Maximum a Posteriori Decoding System Using an ARMA Process for a Partial Response Channel with Burst Errors

Hidetoshi Saito[1], Ryuji Kohno[2]

[1] Division of Electrical Engineering, Faculty of Engineering, Kogakuin University, 1-24-2 Nishi-Shinjuku, Shinjyuku, Tokyo, 163-8677 Japan, E-mail:h-saito@cc.kogakuin.ac.jp

[2] Division of Electrical and Computer Engineering, Faculty of Engineering, Yokohama National University, 79-5 Tokiwadai, Hodogaya, Yokohama, Kanagawa, 240-8501 Japan, E-mail:kohno@ynu.ac.jp

Abstract

One of new idea of maximum a posteriori (MAP) decoding with an autoregressive and moving-average (ARMA) process is introduced into the iterative (recursive) decoding processes between MAP decoding and turbo or sum-product decoding in a perpendicular magnetic recording channel. In this paper, MAP decoding with linear prediction is called MAP-ARMA decoding. The estimation techniques based on multivariate time series analysis, stochastic models and processes are effective for burst errors or continuous erasures. It is shown that this estimation problem for missing data values (observations) is solved by maximum likelihood estimation for an ARMA process with missing values by using multivariate statistical techniques. For the given (4608,4097) low-density parity check (LDPC) coding system, the error rate performance using proposed channel predicting and MAP decoding systems in the channel with both a burst of length 70 or less and random errors attains the almost same performance of conventional MAP decoding in the channel with only random errors.

1 Introduction

Recently, the basic ideas of time series analysis and stochastic processes are introduced into the modeling of the actual perpendicular magnetic recording channel. The original idea of modeling by using an autoregressive (AR) process for the actual channel is called AR channel model in the original paper [1]. This channel modeling method is applied to the related signal processing techniques. For a representative example, Viterbi decoding with linear prediction based on the AR process in a Partial Response (PR) channel is called a PRML-AR system or PRML system using AR channel model [2]. It is known that the bit error rate (BER) performance of a PRML-AR system is superior to that of the conventional PRML system in a high density magnetic recording channel.

Further more, the model of the PRML-AR system is extended to the model with an ARMA process. It is able to use a PRML-ARMA system if a PRML-MA system based on a moving-average (MA) process is defined such as the PRML-AR system. In this PRML-ARMA system, decoding with linear prediction is used for a kind of maximum likelihood decoding with an ARMA process to estimate the codeword sequences through the magnetic recording channel. If turbo coding or low-density parity check (LDPC) coding are used, a decoder in a PR channel adopts maximum a posteriori (MAP) decoding with ARMA process and is called MAP-ARMA decoding in a similar manner.

In this paper, the previous ideas of linear prediction for a PR channel are reconsidered in terms of multivariate normal distribution. It is shown that a new method to estimate missing data values in the transmitted sequences such as burst errors or continuous erasures in an ARMA process.

2 Linear prediction and stationary process

If $\{x_k\}$ denotes a discrete-time sequence at time k, then the column vector collecting samples $x_i, x_{i+1}, \cdots, x_j$ for $i \leq j$ is denoted by $\boldsymbol{x}_i^j = [x_i, x_{i+1}, \cdots, x_j]'$. The vector $\boldsymbol{\alpha}$ denotes a binary column vector of size I ($I \geq 1$). From the original idea given by kavčić et al.[1], the block diagram of an AR channel model is shown in Fig.1 and the model is defined as follows. The sampled channel output is $y_k = x_k + n_k$, where $\{a_k\}$ is a binary input sequence, x_k is a noiseless channel output which depends on the I input symbols $\boldsymbol{\alpha} = \boldsymbol{a}_{k-I+1}^k = [a_{k-I+1}, a_{k-I}, \cdots, a_k]'$, I is the data memory length, D is a delay operator to represent a delay of 1 symbol interval and $\{n_k\}$ is an additive noise sequence. To include channel non-linearities, x_k is constructed as a look-up-table rather than the more common convolution between the I input symbols in the data memory and a linear channel-response. The noise term n_k is the output of a signal-dependent autoregressive filter whose input is a zero-mean unit-variance white Gaussian noise sequence $\{n_k^e\}$.

$$n_k = \sum_{i=1}^{L} \phi_i(\boldsymbol{a}_{k-I+1}^k) n_{k-i} + \sigma(\boldsymbol{a}_{k-I+1}^k) n_k^e \quad (1)$$

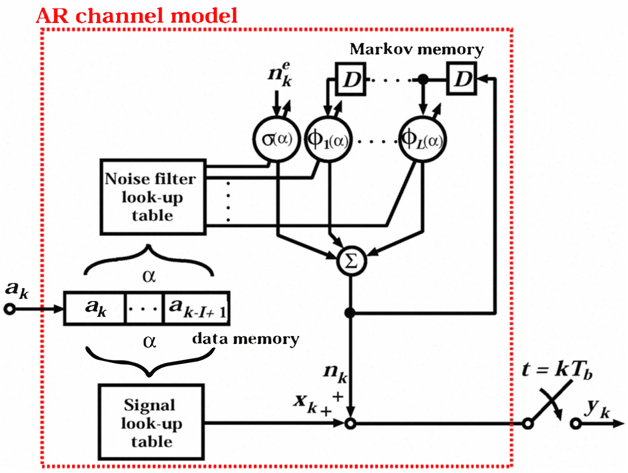

Fig. 1. Block diagram of an autoregressive channel model.

