On the Capacity of Write-Unidirectional Memories With Nonperiodic Codes

Fang-Wei Fu, A. J. Han Vinck, *Senior Member, IEEE*,
Victor K. Wei, *Fellow, IEEE*, and Raymond W. Yeung, *Fellow, IEEE*

***Abstract*—Write-unidirectional memories (WUMs) were introduced by Willems, Vinck, and Borden as an information-theoretic model for storing and updating information on a rewritable medium with the writing constraints: During the odd (resp., even) cycles of updating information, the encoder can only write 1's (resp., 0's) in selected bit positions of WUMs, and not change the contents of other positions. In this correspondence, motivated by the research works of Wolf, Wyner, Ziv, and Körner on write-once memories (WOMs), we study the problem of how to reuse a WUM for fixed T successive cycles with nonperiodic codes (i.e., all coding strategies are permitted for every cycle). For the situation where the encoder knows and the decoder does not know the previous content of the memory, we determine the zero-error capacity region, the average capacity, and the maximum total number of information bits stored in the WUM for fixed T successive cycles. Motivated by the research works of Heegard on WOMs with symmetric input noise, we introduce two models of WUMs with symmetric or asymmetric input noise. By using ε-error as performance criterion, we extend the above results for WUMs to the two models of WUMs with symmetric or asymmetric input noise.**

***Index Terms*—Capacity, capacity region, codes, noisy WUMs, reusable memories, write-unidirectional memories.**

I. Introduction

Write-unidirectional memories (WUMs) were introduced by Willems, Vinck [1] and Borden [2] as an information-theoretic model for storing and updating information on a rewritable medium with the writing constraints: During the odd (resp., even) cycles of updating information, the encoder can only write 1's (resp., 0's) in selected bit positions of WUMs, and not change the contents of other positions. The writing constraints are due to the storage technology in the 1980s for a class of rewritable digital disks (see [27]). The capacity and code constructions of WUMs have been extensively studied since then. For research works on this topic, see [1]–[14]. These research works help us to have a better understanding on rewritable storage medium from an information-theoretic point of view.

One can consider the previous content of the memory as side information available to the encoder and/or the decoder. By using zero-error and ε-error as performance criteria, the authors of [1]–[14] studied the capacity problems of WUMs in the following four cases:

(E_+, D_+): both the encoder and decoder know the previous content of the memory;
(E_+, D_-): the encoder knows and the decoder does not know the previous content of the memory;
(E_-, D_+): the encoder does not know and the decoder knows the previous content of the memory;
(E_-, D_-): both the encoder and decoder do not know the previous content of the memory.

In this correspondence, we only consider the case (E_+, D_-). In the sequel, all definitions, notations, and results are described for this case. For the situation where the encoder and decoder use the same coding scheme (period-1 codes) for all cycles, Borden [2] showed that the capacity of WUMs is $\log[(\sqrt{5}+1)/2]$. For the situation where the encoder and decoder use two coding schemes (period-2 codes) alternatively for odd and even cycles, Overveld [4] determined the capacity region of WUMs. The WUMs are closely related to write-once memories (WOMs) defined by Rivest and Shamir [20] in 1982. For research works on the WOMs, see [20]–[26]. Wolf, Wyner, Ziv, and Körner [22] studied the WOMs from an information-theoretic point of view. They determined the capacity region and the maximum total number of information bits stored in the WOMs for fixed T successive cycles by using arbitrary codes for every cycle. In this correspondence, motivated by the research works of Wolf, Wyner, Ziv, and Körner on the WOMs, we study the problem of how to reuse a WUM for fixed T successive cycles by using nonperiodic codes, i.e., all coding strategies are permitted for every cycle. We determine the zero-error capacity region and the average capacity for the WUMs. We show that the maximum total number of information bits stored in one storage cell of the WUMs for fixed T successive cycles is given by $C(T) = \log a_T$, where a_T is the Tth Fibonacci number.

Heegard [16] studied the WOMs with symmetric input noise. He determined the ε-error capacity region and the maximum total number of information bits stored in the memories for fixed T successive cycles by using arbitrary codes for every cycle. Kuznetsov [9] investigated the WUMs with symmetric output noise. By using the same coding scheme (period-1 codes) for every cycle, he presented a lower bound for the zero-error capacity of the WUMs with symmetric output noise. In this correspondence, motivated by the research works of Heegard on the WOMs with symmetric input noise, we introduce two models of WUMs with symmetric or asymmetric input noise. By using ε-error as performance criterion, we determine the ε-error capacity region, the average capacity, and the maximum total number of information bits stored in the memories for fixed T successive cycles by using arbitrary codes for every cycle.

