as transmitted data symbols and $\widehat{\boldsymbol{y}_i}$ as demodulated data symbols. It is sometimes assumed that SS codes are balanced i.i.d. random variables. This assumption makes distribution of $(W\boldsymbol{x}^{(k)})_i/\sqrt{N}$ Gaussian with mean $\sqrt{N}y_i^{(k)}$ and variance $(K-1)$. Thus probability of incorrect demodulation, i.e., bit error rate (BER) is

$$\text{Prob}(\widehat{y}_i^{(k)} \neq y_i^{(k)}) = Q\left(\sqrt{\frac{N}{K-1}}\right), \quad (8)$$

where $Q(u) = \frac{1}{\sqrt{2\pi}} \cdot \int_u^{\infty} \exp(-x^2/2)dx$.

Note that the capacity of *auto*-associative memory is said to be $0.14N$, having attracted considerable attention. An auto-associative memory model is given by putting $N = L$ and $\boldsymbol{y} = \boldsymbol{x}$ in Fig 1 and introducing a dynamics as $\boldsymbol{x}(t+1) = \text{sgn}(W\boldsymbol{x}(t))$ for $t = 0, 1, 2, \ldots$. We leave discussions on such dynamics in connection with CDMA communications as a future work.

3. Asynchronous CDMA System

The standard associative memory model given by Eq. (6) corresponds to a synchronous CDMA system. We would like to consider an associative memory corresponding to an *asynchronous* CDMA system, where input patterns have variable overlaps.

Consider an asychronous DS/CDMA system with K users. Data and chip durations are T_d and $T_c = T_d/N$, where N is a spreading factor. Assume $T_c = 1$ without loss of generality. We denote the data and SS code signals for k-th user by $y^{(k)}(t) = \sum_{p=0}^{L-1} Y_p^{(k)} u_{T_d}(t - pT_d)$ and $x^{(k)}(t) = \sum_{n=0}^{N-1} X_n^{(k)} u_{T_c}(t - nT_c)$ respectively, where $u_{T_c}(t) = 1$ for $0 \leq t < T_c$ and $u_{T_c}(t) = 0$ otherwise. Assume that data symbols and SS codes are antipodal binary, i.e., $Y_p^{(k)}, X_n^{(k)} \in \{+1, -1\}$. Then a received signal is

$$r(t) = \sum_{k=1}^{K} y^{(k)}(t - t_k)x^{(k)}(t - t_k) + n(t), \quad (9)$$

where t_k is the k-th user's time delay and $n(t)$ denotes the external noise. The output of k-th user's correlator is given by

$$z_p^{(k)} = \int_0^{T_d} r(t + t_k + pT_d)x^{(k)}(t)dt. \quad (10)$$

Let the (row) vectors of data symbols, correlor outputs and noise be denoted by respectively,

$$\boldsymbol{y} = (y_0^{(1)}, \ldots y_0^{(K)}, y_1^{(1)}, \ldots y_1^{(K)}, \ldots y_{L-1}^{(1)}, \ldots y_{L-1}^{(K)}), \quad (11)$$
$$\boldsymbol{z} = (z_0^{(1)}, \ldots z_0^{(K)}, z_1^{(1)}, \ldots z_1^{(K)}, \ldots z_{L-1}^{(1)}, \ldots z_{L-1}^{(K)}), \quad (12)$$
$$\boldsymbol{n} = (n_0^{(1)}, \ldots n_0^{(K)}, n_1^{(1)}, \ldots n_1^{(K)}, \ldots n_{L-1}^{(1)}, \ldots n_{L-1}^{(K)}), \quad (13)$$