$$= \boldsymbol{\phi}(\boldsymbol{\alpha})' \boldsymbol{n}_{k-L}^{k-1} + \sigma(\boldsymbol{\alpha}) n_k^e ,$$

where a coefficient of the autoregressive filter at time k is composed of the vector of tap-weights $\boldsymbol{\phi}(\boldsymbol{\alpha}) = [\phi_L(\boldsymbol{\alpha}), \cdots, \phi_1(\boldsymbol{\alpha})]'$ and the standard deviation $\sigma(\boldsymbol{\alpha})$. From (1), the noise sequence $\{n_k\}$ is considered to be a signal-dependent Gauss-Markov noise process with Markov memory length L. This means that the sequence $\{n_k\}$ conforms to an autoregressive process of order L (AR(L) process). This AR channel model is applied to the optimal maximum-likelihood sequence detector (MLSD) with a linear predictor based on an AR process in the correlated finite memory intersymbol interference (ISI) channels [2].

Fig.2 shows the proposed MLSD with an ARMA process. This MLSD with two prediction blocks is naturally extended from the original Kavčić MLSD in [2]. This detector consists of a signal look-up table, a metric calculation block, prediction blocks based on an ARMA process. In this prediction blocks are divided to two independent sub-blocks with an AR process and a moving-average (MA) process, respectively. The noise term n_k is written as

$$\begin{aligned} n_k - \sum_{i=1}^{p} \phi_i(\boldsymbol{a}_{k-I+1}^k) n_{k-i} &= e_k + \sum_{j=1}^{q} \theta_j(\boldsymbol{a}_{k-I+1}^k) e_{k-j} , \\ n_k - \boldsymbol{\phi}(\boldsymbol{\alpha})' \boldsymbol{n}_{k-p}^{k-1} &= e_k + \boldsymbol{\theta}(\boldsymbol{\alpha})' \boldsymbol{e}_{k-q}^{k-1} , \\ e_k &= \sigma(\boldsymbol{a}_{k-q+1}^k) n_k^e , \end{aligned} \tag{2}$$

where the vectors of coefficient tap-weights in the both of FIR filters are $\boldsymbol{\phi}(\boldsymbol{\alpha}) = [\phi_p(\boldsymbol{\alpha}), \cdots, \phi_1(\boldsymbol{\alpha})]'$ and $\boldsymbol{\theta}(\boldsymbol{\alpha}) = [\theta_q(\boldsymbol{\alpha}), \cdots, \theta_1(\boldsymbol{\alpha})]'$, respectively. For natural convenience, $\phi_i(\boldsymbol{\alpha}) = 0$ if $i > p$ and $\theta_j(\boldsymbol{\alpha}) = 0$ if $j > q$. The metric calculation block is composed of a branch metric calculator, an add-compare-select (ACS) circuit and a state selector for a trellis diagram. The signal look-up table is given by averaging the output sequences of the correlated finite memory ISI channel for each input sequence vector $\boldsymbol{\alpha}$ with data memory

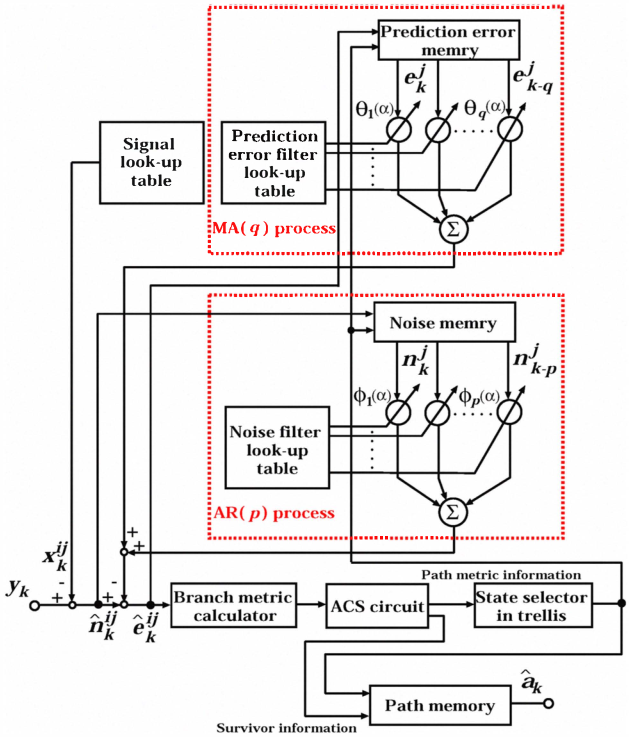

Fig. 2. Block diagram of the maximum-likelihood sequence detector with an autoregressive and moving-average process.