The rest of this correspondence is organized as follows. In Section II, we present the basic definitions and notations, and the main results for the WUMs. In Sections III and IV, we present the proofs for the main results of the WUMs. In Section V, we discuss the model of WUMs. An equivalent definition of WUM codes is given. In Section VI, from an information-theoretic point of view, we discuss the model of reusable memories. Then, we introduce two specific classes of reusable memories, WUMs with symmetric or asymmetric input noise. Main results on the capacity of WUMs with symmetric or asymmetric input noise are given. In Section VII, we present the proofs of the main results for the WUMs with symmetric or asymmetric input noise. Finally, we conclude the correspondence in Section VIII.

II. WUMs With Nonperiodic Codes

Assume that a WUM consists of n storage cells. The initial state of all the storage cells is "0." We want to reuse the WUM for T successive

Manuscript received January 4, 2002; revised September 30, 2003. This work was supported in part by the DSTA project (POD 0103223), the National Natural Science Foundation of China under Grant 60172060, the Trans-Century Training Program Foundation for the Talents of the Education Ministry of China, and the Foundation for University Key Teacher of the Education Ministry of China. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Ulm, Germany, June/July 1997.

F.-W. Fu is with the Temasek Laboratories, National University of Singapore, Singapore 119260, Republic of Singapore, on leave from the Department of Mathematics, Nankai University, Tianjin 300071, China (e-mail: tslfufw@nus.edu.sg).

A. J. Han Vinck is with the Institute for Experimental Mathematics, University of Duisburg-Essen, 45326 Essen, Germany (e-mail: vinck@exp-math.uni-essen.de).

V. K. Wei and R. W. Yeung are with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong (e-mail: kwwei@ie.cuhk.edu.hk; whyeung@ie.cuhk.edu.hk).

Communicated by P. Narayan, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2004.825047

cycles. For every cell of the memory, the encoder has three choices: write a 1, write a 0, or leave the content of the cell unchanged. Hence, the action of the encoder at this cell can be described by a symbol over the alphabet $\{0, 1, \triangle\}$ where $\triangle$ refers to the case when the content of the cell is left unchanged. If the previous content of the cell is $y_- \in \{0, 1\}$, the input (encoder's action) of the cell is $x \in \{0, 1, \triangle\}$, then the new content of the cell $y \in \{0, 1\}$ is given by $y = \Phi(x, y_-)$, where the writing function

$$\Phi: \{0, 1, \triangle\} \times \{0, 1\} \longmapsto \{0, 1\}$$

is defined by Simonyi [3] as follows:

$$\Phi(x, y_-) = y = \begin{cases} y_-, & \text{if } x = \triangle, \\ x, & \text{if } x = 0, 1. \end{cases} \quad (1)$$

Note that for the odd (resp., even) cycles, the input (encoder's action) of the cell is "1" (resp., "0") or "$\triangle$." Below we give the definitions of WUM codes, the capacity region, the average capacity, and the maximum total number of information bits stored in the memories for fixed T successive cycles. We only consider the following case: The encoder knows and the decoder does not know the previous content of the memory. The encoder and decoder can use arbitrary codes for every cycle, and there are no decoding errors (zero-error case).

For

$$x^n = (x_1, \dots, x_n) \in \{0, 1, \triangle\}^n$$
$$y^n = (y_1, \dots, y_n) \in \{0, 1\}^n$$

we denote

$$\Phi(x^n, y^n) = (\Phi(x_1, y_1), \dots, \Phi(x_n, y_n)).$$

For

$$a^n = (a_1, \dots, a_n) \in \{0, 1\}^n$$
$$b^n = (b_1, \dots, b_n) \in \{0, 1\}^n$$

we say that $a^n \le b^n$, if $a_i \le b_i, \forall\, i = 1, \dots, n$. Denote the message index set at the tth cycle as $I_t = \{1, \dots, M_t\}, t = 1, 2, \dots, T$.

Definition 1: An $(n, M_1, \dots, M_T)$ code consists of T pairs of encoding and decoding functions

$$f_t: I_t \times \{0, 1\}^n \longmapsto \begin{cases} \{\triangle, 1\}^n, & \text{if } t \text{ is odd} \\ \{\triangle, 0\}^n, & \text{if } t \text{ is even} \end{cases}$$
$$g_t: \{0, 1\}^n \longmapsto I_t, \qquad t = 1, 2, \dots, T$$

such that for every $t = 1, 2, \dots, T$ and $m_t \in I_t$, $g_t(y_t^n) = m_t$ where $y_0^n = \mathbf{0}$ (the zero vector of length n) and $x_t^n = f_t(m_t, y_{t-1}^n)$, $y_t^n = \Phi(x_t^n, y_{t-1}^n)$, $t = 1, \dots, T$.