where $n_p^{(k)} = \int_0^{T_d} n(t + t_k + pT_d)x^{(k)}(t)dt$. Then

$$\boldsymbol{z} = \boldsymbol{y}R + \boldsymbol{n}, \quad (14)$$

where R is a $KL \times KL$ matrix given by $(R)_{ij} = \int_{\max\{t_i,t_j\}}^{T_d+\min\{t_i,t_j\}} x^{(i)}(t - t_i)x^{(j)}(t - t_j)dt$. A single user receiver gives the estimate of $\boldsymbol{y}$ as $\hat{\boldsymbol{y}} = \text{sgn}(\boldsymbol{z})$. Assume that users are ordered in descending order of time delay such that $0 \leq t_1 \leq t_2 \leq \ldots \leq t_K \leq T_d$. Then R is given by [7]

$$R = \begin{bmatrix} R(0) & R(1)^T & & 0 \\ R(1) & R(0) & \ddots & \\ & \ddots & \ddots & R(1)^T \\ 0 & & R(1) & R(0) \end{bmatrix}, \quad (15)$$

where

$$R(0) = \begin{bmatrix} N & R_{12}(t_{12}) & \ldots & R_{1K}(t_{1K}) \\ R_{21}(t_{21}) & N & \ddots & R_{2K}(t_{2K}) \\ \vdots & \ddots & \ddots & \vdots \\ R_{K1}(t_{K1}) & R_{K2}(t_{K2}) & \ldots & N \end{bmatrix},$$

$$R(1) = \begin{bmatrix} 0 & R_{12}(t_{12} + T_d) & \ldots & R_{1K}(t_{1K} + T_d) \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_{K-1,K}(t_{K-1,K} + T_d) \\ 0 & \ldots & 0 & 0 \end{bmatrix}.$$

Relative time delay is denoted by $t_{jk} = t_j - t_k$. For an integer time delay $t_{jk} = \ell$, Pursley's aperiodic cross-correlation function $R_{jk}(\ell)$ is defined in [8] as

$$R_{jk}(\ell) = \sum_{n=0}^{N-\ell-1} x_n^{(j)} x_{n+\ell}^{(k)} \quad \text{for } \ell > 0. \quad (16)$$

For a negative ℓ, $R_{jk}(\ell) = R_{kj}(-\ell)$. For a non-integer delay, $R_{jk}(\ell + \varepsilon) = (1 - \varepsilon) \cdot R_{jk}(\ell) + \varepsilon \cdot R_{jk}(\ell + 1)$, where $0 < \varepsilon < 1$.

Let we assume that relative time delays are expressed as $t_{jk} = \ell + k/M$ for $0 \le \ell < N-1$, $M \ge 1$ and $0 \le k < M-1$, where M is called up-sampling rate. Then we can observe that the correlation matrix for an asynchronous DS/CDMA system R is given by $R = \frac{1}{M}X^T X$, where $X = (\hat{\boldsymbol{x}}^{(1)}, \ldots, \hat{\boldsymbol{x}}^{(LK)})$ and

$$\hat{\boldsymbol{x}}^{(k+pK)} = (\underbrace{0,\ldots,0}_{pNM+\ell M+k}, \underbrace{x_0^{(k)}, x_0^{(k)}, \ldots, x_0^{(k)}}_{M}, \underbrace{x_1^{(k)}, \ldots, x_1^{(k)}}_{M}, \ldots, \underbrace{x_{N-1}^{(k)}, \ldots, x_{N-1}^{(k)}}_{M}, \underbrace{0,\ldots\ldots,0}_{(L-p-1)NM-\ell M-k})^T. \quad (17)$$

Hence the asynchronous DS/CDMA system described by Eqs.(14) and (15) is identical to an associative memory model, where the input vectors are up-sampled by a factor of M and memorized with overlapping. The connection weight matrix is given by

$$W_c = \tfrac{1}{M}\boldsymbol{y}X^T, \quad (18)$$

which corresponds to single user detection. The transmitted data is demodulated as