length I and decides the number of 2^{I-1} states in the MLSD. In decoding process, the state transition from $S(k) = S_i$ to $S(k+1) = S_j$ gives the signal x_k^{ij} at time k from the signal look-up table, where $S(k) = S_i$ means that the state at time k is the i-th state of a trellis diagram in the MLSD, were $0 \leq i, j \leq 2^{L-1}$. At first, the tentative estimated noise $\hat{n}_k^{ij}$ at time k for each state transition is obtained by subtracting each signal x_k^{ij} from the decoder input signal y_k and is a candidate of input to the 2^{L-1} Markov memory as n_k^j. From (2), the tentative prediction error $\hat{e}_k^{ij}$ at time k for each state transition is obtained by subtracting each $\hat{n}_k^{ij}$ from the both FIR filter outputs $\sum_{\ell=1}^{p} \phi_\ell(\boldsymbol{\alpha}) n_{k-\ell}$ and $\sum_{m=1}^{q} \theta_m(\boldsymbol{\alpha}) e_{k-m}$. This tentative prediction error $\hat{e}_k^{ij}$ is a candidate of input to the 2^{L-1} Markov memory as e_k^j. In this paper, the Markov memory to keep previous p noise samples n_{k-p} through n_{k-1} and previous q prediction error samples e_{k-q} through e_{k-1} are called the noise memory and prediction error memory, respectively. The lengths p and q are corresponding to the noise memory and prediction error memory lengths. The sequences in the noise and prediction error memories are replaced with more likely sequences by the metric calculation block as well as the path memory.

3 Multivariate normal distribution

In the previous section, the determination of an appropriate ARMA process represents observed stationary time series which are composed of noise n_k or prediction error e_k at time k in a data sequence,

$k = 1, 2, \cdots, N$. This theory of univariate time series extends to N time series $\{n_{ti}\}$ and $\{e_{ti}\}$ with means $En_{ti}^2 < \infty$ and $Ee_{ti}^2 < \infty$, respectively. If all infinite dimensional joint distributions of the random variables $\{X_{ti}\}$ which represents $\{n_{ti}\}$ or $\{e_{ti}\}$ were multivariate normal, then the distributional properties of $\{X_{ti}\}$ would be completed determined by the means,

$$\mu_{ti} = EX_{ti}, \tag{3}$$

and covariances,

$$\gamma_{ij}(t+h, t) = E\left[(X_{t+h,i} - \mu_{t+h,i})(X_{tj} - \mu_{tj})\right]. \tag{4}$$

It is more convenient when dealing with N interrelated series to use vector notation. It is defined an N-variate random vector as

$$\boldsymbol{X}_t = (X_{t1}, X_{t2}, \cdots, X_{tN})', \quad t = 0, 1, 2, \cdots \tag{5}$$

The second-order properties of the multivariate time series $\{\boldsymbol{X}_{ti}\}$ are then specified by the mean vectors,

$$\boldsymbol{\mu}_t = E\boldsymbol{X}_t = (\mu_{t1}, \mu_{t2}, \cdots, \mu_{tN})', \tag{6}$$

and covariance matrices,

$$\begin{aligned} \boldsymbol{\Gamma}(t+h, t) &= E\left[(\boldsymbol{X}_{t+h} - \boldsymbol{\mu}_{t+h})(\boldsymbol{X}_t - \boldsymbol{\mu}_t)'\right] \\ &= \left[\gamma_{ij}(t+h, t)\right]_{i,j=1}^{N}. \end{aligned} \tag{7}$$

The series $\{n_{ti}\}$ and $\{e_{ti}\}$ with means and covariances are stationary because $\boldsymbol{\mu}_t$ and $\Gamma(t+h, t)$, $t = 0, 1, 2, \cdots$, are independent of t in these series. Therefore, it is able to consider N-variate random vectors of n_k and e_k. It is defined as multivariate white noise which is written as $\{\boldsymbol{Z}_t\} \sim \mathcal{WN}(\boldsymbol{0}, \boldsymbol{\Sigma})$, if the N-variate series $\{\boldsymbol{Z}_t\}$ is stationary with mean vector $\boldsymbol{0}$ and covariance matrix function,

$$\boldsymbol{\Gamma}(h) = \begin{cases} \boldsymbol{\Sigma} & \text{if } h = 0. \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

4 Estimation of missing values in multivariate time series

Suppose it is given N data values $Z_{i_1}, \cdots, Z_{i_r}, U_{i_{r+1}}, \cdots, U_{i_N}$ $(1 \le i_{r+1} < i_{r+2} \cdots < i_{N-1} \le i_N)$ of an ARMA(p,q) process. There are r missing values and $N-r$ existing values in this data values. It is defined that the r-variate random vector of missing values is $\boldsymbol{Z} = (Z_{i_1}, \cdots, Z_{i_r})'$ and $(N-r)$-variate random vector of existing values is $\boldsymbol{U} = (U_{i_{r+1}}, \cdots, U_{i_N})'$. It wishes to find the optimal linear estimates of missing values Z_j, $j \in \{i_1, i_2, \cdots, i_r\}$, in terms of U_k, $k \in \{i_{r+1}, i_{r+2}, \cdots, i_N\}$ and the components of $\boldsymbol{Z}$. If $\boldsymbol{N}$ is the N-variate random vector of N data values, $\boldsymbol{N} = (\boldsymbol{Z}', \boldsymbol{U}')'$. It is represented that $\{\boldsymbol{N}\} \sim \mathcal{WN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ because the mean vector of $\boldsymbol{N} = \boldsymbol{\mu}$ and the $N \times N$ covariance matrix $\boldsymbol{\Sigma} = [\sigma_{ij}]_{i,j=1}^{N}$. At first, $\boldsymbol{Y}$ is an N-variate random vector as follows:

$$\begin{aligned} \boldsymbol{Y} &= \begin{bmatrix} \boldsymbol{Y}_1 \ (r \times 1) \\ \boldsymbol{Y}_2 \ ((n-r) \times 1) \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{Z} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{U} \\ \boldsymbol{U} \end{bmatrix}, \end{aligned} \tag{9}$$

where $\boldsymbol{Y}_1 = (Y_{\ell_1}^1, \cdots, Y_{\ell_r}^1)'$, $\boldsymbol{Y}_2 = (Y_{\ell_{r+1}}^2, \cdots, Y_{\ell_N}^2)'$ and the covariance matrix is represented by