Note that for t odd, $y_{t-1}^n \le y_t^n$ and for t even, $y_t^n \le y_{t-1}^n$. Denote

$$R_t = \frac{1}{n} \log M_t, \qquad t = 1, \dots, T.$$

The T-tuple $(R_1, \dots, R_T)$ is called the rate vector of this code. The closure of the set of all rate-vectors $\boldsymbol{\mathcal{C}}_T$ is called the capacity region of the WUMs. The maximum total number of information bits stored in one storage cell of the WUMs during the T updating cycles is defined by

$$C(T) = \max\left\{ \sum_{t=1}^{T} R_t \middle| (R_1, R_2, \dots, R_T) \in \boldsymbol{\mathcal{C}}_T \right\}.$$

The average capacity of the WUMs is defined by

$$C = \lim_{T\to\infty} \frac{C(T)}{T}.$$

Let $h(x)$ be the binary entropy function, i.e.,

$$h(x) = -x \log x - (1 - x) \log(1 - x), \qquad x \in [0, 1].$$

Let $\boldsymbol{\mathcal{R}}_T$ be the closure of the set consisting of all T-tuples $(R_1, R_2, \dots, R_T)$ satisfying the following condition: There exist $0 \le p_1, \dots, p_T \le 1$ such that

$$0 \le R_t \le q_{t-1} h(p_t)$$

where

$$q_0 = 1, \quad q_t = 1 - p_t q_{t-1}, \qquad t = 1, \dots, T.$$

We now state our main results for the WUMs as the following theorems.

Theorem 1: $\boldsymbol{\mathcal{C}}_T = \boldsymbol{\mathcal{R}}_T$.

Theorem 2: $C(T) = \log a_T$, where a_T is the Tth Fibonacci number, i.e., $a_{-1} = a_0 = 1, a_i = a_{i-1} + a_{i-2}, i \ge 1$.

As a direct consequence of Theorem 2, we have the following.

Corollary 1: $C = \log \frac{\sqrt{5}+1}{2}$.

Corollary 1 shows that the average capacity is equal to the capacity of period-1 codes. The proof of Theorem 1 will be given in Section III. In the proof, we can see that $\boldsymbol{\mathcal{R}}_T$ is a convex set. The proof of Theorem 2 will be given in Section IV.

III. Proof of Theorem 1

In this section, we modify the methods of Wolf, Wyner, Ziv, and Körner [22] for proving coding theorems of the WOMs to prove Theorem 1.The proof of Theorem 1 is divided into two parts: 1) proof of the direct part of Theorem 1, and 2) proof of the converse part of Theorem 1.

A. Proof of the Direct Part of Theorem 1

The Direct Part of Theorem 1: For $0 \le p_1, \dots, p_T \le 1$ let $q_0 = 1$, $q_t = 1 - p_t q_{t-1}, 1 \le t \le T$. For any fixed $\delta > 0$ (very small), set

$$M_t = 2^{n[q_{t-1} h(p_t) - \delta]}, \qquad t = 1, 2, \dots, T.$$

For n sufficiently large, there exists an $(n, M_1, \dots, M_T)$ code for the WUMs.

Remark: $M_1, \dots, M_T$ are assumed to be positive integers. This assumption does not affect the asymptotic behavior of the code.

Proof: Let $w_t = n q_t$ for t odd and $w_t = n(1 - q_t)$ for t even. We assume that $w_1, w_2, \dots, w_T$ are positive integers. This assumption does not affect the asymptotic behavior of the following code construction. For every $t = 1, 2, \dots, T$, denote

$$B(w_t) = \left\{ x \in \{0, 1\}^n | \boldsymbol{w}_H(x) = w_t \right\}$$

where $\boldsymbol{w}_H(\cdot)$ is the Hamming weight of the binary vector. For every $t = 1, 2, \dots, T$, let $\{A_m^t\}_{m=1}^{M_t}$ be a partition of $B(w_t)$, i.e.,

$$A_m^t \cap A_{m'}^t = \emptyset, \quad m \ne m', \text{ and } \bigcup_{m=1}^{M_t} A_m^t = B(w_t).$$

There exists an $(n, M_1, \dots, M_T)$ code for the WUMs, if this set of partitions satisfies the following property.

Property ♣ : For every $t = 1, 2, \dots, T$, every $u \in B(w_{t-1})$, and every $m \in I_t = \{1, \dots, M_t\}$, there exists a vector $y(m, u) \in A_m^t$ such that $u \le y(m, u)$, for t odd, and $y(m, u) \le u$, for t even.

The encoding and decoding functions can be defined as follows. Because $u \le y(m, u)$, for t odd, and $y(m, u) \le u$, for t even, we can choose $x(m,u) \in \{\triangle, 1\}^n$ for t odd, and $x(m,u) \in \{\triangle, 0\}^n$ for t even, such that $y(m,u) = \Phi(x(m,u), u)$. We define the encoding function $f_t(m,u) = x(m,u)$ and the decoding function $g_t(y) = m$, if $y \in A_m^t$. Below we show that there exists such a set of partitions, by using a random coding method. With every $b \in B(w_t)$, we connect a random index r_b, which is uniformly distributed over the message set $I_t = \{1, \ldots, M_t\}$, and all these random indexes are independent. Define

$$A_m^t = \{b \in B(w_t) | r_b = m\}, \qquad m = 1, 2, \ldots, M_t.$$

Then $\{A_m^t\}_{m=1}^{M_t}$ forms a random partition of $B(w_t)$. For a fixed $u \in B(w_{t-1})$, set

$$G_t(u) = \{y \in B(w_t) | u \le y\}, \qquad t \text{ is odd}$$
$$G_t(u) = \{y \in B(w_t) | y \le u\}, \qquad t \text{ is even.}$$