$$\begin{aligned}\widehat{y}_p^{(k)} = \mathrm{sgn}\Big[&Ny_p^{(k)} + \sum_{j\neq k} y_p^{(j)}\Big\{(1-\tfrac{m_{jk}}{M})R_{jk}(\ell_{jk}) + \tfrac{m_{jk}}{M}R_{jk}(\ell_{jk}+1)\Big\} \\ &+ \sum_{j<k} y_{p-1}^{(j)}\Big\{(1-\tfrac{m_{jk}}{M})R_{jk}(\ell_{jk}+N) + \tfrac{m_{jk}}{M}R_{jk}(\ell_{jk}+N+1)\Big\} \\ &+ \sum_{j>k} y_{p+1}^{(j)}\Big\{(1-\tfrac{m_{jk}}{M})R_{jk}(\ell_{jk}+N) + \tfrac{m_{jk}}{M}R_{jk}(\ell_{jk}+N+1)\Big\}\Big],\end{aligned}$$

where $\ell_{jk}M + m_{jk}$ is relative time delay between k-th and j-th users ($0 \le \ell_{jk} \le N-1$, $0 \le m_{jk} \le M-1$). If $x_n^{(k)}$ is balanced i.i.d. random variables, it is easy to see $\mathbf{E}\left[\{R_{kj}(\ell)\}^2\right] = N - |\ell|$ and $\mathbf{E}[R_{kj}(\ell)R_{kj}(\ell+1)] = 0$. It is therefore observed

$$\mathrm{Prob}(\widehat{y}_i^{(k)} \neq y_i^{(k)}) = Q\left(\sqrt{\frac{N}{(K-1)\sigma^2_{\mathrm{MAI}}}}\right), \quad (19)$$

where

$$\begin{aligned}\sigma^2_{\mathrm{MAI}} =& \mathrm{var}\left[(1-\tfrac{m_{jk}}{M})R_{jk}(\ell_{jk}) + \tfrac{m_{jk}}{M}R_{jk}(\ell_{jk}+1)\right] \\ &+ \mathrm{var}\left[(1-\tfrac{m_{jk}}{M})R_{jk}(\ell_{jk}+N) + \tfrac{m_{jk}}{M}R_{jk}(\ell_{jk}+N+1)\right] \\ =& \tfrac{2}{3} + \tfrac{1}{3M^2} \to \tfrac{2}{3} \quad (M \to \infty). \quad (20)\end{aligned}$$

Thus, compared with synchronous system, asynchronous system can acquire 50% more pairs of patterns with the same BER. Note that $M = 1$ implies chip-synchronous system to give $\sigma^2_{\mathrm{MAI}} = 1$. Practically $M = 5$ is found to be sufficient.

We have so far considered that the connection weight matrix is constructed by (18), which corresponds to single user detection. On the other hand,

$$W_p = \boldsymbol{y}(X^T X)^{-1}X^T = YX^\dagger \quad (21)$$

gives decorrelating detector, one of the multi-user detection (MUD) receivers. Moreover, consider the input vector is deviated from the original pattern by additive noise so that $\boldsymbol{x} = \boldsymbol{x}^{(k)} + \eta$, where η is a random variable with mean zero and variance σ^2. Then minimum mean square error (MMSE) criterion gives

$$W_h = \boldsymbol{y}(X^T X + N\sigma^2 I)^{-1}X^T. \quad (22)$$

Hence it is observed that connection matrix W in associative memory is related to the receiver's structure in CDMA systems. Matrices W_c, W_p and W_h are called correlation type, projection type and their hybrid type.

4. Markov patterns

In this section, we consider the input vectors are generated by a Markov chain. It is found that Markov chain with eigenvalue $-2+\sqrt{3}$ can be accommodated in the same associative memory better than i.i.d. input vectors by a factor of $2/\sqrt{3}$. Let us suppose that $x_n^{(k)} \in \{+1, -1\}$ ($n = 0, 1, \ldots$) is generated by a Markov chain with transition matrix

$$P = \frac{1}{2}\begin{pmatrix} 1+\lambda & 1-\lambda \\ 1-\lambda & 1+\lambda \end{pmatrix}, \quad (|\lambda| < 1) \quad (23)$$

where λ is the eigenvalue of P except 1. For $0 \le \ell < N-1$, we have [9, 10]