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}. \tag{10}$$

In this matrix $\boldsymbol{\Sigma}$, the notations $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12}$, $\boldsymbol{\Sigma}_{21}$ and $\boldsymbol{\Sigma}_{22}$ are the $r \times r$, $r \times (N-r)$, $(N-r) \times r$ and $(N-r) \times (N-r)$ submatrices, respectively. If $\boldsymbol{C}$ is an $N \times N$ matrix and written as

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}, \tag{11}$$

it is shown that $\boldsymbol{Y} = \boldsymbol{CN}$. From (11), it gives

$$\begin{aligned} \mathbf{mod}|\boldsymbol{C}^{-1}| &= \frac{\mathbf{1}}{\mathbf{mod}|\boldsymbol{C}|} = \sqrt{\frac{\mathbf{1}}{|\boldsymbol{C}|^{\mathbf{2}}}} \\ &= \sqrt{\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{C}| \cdot |\boldsymbol{\Sigma}| \cdot |\boldsymbol{C}'|}} = \sqrt{\frac{|\boldsymbol{\Sigma}|}{|\boldsymbol{C\Sigma C}'|}}, \end{aligned} \tag{12}$$

where the notation mod is an absolute value of determinant. Then, it is defined that the probability density function of $\boldsymbol{Y}$ is $f(y_{\ell_1}^1, \cdots, y_{\ell_r}^1, y_{\ell_{r+1}}^2, \cdots, y_{\ell_N}^2)$ and written as

$$f(\boldsymbol{y}) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})\right\}, \tag{13}$$

where $\{\boldsymbol{Y}\} \sim \mathcal{WN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and the mean vector $\boldsymbol{\mu} = E(\boldsymbol{Y}) = (\mu_1, \mu_2, \cdots, \mu_N)'$. From (13),

$$\begin{aligned} Q &= (\boldsymbol{n} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{n} - \boldsymbol{\mu}) \\ &= (\boldsymbol{C}^{-1}\boldsymbol{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}\boldsymbol{y} - \boldsymbol{\mu}) \\ &= (\boldsymbol{C}^{-1}\boldsymbol{y} - \boldsymbol{C}^{-1}\boldsymbol{C\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{C}^{-1}\boldsymbol{y} - \boldsymbol{C}^{-1}\boldsymbol{C\mu}) \\ &= \left[(\boldsymbol{C}^{-1}(\boldsymbol{y} - \boldsymbol{C\mu})\right]'\boldsymbol{\Sigma}^{-1}\left[(\boldsymbol{C}^{-1}(\boldsymbol{y} - \boldsymbol{C\mu})\right] \\ &= (\boldsymbol{y} - \boldsymbol{C\mu})'(\boldsymbol{C}^{-1})'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}^{-1}(\boldsymbol{y} - \boldsymbol{C\mu}) \\ &= (\boldsymbol{y} - \boldsymbol{C\mu})'(\boldsymbol{C\Sigma C}')^{-1}(\boldsymbol{y} - \boldsymbol{C\mu}). \end{aligned} \tag{14}$$

From (13) and (14), a probability density function $\boldsymbol{Y}$ is written as

$$\begin{aligned} f(\boldsymbol{y}) = &\frac{1}{\sqrt{(2\pi)^N |\boldsymbol{C\Sigma C}'|}} \\ &\exp\left\{-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{C\mu})'(\boldsymbol{C\Sigma C}')^{-1}(\boldsymbol{y} - \boldsymbol{C\mu})\right\} \end{aligned} \tag{15}$$

and it is found that $\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{N} \sim \mathcal{WN}(\boldsymbol{C}\boldsymbol{\mu}, \boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^{'})$. The mean vector and covariance matrix of $\boldsymbol{Y}$ are written as follows.

$$\boldsymbol{C}\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\nu_1} \\ \boldsymbol{\nu_2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu_1} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu_2} \\ \boldsymbol{\mu_2} \end{bmatrix}. \tag{16}$$

$$\begin{aligned} &\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^{'} \\ &= \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \boldsymbol{I} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}. \end{aligned} \tag{17}$$

If $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$,

$$\begin{aligned} &\begin{bmatrix} \boldsymbol{\Sigma}_{11.2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_{11.2}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \\ &\begin{vmatrix} \boldsymbol{\Sigma}_{11.2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22} \end{vmatrix} = |\boldsymbol{\Sigma}_{11.2}| \cdot |\boldsymbol{\Sigma}_{22}|. \end{aligned} \tag{18}$$