If t is odd, then $w_{t-1} = n(1-q_{t-1})$ and $w_t = nq_t = n - np_t q_{t-1}$. Therefore,

$$k_t = |G_t(u)| = \binom{n - w_{t-1}}{w_t - w_{t-1}} = \binom{nq_{t-1}}{np_t q_{t-1}} = 2^{n[q_{t-1}h(p_t) + \boldsymbol{o}(n)]}$$

where $\boldsymbol{o}(n) \to 0$, for $n \to \infty$.

If t is even, then $w_{t-1} = nq_{t-1}$ and $w_t = n(1-q_t) = np_t q_{t-1}$. Therefore,

$$k_t = |G_t(u)| = \binom{w_{t-1}}{w_t} = \binom{nq_{t-1}}{np_t q_{t-1}} = 2^{n[q_{t-1}h(p_t) + \boldsymbol{o}(n)]}.$$

Fix $m \in I_t = \{1, \ldots, M_t\}$ and $u \in B(w_{t-1})$, then

$$\begin{aligned} Pr\left\{A_m^t \cap G_t(u) = \emptyset\right\} &= Pr\left\{\text{for every } b \in G_t(u), r_b \ne m\right\} \\ &= \left[1 - \frac{1}{M_t}\right]^{k_t} \\ &\le \exp\{-k_t/M_t\} \\ &= \exp\{-2^{n[\delta - \boldsymbol{o}(n)]}\}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\Pr\left\{A_m^t \cap G_t(u) = \emptyset \text{ for some } 1 \le t \le T, m \in I_t, u \in B(w_{t-1})\right\} \\ &\quad \le T \cdot 2^{2n} \cdot \exp\{-2^{n[\delta - \boldsymbol{o}(n)]}\} \to 0, \qquad n \to \infty. \end{aligned}$$

This implies that when n is sufficiently large, there exists a set of partitions $\{A_m^t\}_{m=1}^{M_t}$ of $B(w_t)$, $t = 1, 2, \ldots, T$. This set of partitions satisfies Property ♣. This completes the proof of the direct part of Theorem 1. □

B. Proof of the Converse Part of Theorem 1

The Converse Part of Theorem 1: For an $(n, M_1, \ldots, M_T)$ WUM code, the rate vector

$$\left(\frac{1}{n}\log M_1, \frac{1}{n}\log M_2, \ldots, \frac{1}{n}\log M_T\right) \in \boldsymbol{\mathcal{R}}_T. \tag{2}$$

Proof: Let $W_1, W_2, \ldots, W_T$ be independent random variables, uniformly distributed over the message set $I_t = \{1, \ldots, M_t\}$, $t = 1, 2, \ldots, T$, respectively. Let $\{(f_t, g_t)\}_{t=1}^T$ be the T pairs of encoding and decoding functions for the $(n, M_1, \ldots, M_T)$ code. Denote $Y_0^n = \mathbf{0}$ (the zero vector) and for $t = 1, \ldots, T$

$$X_t^n = (X_{t,1}, X_{t,2}, \ldots, X_{t,n}) = f_t\left(W_t, Y_{t-1}^n\right)$$
$$Y_t^n = (Y_{t,1}, Y_{t,2}, \ldots, Y_{t,n}) = \Phi\left(X_t^n, Y_{t-1}^n\right).$$

It follows that $g_t(Y_t^n) = W_t$ for $t = 1, \ldots, T$ and

$$Y_{t,i} = \Phi\left(X_{t,i}, Y_{t-1,i}\right), \qquad i = 1, 2, \ldots, n.$$

Since W_t and Y_{t-1}^n are independent, we have

$$\begin{aligned} H(W_t) &= H\left(W_t | Y_{t-1}^n\right) \\ &= H\left(W_t, Y_{t-1}^n | Y_{t-1}^n\right) \\ &\overset{(a)}{\ge} H\left(X_t^n | Y_{t-1}^n\right) = H\left(X_t^n, Y_{t-1}^n | Y_{t-1}^n\right) \\ &\overset{(b)}{\ge} H\left(Y_t^n | Y_{t-1}^n\right) \\ &\overset{(c)}{\ge} H\left(W_t | Y_{t-1}^n\right) = H(W_t), \end{aligned}$$

where (a) follows from the fact that $X_t^n = f_t\left(W_t, Y_{t-1}^n\right)$, and (b) follows from the fact that $Y_t^n = \Phi\left(X_t^n, Y_{t-1}^n\right)$, and (c) follows from the fact that $g_t\left(Y_t^n\right) = W_t$. Therefore,

$$H(W_t) = H\left(Y_t^n | Y_{t-1}^n\right), \qquad t = 1, 2, \ldots, T.$$

This implies that

$$\begin{aligned} \frac{1}{n}\log M_t = \frac{1}{n}H(W_t) &= \frac{1}{n}H\left(Y_t^n | Y_{t-1}^n\right) \\ &\le \frac{1}{n}\sum_{i=1}^n H\left(Y_{t,i} | Y_{t-1,i}\right). \end{aligned} \tag{3}$$