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{E}_{XY}[\{R_{jk}(\ell)\}^2 + \{R_{kj}(N-\ell)\}^2] = \frac{1+\lambda^2}{1-\lambda^2} \quad (24)$$

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N}\mathrm{E}_{XY}[&R_{jk}(\ell)\cdot R_{jk}(\ell+1) \\ &+ R_{kj}(N-\ell)\cdot R_{kj}(N-\ell-1)] = \frac{2\lambda}{1-\lambda^2}. \quad (25)\end{aligned}$$

It is worth noting that $R_{jk}(\ell)+R_{jk}(\ell-N)$ and $R_{jk}(\ell)-R_{jk}(\ell-N)$ are called even and odd cross-correlation functions, respectively, and that both cross-correlations asymptotically follows the same distributions [9, 11]. Eqs.(24) and (25) imply

$$\sigma^2_{\mathrm{MAI}} = \frac{2}{3}\frac{1+\lambda^2}{1-\lambda^2} + \frac{1}{3}\frac{2\lambda}{1-\lambda^2} \le \frac{1}{\sqrt{3}}. \quad (26)$$

We see σ^2_{MAI} is minimized by $\lambda = -2+\sqrt{3}$, which is coincide with asynchronous DS/CDMA systems. The capacity of this memory is increased by a factor of $\frac{2}{3}/\frac{1}{\sqrt{3}} = \frac{2}{\sqrt{3}}$.

5. Simulation Results

Numerical simulation was performed for completely asynchronous model, where up-sampling rate is $M = 5$ and Markov and i.i.d. patterns were randomly generated and embeded into a network using the connection weight matrices given by (21) and (22). Markov patterns with the optimum eigenvalue $-2+\sqrt{3}$ are shown in Fig. 2 to have lower bit error rate (BER) than i.i.d. patterns for both decorrelator and MMSE cases. Fig. 3 implies that Markov patterns can be memorized more than i.i.d. patterns with the same error

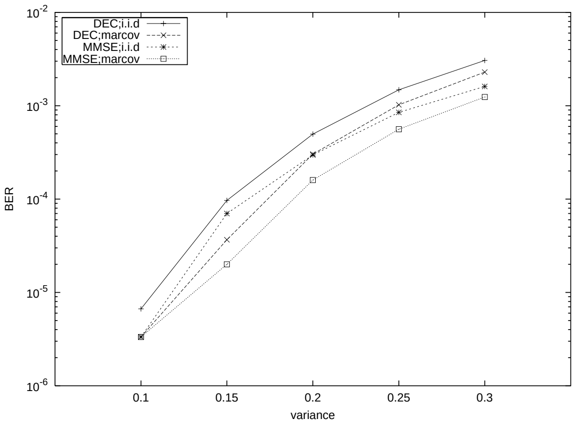

Figure 2: Bit error rates versus the variance of noise: $N = 20$, $L = 10$, and the number of patterns is $K = 30$.

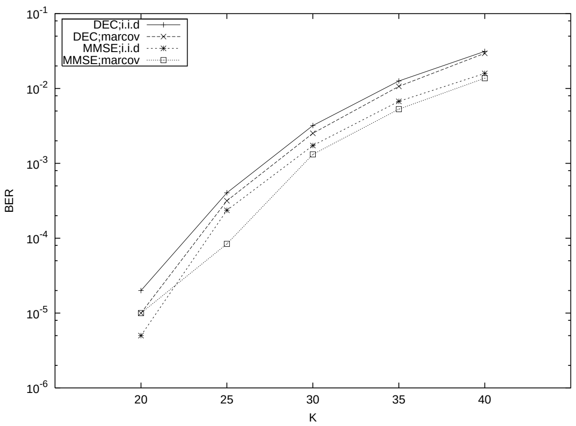

Figure 3: Bit error rates versus the number of users: $N = 20$, $L = 10$, and the variance of noise is $\sigma^2 = 0.3$.

probabilities. BER performance against eigenvalue parameter λ is illustrated in Fig.4. It is shown that BER of Markov pattern with a negative eigenvalue λ is less than that of i.i.d. pattern ($\lambda = 0$) for both decorrelator and MMSE cases. It is numerically verified that Markov patterns with $\lambda = -2+\sqrt{3}$ are optimum in terms of BER.