From (15),

$$\begin{aligned} Q &= (\boldsymbol{y} - \boldsymbol{C}\boldsymbol{\mu})^{'}(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}^{'})^{-1}(\boldsymbol{y} - \boldsymbol{C}\boldsymbol{\mu}) \\ &= \left[(\boldsymbol{y}_1 - \boldsymbol{\nu}_1)^{'} \ (\boldsymbol{y}_2 - \boldsymbol{\nu}_2)^{'}\right] \begin{bmatrix} \boldsymbol{\Sigma}_{11.2}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_1 - \boldsymbol{\nu}_1 \\ \boldsymbol{y}_2 - \boldsymbol{\nu}_2 \end{bmatrix} \\ &= (\boldsymbol{y}_1 - \boldsymbol{\nu}_1)^{'}\boldsymbol{\Sigma}_{11.2}^{-1}(\boldsymbol{y}_1 - \boldsymbol{\nu}_1) + (\boldsymbol{y}_2 - \boldsymbol{\nu}_2)^{'}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{y}_2 - \boldsymbol{\nu}_2) \\ &= Q_1 + Q_2. \end{aligned} \tag{19}$$

In consequence, $\boldsymbol{Y}$ in (15) is written as

$$\begin{aligned} f(\boldsymbol{y}) &= \frac{1}{\sqrt{(2\pi)^r|\boldsymbol{\Sigma}_{11.2}|}} \exp\left\{-\frac{Q_1}{2}\right\} \\ &\quad + \frac{1}{\sqrt{(2\pi)^{N-r}|\boldsymbol{\Sigma}_{22}|}} \exp\left\{-\frac{Q_2}{2}\right\} \\ &= f(\boldsymbol{y}_1, \boldsymbol{y}_2). \end{aligned} \tag{20}$$

In (20), the r-variate random vector $\boldsymbol{Y}_1$ and the $(n-r)$-variate vector $\boldsymbol{Y}_2$ are independently and identically distributed. These functions are marginal distribution functions of $\boldsymbol{Y}$ and satisfies that

$$\begin{aligned} \boldsymbol{Y}_1 &\sim \mathcal{WN}(\boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{11.2}), \\ \boldsymbol{Y}_2 &\sim \mathcal{WN}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}). \end{aligned} \tag{21}$$

If $\boldsymbol{N}$ is transformed into $\boldsymbol{Y}$ bijectively, it is defined that

$$\begin{aligned} Y_{\ell_j}^1 &= Y_{\ell_j}^1(Z_{i_1}, Z_{i_2}, \cdots, Z_{i_r}, U_{i_{r+1}}, U_{i_{r+2}}, \cdots, U_{i_N}), \\ &\qquad j = 1, \cdots, r, \\ Y_{\ell_k}^2 &= Y_{\ell_k}^2(Z_{i_1}, Z_{i_2}, \cdots, Z_{i_r}, U_{i_{r+1}}, U_{i_{r+2}}, \cdots, U_{i_N}), \\ &\qquad k = r+1, \cdots, N \end{aligned} \tag{22}$$

and the inverse transforms are defined as

$$\begin{aligned} Z_{i_j} &= Z_{i_j}(Y_{\ell_1}^1, Y_{\ell_2}^1, \cdots, Y_{\ell_r}^1, Y_{\ell_{r+1}}^2, Y_{\ell_{r+2}}^2, \cdots, Y_{\ell_N}^2), \\ &\qquad j = 1, \cdots, r, \\ U_{i_k} &= U_{i_k}(Y_{\ell_1}^1, Y_{\ell_2}^1, \cdots, Y_{\ell_r}^1, Y_{\ell_{r+1}}^2, Y_{\ell_2}^2, \cdots, Y_{\ell_N}^2), \\ &\qquad k = r+1, \cdots, N. \end{aligned} \tag{23}$$

From (22) and (23), the probability density function $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_N)$ is given as

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{z}, \boldsymbol{u}) \\ &= g(Y_{1_1}(\boldsymbol{n}), \cdots, Y_{1_r}(\boldsymbol{n}), Y_{2_1}(\boldsymbol{n}), \cdots, Y_{2_{N-r}}(\boldsymbol{n})) \\ &\qquad J(\boldsymbol{n} : \boldsymbol{y}) \\ &= g(Y_{1_1}(\boldsymbol{z}, \boldsymbol{u}), \cdots, Y_r(\boldsymbol{z}, \boldsymbol{u}), Y_{2_1}(\boldsymbol{z}, \boldsymbol{u}), \cdots, Y_{2_{N-r}}(\boldsymbol{z}, \boldsymbol{u})) \\ &\qquad J(\boldsymbol{z}, \boldsymbol{u} : \boldsymbol{y}_1, \boldsymbol{y}_2), \end{aligned} \tag{24}$$