For a fixed $i = 1, 2, \ldots, n$, set $q_{0,i} = 1$ and

$$q_{t,i} = \begin{cases} \Pr\{Y_{t,i} = 1\}, & \text{if } t \text{ is odd} \\ \Pr\{Y_{t,i} = 0\}, & \text{if } t \text{ is even} \end{cases}$$
$$p_{t,i} = \begin{cases} \Pr\{X_{t,i} = \triangle | Y_{t-1,i} = 0\}, & \text{if } t \text{ is odd} \\ \Pr\{X_{t,i} = \triangle | Y_{t-1,i} = 1\}, & \text{if } t \text{ is even.} \end{cases}$$

Note that $X_{t,i} \in \{\triangle, 1\}$ for t odd and $X_{t,i} \in \{\triangle, 0\}$ for t even, and $Y_{t,i} = \Phi(X_{t,i}, Y_{t-1,i})$. It is not hard to show that for every $t = 1, 2, \ldots, T$

$$q_{t,i} = 1 - p_{t,i}q_{t-1,i} \tag{4}$$
$$H(Y_{t,i} | Y_{t-1,i}) = q_{t-1,i}h(p_{t,i}). \tag{5}$$

By (3) and (5), we have

$$\frac{1}{n}\log M_t \le \frac{1}{n}\sum_{i=1}^n q_{t-1,i}h(p_{t,i}). \tag{6}$$

Denote

$$q_t = \frac{1}{n}\sum_{i=1}^n q_{t,i}, \qquad t = 0, 1, \ldots, T$$
$$p_t = \frac{\sum_{i=1}^n p_{t,i}q_{t-1,i}}{\sum_{i=1}^n q_{t-1,i}} = \frac{\sum_{i=1}^n p_{t,i}q_{t-1,i}}{nq_{t-1}}, \qquad t = 1, 2, \ldots, T.$$

Then, $q_0 = 1$ since $q_{0,i} = 1$ for every $i = 1, 2, \ldots, n$. It follows from (4) that

$$q_t = \frac{1}{n}\sum_{i=1}^n q_{t,i} = 1 - \frac{1}{n}\sum_{i=1}^n p_{t,i}q_{t-1,i} = 1 - p_t q_{t-1}. \tag{7}$$

Note that the binary entropy function $h(x)$ is a convex function. By (4)

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n q_{t-1,i}h(p_{t,i}) &= q_{t-1} \cdot \left[\sum_{i=1}^n \frac{q_{t-1,i}}{nq_{t-1}}h(p_{t,i})\right] \\ &\le q_{t-1}h\left(\frac{\sum_{i=1}^n p_{t,i}q_{t-1,i}}{nq_{t-1}}\right) \\ &= q_{t-1}h(p_t). \end{aligned} \tag{8}$$

By (6) and (8), we have

$$\frac{1}{n}\log M_t \le q_{t-1} h(p_t). \tag{9}$$

Hence, (2) follows from (7) and (9). This completes the proof. □

Remark: By using the same argument deriving (7) and (8) in the above proof, one can prove that the capacity region $\boldsymbol{\mathcal{R}}_T$ is a convex set.

IV. Proof of Theorem 2

Lemma 1: Suppose r is a real number. For $x \in [0,1]$, we have

$$h(x) + rx \le \log(1+2^r)$$

where equality holds only when $x = \frac{2^r}{1+2^r}$.

Theorem 2 can be easily derived from Lemma 2 below. Lemma 2 also gives the relationship between the Fibonacci sequences and the binary entropy function.

Lemma 2: For $0 \le p_1, \dots, p_T \le 1$, let $q_0 = 1$, $q_t = 1 - p_t q_{t-1}$, $t = 1, \dots, T$. Then

$$\sum_{t=1}^{T} q_{t-1} h(p_t) \le \log a_T$$

where a_T is the Tth Fibonacci number, i.e.,

$$a_{-1} = a_0 = 1, \quad a_i = a_{i-1} + a_{i-2}, \qquad i \ge 1.$$

The equality holds iff

$$p_{T-i} = \frac{a_{i-1}}{a_{i+1}}, \qquad i = 0, 1, \dots, T-1.$$

Proof: Actually, we can prove that for every $m = 1, \dots, T$

$$\sum_{t=T-m+1}^{T} q_{t-1} h(p_t) \le q_{T-m} \log \frac{a_m}{a_{m-1}} + \log a_{m-1} \tag{10}$$

where equality holds iff $p_{T-i} = \frac{a_{i-1}}{a_{i+1}}$, $i = 0, 1, \dots, m-1$.