6. Concluding Remarks

We gave an associative memory model which can be described by an asynchronous DS/CDMA system. Three types of connection matrices, i.e. W_c, W_p and W_h, are discussed. A connection matrix W_c represents a single-user matched filter receiver and we gave the estimate of the bit error rate for this case. The other two matrices correspond to multi-user detection (MUD) receivers. Numerical simulations show that Markov patterns are easier to learn than i.i.d. patterns by an associative memory with asynchrnously overlapping connections. This result implies that Markov patterns are promising for MUD receivers in asynchronous CDMA systems as well.

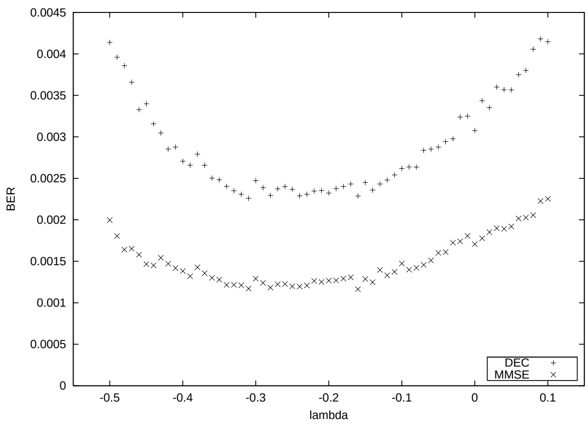

Figure 4: Bit error rates versus the eigenvalue of the Markov chain: $N = 20$, $L = 10$, the number of patterns is $K = 30$ and the variance of noise is $\sigma = 0.3$.

References

[1] D. Guo and S. Verdu, "Randomly spread CDMA: asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983 – 2010, 2005.

[2] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[3] T. Tanaka and M. Okada, "Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 700–706, Feb. 2005.

[4] G. Mazzini, G. Setti, and R. Rovatti, "Chaotic complex spreading sequences for asynchronous DS-CDMA -part I: System modeling and results," *IEEE Trans. Circuits Syst. I*, vol. 44, no. 10, pp. 937–947, 1997.

[5] G. Mazzini, R.Rovatti, and G. Setti, "Interference minimisation by autocorrelation shaping in asynchronous DS/CDMA systems: Chaos based spreading is nearly optimal," *Electronics Letters, IEE*, vol. 35, pp. 1054–55, 1999.

[6] G. S. Rovatti, R.; Mazzini, "On the ultimate limits of chaos-based asynchronous DS-CDMA-I: basic definitions and results," *IEEE Trans. Circuits Syst. I*, vol. 51, no. 7, pp. 1336–1347, July 2004.

[7] S. Verdu, *Multiuser Detection.* Cambridge, UK: Cambridge Univ. Press, 1998.

[8] M. B. Pursley, "Performance evaluation for phase-coded SS multiple-access communication-part-I: system analysis," *IEEE Trans. Commun.*, vol. 25, no. 8, pp. 795–799, Aug. 1977.

[9] T. Kohda and H. Fujisaki, "Variances of multiple access interference: Code average against data average," *Electronics Letters, IEE*, vol. 36, pp. 1717–1719, 2000.

[10] Y. Jitsumatsu and T. Kohda, "Reduction of MAI in asynchronous DS/CDMA systems using post-filter," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 9, pp. 2301–2307, Sept. 2004.

[11] T. Kohda, "Information sources using chaotic dynamics," *Proc. IEEE*, vol. 90, no. 5, pp. 641–661, May 2002.