where $J(\boldsymbol{n} : \boldsymbol{y})$ is the Jacobian which is given by

$$\begin{aligned} &J(\boldsymbol{z}, \boldsymbol{u} : \boldsymbol{y}_1, \boldsymbol{y}_2) \\ &= \mathrm{mod} \begin{vmatrix} \frac{\partial z_{i_1}}{\partial y_{\ell_1}^1} \cdots \frac{\partial z_{i_1}}{\partial y_{\ell_r}^1} & \frac{\partial z_{i_1}}{\partial y_{\ell_{r+1}}^2} \cdots \frac{\partial z_{i_1}}{\partial y_{\ell_N}^2} \\ \cdots \ \cdots \ \cdots & \cdots \ \cdots \ \cdots \\ \frac{\partial z_{i_r}}{\partial y_{\ell_r}^1} \cdots \frac{\partial z_{i_r}}{\partial y_{\ell_r}^1} & \frac{\partial z_{i_r}}{\partial y_{\ell_{r+1}}^2} \cdots \frac{\partial z_{i_r}}{\partial y_{\ell_N}^2} \\ \frac{\partial u_{i_{r+1}}}{\partial y_{\ell_1}^1} \cdots \frac{\partial u_{i_{r+1}}}{\partial y_{\ell_r}^1} & \frac{\partial u_{i_{r+1}}}{\partial y_{\ell_{r+1}}^2} \cdots \frac{\partial u_{i_{r+1}}}{\partial y_{\ell_N}^2} \\ \cdots \ \cdots \ \cdots & \cdots \ \cdots \ \cdots \\ \frac{\partial u_{i_N}}{\partial y_{\ell_1}^1} \cdots \frac{\partial u_{i_N}}{\partial y_{\ell_r}^1} & \frac{\partial u_{i_N}}{\partial y_{\ell_{r+1}}^2} \cdots \frac{\partial u_{i_N}}{\partial y_{\ell_N}^2} \end{vmatrix} \\ &= 1. \end{aligned} \tag{25}$$

The probability density function of $\boldsymbol{Z}$ and $\boldsymbol{U}$ is given by

$$\begin{aligned} &f(\boldsymbol{z}, \boldsymbol{u}) = \frac{1}{\sqrt{(2\pi)^r|\boldsymbol{\Sigma}_{11.2}|}} \\ &\exp\Big\{-\frac{1}{2}\left[(\boldsymbol{z} - \boldsymbol{\mu}_1) - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u} - \boldsymbol{\mu}_2)\right]^{'} \\ &\boldsymbol{\Sigma}_{11.2}\left[(\boldsymbol{z} - \boldsymbol{\mu}_1) - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u} - \boldsymbol{\mu}_2)\right]\Big\} \\ &\frac{1}{\sqrt{(2\pi)^{N-r}|\boldsymbol{\Sigma}_{22}|}} \exp\Big\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\mu}_2)^{'}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u} - \boldsymbol{\mu}_2)\Big\}. \end{aligned} \tag{26}$$

From (13), it is given that

$$\frac{1}{\sqrt{(2\pi)^{N-r}|\boldsymbol{\Sigma}_{22}|}} \exp\Big\{-\frac{1}{2}(\boldsymbol{u} - \boldsymbol{\mu}_2)^{'}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u} - \boldsymbol{\mu}_2)\Big\}. \tag{27}$$

Then, the conditional probability density function of $\boldsymbol{Z}$ given $\boldsymbol{U}$ is defined to be

$$f(\boldsymbol{z}|\boldsymbol{u}) = \frac{1}{(2\pi)^{\frac{r}{2}}|\boldsymbol{\Sigma}_{11.2}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\left[(\boldsymbol{z}-\boldsymbol{\mu}_1)-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u}-\boldsymbol{\mu}_2)\right]' \boldsymbol{\Sigma}_{11.2}^{-1}\left[(\boldsymbol{z}-\boldsymbol{\mu}_1)-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u}-\boldsymbol{\mu}_2)\right]\right\}. \quad (28)$$

If $\det|\boldsymbol{\Sigma}_{22}| > 0$, the conditional distribution of $\boldsymbol{Z}$ given $\boldsymbol{U}$ is $\mathcal{WN}(\boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u}-\boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11.2})$. It is shown that

$$\begin{aligned} E\{\boldsymbol{Z}|\boldsymbol{U}=\boldsymbol{u}\} &= \boldsymbol{\mu}_1+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{u}-\boldsymbol{\mu}_2), \\ V\{\boldsymbol{Z}|\boldsymbol{U}=\boldsymbol{u}\} &= \boldsymbol{\Sigma}_{11.2}. \end{aligned} \quad (29)$$

It is clear that the values $f(\boldsymbol{z},\boldsymbol{u})$ and $f(\boldsymbol{z}|\boldsymbol{u})$ are maximized if it is satisfied

$$\boldsymbol{z} = E\{\boldsymbol{Z}|\boldsymbol{U}=\boldsymbol{u}\}. \quad (30)$$

This means that the optimal estimation problem of $\boldsymbol{Z}$ using $\boldsymbol{U}$ results in the problem of maximization of the joint density function $f(\boldsymbol{z},\boldsymbol{u})$ for $\boldsymbol{z}$. Then, it is assumed each optimal estimated value $Z_{i_k}^*$ for Z_{i_k} $(1 \le k \le r)$ in terms of $\boldsymbol{U}$ is defined as linear combination of 1 and $\boldsymbol{U}$ with real coefficients c_i $(0 \le i \le N-r)$.

$$\begin{aligned} Z_{i_k}^* &= c_0 + c_1U_{i_{r+1}} + c_2U_{i_{r+2}} + \cdots + c_{N-r}U_{i_N} \\ &= c_0 + \boldsymbol{c}'\mathrm{U}. \end{aligned} \quad (31)$$

In (31), the linear combination of 1, $U_{i_{r+1}}$, $U_{i_{r+2}}$, $\cdots$, U_{i_N} is the optimal approximate of Z_{i_k} in the sense that following equation is minimized.