If we take $m = T$ in (10), we obtain Lemma 2. Below we prove (10) by induction on m. Clearly, $q_{T-1} h(p_T) \le q_{T-1}$, and equality holds iff $p_T = 1/2$. This means that (10) holds for $m = 1$. Assume that (10) holds for $m = k$. We want to prove that (10) holds for $m = k+1$.

$$\begin{aligned}
&\sum_{t=T-k}^{T} q_{t-1} h(p_t) \\
&\stackrel{1)}{=} q_{T-k-1} h(p_{T-k}) + \sum_{t=T-k+1}^{T} q_{t-1} h(p_t) \\
&\stackrel{2)}{\le} q_{T-k-1} h(p_{T-k}) + q_{T-k} \log \frac{a_k}{a_{k-1}} + \log a_{k-1} \\
&\stackrel{3)}{=} q_{T-k-1} \left[h(p_{T-k}) + p_{T-k} \log \frac{a_{k-1}}{a_k} \right] + \log a_k \\
&\stackrel{4)}{\le} q_{T-k-1} \log \left(1 + \frac{a_{k-1}}{a_k} \right) + \log a_k \\
&\stackrel{5)}{=} q_{T-k-1} \log \frac{a_{k+1}}{a_k} + \log a_k.
\end{aligned}$$

Step 2) follows from the induction assumption; Step 3) follows from the fact that $q_{T-k} = 1 - p_{T-k} q_{T-k-1}$; Step 4) follows from Lemma 1; Step 5) follows from the fact that a_k is the kth Fibonacci number. From the induction assumption and Lemma 1, we can easily see that the equality holds iff $p_{T-i} = a_{i-1}/a_{i+1}$, $i = 0, 1, \dots, k$. This implies that (10) holds for $m = k+1$. This completes the proof. □

Remark: By carefully reading [3], one can see that Simonyi and Tardos actually presented a combinatorial proof for the following result: If there exists an $(n, M_1, \dots, M_T)$ WUM code, then $\prod_{i=1}^{T} M_i \le (a_T)^n$, where a_T is the Tth Fibonacci number. Our proof of Theorem 2 is an information-theoretic proof. It also holds for ε-error codes.

V. Discussion of the WUM Model

From Theorem 1, we see that the expression of the capacity region of WUM is uniform for the odd and even cycles. This phenomenon can be explained by the following facts which were first observed by Willems [10].

We use $(t)_2$ to denote the residue number of t modulo 2, i.e., $(t)_2 = 0$, if t is even and $(t)_2 = 1$, if t is odd. For a particular storage cell in the WUM at the tth cycle, we denote its input as $x(t) \in \{\triangle, (t)_2\}$, the previous content (state) as $y(t-1)$, and the new content as $y(t) = \Phi(x(t), y(t-1))$. The writing function $\Phi(\cdot,\cdot)$ is defined by (1) in Section I. Define

$$\overline{x}(t) = \begin{cases} 1, & \text{if } x(t) = \triangle \\ 0, & \text{if } x(t) \ne \triangle \end{cases}$$

and

$$\overline{y}(t) = (t)_2 - y(t)$$

where the operations are over the binary additive group $\{0,1\}$, i.e., $a + b = (a+b)_2$. From the fact that $y(t) = \Phi(x(t), y(t-1))$, we know that $\overline{y}(t) = \Omega(\overline{x}(t), \overline{y}(t-1))$, where

$$\Omega : \{0,1\} \times \{0,1\} \longmapsto \{0,1\}$$

is defined by

$$\Omega(x,s) = \begin{cases} 0, & x = 0, \ s = 0 \\ 1, & x = 1, \ s = 0 \\ 0, & x = 0, \ s = 1 \\ 0, & x = 1, \ s = 1. \end{cases}$$

Note that at the tth cycle, there is a one-to-one correspondence between $x(t)$ and $\overline{x}(t)$, and as well as between $y(t)$ and $\overline{y}(t)$. We can obtain $x(t)$ from $\overline{x}(t)$ by

$$x(t) = \begin{cases} \triangle, & \text{if } \overline{x}(t) = 1 \\ 0, & \text{if } \overline{x}(t) = 0 \text{ and } t \text{ is even} \\ 1, & \text{if } \overline{x}(t) = 0 \text{ and } t \text{ is odd} \end{cases}$$

and $y(t)$ from $\overline{y}(t)$ by $y(t) = (t)_2 - \overline{y}(t)$. From the above analysis, we can redefine a WUM code as follows.