$$\begin{aligned} &\mathcal{F}(c_0,c_1,\cdots,c_{N-r}) \\ &= E\Big[(Z_{i_k}-c_0-c_1U_{i_{r+1}}-\cdots-c_{N-r}U_{i_N})^2\Big] \\ &= E\Big[(Z_{i_k}-c_0-\boldsymbol{c}'\boldsymbol{U})^2\Big]. \end{aligned} \quad (32)$$

If $\boldsymbol{N}$ consists of $\boldsymbol{Z} = (Z_{i_k})$ and $\boldsymbol{U} = (U_{i_{r+1}}, U_{i_{r+2}}, \cdots, U_{i_N})'$, it is defined as

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \boldsymbol{\sigma}_1' \\ \boldsymbol{\sigma}_1 & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \quad (33)$$

where $\boldsymbol{\sigma}_1'$, $\boldsymbol{\sigma}_1$ and $\boldsymbol{\Sigma}_{22}$ are the $1 \times (N-r)$, $(N-r) \times 1$ and $(N-r) \times (N-r)$ submatrices, respectively. From (32), it gives

$$\begin{aligned} &\mathcal{F}(c_0,c_1,\cdots,c_{N-r}) \\ &= E\Big[\Big\{(Z_{i_k}-\mu_1)-\boldsymbol{c}'(\boldsymbol{U}-\boldsymbol{\mu}_2)-(c_0-\mu_1+\boldsymbol{c}'\boldsymbol{\mu}_2)\Big\}^2\Big] \\ &= \sigma_{11}+\boldsymbol{c}'\boldsymbol{\Sigma}_{22}\boldsymbol{c}-2\boldsymbol{c}'\boldsymbol{\sigma}_1+(c_0-\mu_1+\boldsymbol{c}'\boldsymbol{\mu}_2)^2 \\ &= (\sigma_{11}-\boldsymbol{\sigma}_1'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_1)+(\boldsymbol{c}-\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_1)'\boldsymbol{\Sigma}_{22}(\boldsymbol{c}-\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_1) \\ &\quad + (c_0-\mu_1+\boldsymbol{c}'\boldsymbol{\mu}_2)^2. \end{aligned} \quad (34)$$

If it is given as

$$\begin{aligned} \boldsymbol{c}' &= \boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_1, \\ c_0 &= \mu_1-\boldsymbol{c}'\boldsymbol{\mu}_2, \end{aligned} \quad (35)$$

it is found that the minimum optimal value of $\mathcal{F}(c_0,c_1,\cdots,c_{N-r})$ is obtained as $\sigma_{11}-\boldsymbol{\sigma}_1'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_1$ and all terms in (34) take nonnegative values. From (31) and (35),

$$Z_{i_k}^* = \mu_1+\boldsymbol{c}'(\boldsymbol{U}-\boldsymbol{\mu}_2) \quad (36)$$

and it is shown that (36) is equivalent to (29) or (30). In an ARMA process, the solutions of missing n_{i_k} in (2) correspond to $Z_{i_k}^*$ $(1 \le k \le r)$ and these solutions are obtained by using the results of (35), (36) and the autocovariance functions of an ARMA process. (The method of calculating the autocovariance functions of an ARMA process refers to [3], etc.)

5 Simulation system in a perpendicular magnetic recording channel

Fig.3 shows the block diagram of the LDPC coded PR system with the proposed MAP decoding system. In this read/write system, a raw data sequence $\{a_k\}$ with bit rate f_b is inputted into a LDPC encoder and the codeword sequence $\{b_{k'}\}$ is generated. The interleaver Π takes each incoming block of the original sequence $\{b_{k'}\}$ and rearranges the sequence in a different temporal order. The interleaved sequence for $\{b_{k'}\}$ with additional symbols which terminate a trellis diagram of the PR system is transformed into the sequence $\{c_{k'}\}$ at the precoder. The sequence $\{c_k\}$ is NRZ-recorded on the perpendicular double-layered medium. Here, an isolated reproducing waveform at the reading point is assumed to be a hyperbolic tangent function-like waveform given by

$$h(t) = A_p \tanh\left(\frac{\ln 3}{T_{50}}t\right), \quad (37)$$

where A_p is a half of the amplitude and T_{50} is the time interval which $h(t)$ needs to rise from $-A_p/2$ to $A_p/2$. The normalized linear density (user density) is defined as

$$K_p = \frac{T_{50}}{T_b}, \quad (38)$$

where T_b is a user bit interval. The reproducing waveform corresponding to the recording sequence read back by reading head is inputted into the equalizer which consists of a low-pass filter with the cut-off frequency x_h normalized by bit rate f_b and the transversal filter with N_t taps. In this read/write system, the noise at reading point consists of additive white Gaussian noise and jitter-like noise. The equalization is performed so that the overall characteristic between the input of recording head and the output of the equalizer is equal to the aimed PR characteristic. It assumes that the noise at the reading point is zero-mean, white Gaussian noise with two-sided power spectral density equal to $N_0/2$

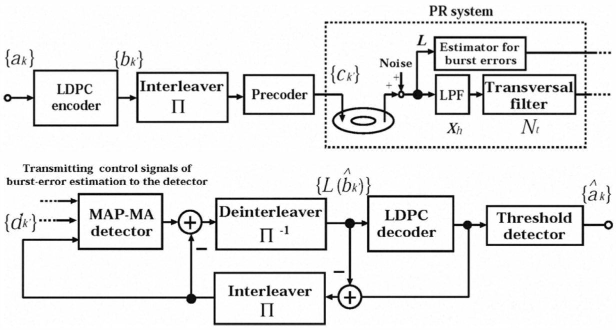

Fig. 3. Block diagram of the LDPC coding and decoding systems.

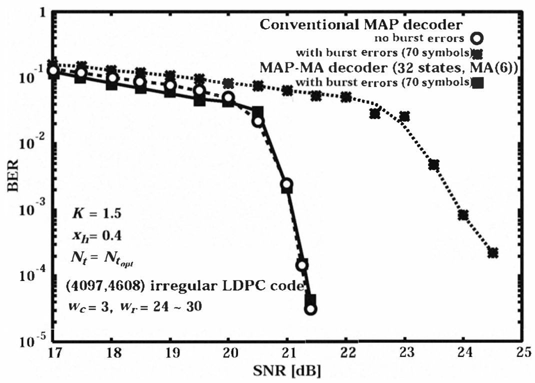

Fig. 4. Bit error rate performances for the LDPC coding and decoding systems.

and the signal-to-noise ratio (SNR) at the reading point is defined as

$$a = \frac{A_p}{\sqrt{0.6 N_0 f_b}}\,, \tag{39}$$

where σ_w^2 is the total noise power within the bandwidth of $0.6f_b$. In decoding process, it is assumed that the data sequence of length L bits in the equalizer output sequence $\{d_{k'}\}$ is missing. The continuous L erasures are cased by thermal asperity (TA) and give sudden increment or boosting of signals during these intervals. In this paper, these erasures are called burst errors. For these deep fade of the signal levels, it is assumed that the burst estimator detects the start point of TA and intervals of this missing data sequence by using the distance-based method with a threshold [4]. During this L bits intervals, erasure flags set up the values of signal levels and the extrinsic log-likelihood ratio (LLR) into MAP detector to zeros. [5]. The equalizer output sequence $\{d_{k'}\}$ is decoded by the proposed MAP decoding algorithm with an ARMA process in the soft decision detector. The log-likelihood ratio (LLR) sequence $\{L(b_{k'})\}$ correspond to the soft output sequence of $\{b_{k'}\}$ is obtained from the deinterleaved MAP decoder output sequence. This sequence $\{L(b_{k'})\}$ is a codeword sequence generated by the LDPC encoder and the codeword sequence is decoded by a revised sum-product algorithm based on a decomposed minimal trellis of each parity check codeword in the LDPC decoder. The output data sequence $\{\hat{a}_{k'}\}$ is obtained after the recursive decoding process between the proposed MAP detector and LDPC decoder. The BER of this read/write system is evaluated by computer simulation.

6 Simulation result

Fig.4 shows the BER performances of the LDPC coded PR systems in the recursive decoding. This read/write system uses an irregular type (4608,4097) Gallager LDPC code as a recording code which column weights w_c are 3 and row weights w_r are between 24 and 30. The symbol ○ indicate the BER performance for the system with conventional MAP detector in the case of only random errors which occur in the PR channel (it is called a random-error channel). The symbol $*$ shows the BER performances with the conventional MAP detector in the case of both of random and burst errors. The symbol □ shows the BER performances with MAP-MA(6) detector where K_p=1.5, $x_h = 0.4$, $q = 6$ and $N_t = N_{t_{opt}}$. The symbol $N_{t_{opt}}$ is the optimum value of N_t which gives the near minimum BER. The interleaver Π is an s-random interleaver, where s is defined as the minimum interleaving distance and s takes at least $q + 1$. As can be seen the Fig.4, the performance of the system in the random-error channel improves the SNR about by 3.0 dB over the system in the channel which contains a combination of both random and burst errors (it is called a compound channel) at a BER of 10^{-4}, where the burst of length L = 70, s = 7 and PR(1,2,1) (PR2) system with 1 symbol delayed precoder is used. The BER performance of the system after the burst error correcting by proposed MAP-MA detector in this compound channel is almost equal to that of the system in the random-error channel.

7 Conclusion

In this paper, it is shown that the maximum a posteriori probability decoding using an ARMA process (MAP-ARMA) is useful to correct a burst of length 70 or less and random errors when the transmission errors occur randomly and in clusters. The results show that the proposed MAP-ARMA decoding outperforms the conventional MAP decoding in the bit error rate performance LDPC coded PR channel with these compound errors.

References

[1] A. Kavčić and A. Patapoutian, IEEE Trans. Magn., vol.35, no.5, pp.2316–2318, Sep. 1999.

[2] A. Kavčić and J. M. F. Moura, IEEE Trans. Inform. Theory, Vol.46, No.1, pp.291–301, Jan. 2000.

[3] P. J. Brockwell and R. A. Davis, " Time Series: Theory and Methods," Springer-Verlag, New York, 2nd edition, 1991.

[4] M. N. Kaynak, T. M. Muman and E. M. Kurtas, IEEE Trans. Magn. vol.40, no.4, pp.3087–3089, Jul. 2004.

[5] W. Tan and J. R. Cruz, IEEE Trans. Magn. vol.39, no.5, pp.2579–2581, Sept. 2003.