Definition 2: An $(n, M_1, \dots, M_T)$ code consists of T pairs of encoding and decoding functions

$$\begin{aligned} f_t &: \{1, 2, \dots, M_t\} \times \{0,1\}^n \longmapsto \{0,1\}^n, \\ g_t &: \{0,1\}^n \longmapsto \{1, 2, \dots, M_t\}, \qquad t = 1, 2, \dots, T \end{aligned}$$

such that for every $t = 1, 2, \dots, T$ and $m_t \in \{1, 2, \dots, M_t\}$, $g_t(y_t^n) = m_t$ where $y_0^n = \mathbf{0}$ and $x_t^n = f_t(m_t, y_{t-1}^n)$, $y_t^n = \Omega(x_t^n, y_{t-1}^n)$, $t = 1, \dots, T$.

Note that in this modified definition, the image space of the encoding functions is always $\{0,1\}^n$, which does not depend on the even or odd cycles (see Definition 1 in Section II). Heegard [16] studied the deterministic reusable memories (named as deterministic WOMs), and determined the ε-error capacity region. By the above analysis, we see that the model of WUMs is equivalent to a specific class of deterministic reusable memories. From the result of Heegard [16] and the above analysis, we can also determine the ε-error capacity region for the WUMs, which is the same as the zero-error capacity region we established here.

VI. THE INFORMATION-THEORETIC MODELS OF REUSABLE MEMORIES AND WUMS WITH SYMMETRIC OR ASYMMETRIC INPUT NOISE

In this section, we discuss the general information-theoretic model of reusable memories, which was studied by Heegard [16] (named as the noisy WOMs) and Ahlswede and Simonyi [14]. The model of WOMs with symmetric input noise, introduced and studied by Heegard [16], is a specific class of reusable memories. Heegard [16] determined the ε-error capacity region and the maximum total number of information bits stored in the WOMs with symmetric input noise for fixed T successive cycles by using arbitrary codes for every cycle. Motivated by the research works of Heegard on the WOMs with symmetric input noise, we introduce two models of WUMs with symmetric or asymmetric input noise which are two specific classes of reusable memories. By using ε-error as performance criterion, we determine the ε-error capacity region, the average capacity, and the maximum total number of information bits stored in the memories for fixed T successive cycles by using arbitrary codes for every cycle.

Heegard [16] as well as Ahlswede and Simonyi [14] proposed the general information-theoretic model of reusable memories with random errors. A probabilistic channel is introduced to describe the state transition mechanism of storage cells in the reusable memories. This general model of reusable memories is highly related to the model of computer memories with defects (see [17]–[19]), which is characterized by a channel with random parameters. Heegard [16] studied the capacity of reusable memories for the situation where the encoder knows and the decoder does not know the previous content of the memories. He presented an inner bound for the ε-error capacity region. Furthermore, he showed that this inner bound is tight for the model of WOMs with symmetric input noise.

Let $\mathcal{X}, \mathcal{Y}$ be two finite sets, called the input and output alphabets. Let $\mathcal{S} = \mathcal{Y}$ be the state alphabet. Let $\boldsymbol{Q} = \{q(y) : y \in \mathcal{Y}\}$ be a probability distribution on $\mathcal{Y}$, called the initial state distribution. Let

$$\boldsymbol{W} = \{W(y|x, s) : x \in \mathcal{X}, s \in \mathcal{Y}, y \in \mathcal{Y}\}$$

be a conditional probability distribution from $\mathcal{X} \times \mathcal{Y}$ to $\mathcal{Y}$.

Assume that the reusable memory consists of n storage cells. Its initial state is

$$s^n = (s_1, s_2, \dots, s_n) \in \mathcal{Y}^n$$

with the probability $\prod_{i=1}^n q(s_i)$. In the process of updating information, if its state is

$$s^n = (s_1, s_2, \dots, s_n) \in \mathcal{Y}^n$$

and its input is

$$x^n = (x_1, x_2, \dots, x_n) \in \mathcal{X}^n$$

then its output is

$$y^n = (y_1, y_2, \dots, y_n) \in \mathcal{Y}^n$$

with the probability $\prod_{i=1}^n W(y_i|x_i, s_i)$. The output y^n also becomes the new state of the memory for the next cycle. Assume that we want to reuse the reusable memory for T successive cycles.

An $(n, M_1, \dots, M_T, \delta)$ code consists of T pairs of encoding and decoding functions

$$f_t : \{1, 2, \dots, M_t\} \times \mathcal{Y}^n \longmapsto \mathcal{X}^n$$
$$g_t : \mathcal{Y}^n \longmapsto \{1, 2, \dots, M_t\}, \qquad t = 1, 2, \dots, T.$$

Let $W_1, W_2, \dots, W_T$ be a sequence of independent messages with W_t uniformly distributed over the set $\{1, 2, \dots, M_t\}$. We obtain a sequence of random vectors $Y_0^n, X_1^n, Y_1^n, \dots, X_T^n, Y_T^n$: The initial state random vector Y_0^n is generated according to the product distribution

$$\Pr\{Y_0^n = s^n\} = \prod_{i=1}^n q(s_i), \qquad s^n \in \mathcal{Y}^n.$$

At the tth cycle, the input random vector is given by

$$X_t^n = f_t(W_t, Y_{t-1}^n)$$

where Y_{t-1}^n is the output random vector at the $(t-1)$th cycle, i.e., the state random vector at the tth cycle. The output random vector $Y^n(t)$ at the tth cycle is generated by the conditional distribution $\boldsymbol{W}^n$ through $X^n(t)$ and $Y^n(t-1)$ as follows:

$$\Pr\Big\{Y^n(t) = y^n | X^n(t) = x^n,\ Y^n(t-1) = s^n\Big\} = \prod_{i=1}^n W(y_i|x_i, s_i).$$

The output random vector $Y^n(t)$ also becomes the state random vector at the $(t+1)$th cycle. Let $\bar{W}_t = g_t(Y^n(t))$, which is the reproduction of the random message W_t. The tth probability of error is defined as

$$P_e(t) \equiv \Pr\Big\{\bar{W}_t \neq W_t\Big\}.$$

The (worst case) probability of error for this code is defined as

$$\delta \equiv \max_{1 \le t \le T} P_e(t).$$

A rate T-tuple $(R_1, R_2, \dots, R_T)$ is said to be ε-achievable if for any $\varepsilon > 0$ there exists an $(n, M_1, \dots, M_T, \delta)$ code for some n, such that $R_t = (1/n) \log M_t$ for $t = 1, 2, \dots, T$ and $\delta < \varepsilon$. The closure of the set of all ε-achievable rate vectors $\boldsymbol{\mathcal{C}}_T^*$ is called the ε-error capacity region. The maximum total number of information bits stored in one storage cell of the memories during the T updating cycles is defined as

$$C^*(T) = \max\left\{\sum_{t=1}^T R_t | (R_1, R_2, \dots, R_T) \in \boldsymbol{\mathcal{C}}_T^*\right\}.$$

The average capacity of the reusable memories is defined as

$$C = \limsup_{T \to \infty} \frac{C^*(T)}{T}.$$

Heegard [16] presented an inner bound for the ε-error capacity region as follows.

Theorem 3: (An achievable rate region) Fix T auxiliary alphabets $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_T$ and choose T conditional distributions

$$\boldsymbol{P}_t = \left\{P_t(u_t, x_t|y_{t-1}) : u_t \in \mathcal{U}_t, x_t \in \mathcal{X}, y_{t-1} \in \mathcal{Y}\right\}$$

for $t = 1, 2, \dots, T$. Let the joint distribution of the random variables $(Y_0, U_1, X_1, Y_1, \dots, U_T, X_T, Y_T)$ take the form

$$q(y_0) \prod_{t=1}^T P_t(u_t, x_t|y_{t-1}) W(y_t|x_t, y_{t-1}).$$

Then $(R_1, R_2, \dots, R_T) \in \boldsymbol{\mathcal{C}}_T^*$ if for $t = 1, 2, \dots, T$

$$0 \le R_t \le I(U_t; Y_t) - I(U_t; Y_{t-1}).$$

The two models of WUMs with symmetric or asymmetric input noise can be described as two specific classes of reusable memories as follows. The input alphabet, the state alphabet, and the output alphabet are given by

$$\mathcal{X} = \mathcal{S} = \mathcal{Y} = \{0, 1\}.$$

The initial state distribution of the memories is given by

$$q(0) = 1 - \beta, \qquad q(1) = \beta.$$

For a particular storage cell in the memories, assume that its state (previous content) is $s \in \{0, 1\}$, its input is $x \in \{0, 1\}$. Due to the symmetric or asymmetric input noise, the input $x \in \{0, 1\}$ may be changed to $x' \in \{0, 1\}$. The conditional probability from x to x' is given by

1) *WUMs with symmetric input noise*

$$W_s(0|0) = W_s(1|1) = 1 - \alpha$$

$$W_s(1|0) = W_s(0|1) = \alpha;$$

2) *WUMs with asymmetric input noise*

$$W_a(0|0) = 1-\alpha, \quad W_a(1|0) = \alpha$$
$$W_a(0|1) = 0, \quad W_a(1|1) = 1.$$

The output of this storage cell is $y = \Omega(x', s)$ where

$$\Omega : \{0,1\} \times \{0,1\} \longmapsto \{0,1\}$$

is defined in Section V. It is easy to see that the conditional probability $W(y|x,s)$ from (x,s) to y for these two classes of reusable memories is given by

1) *WUMs with symmetric input noise*

$$W(\cdot \mid \cdot, s=0) = \begin{pmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{pmatrix} \tag{11}$$

$$W(\cdot \mid \cdot, s=1) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}; \tag{12}$$

2) *WUMs with asymmetric input noise*

$$W(\cdot \mid \cdot, s=0) = \begin{pmatrix} 1-\alpha & \alpha \\ 0 & 1 \end{pmatrix} \tag{13}$$

$$W(\cdot \mid \cdot, s=1) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}. \tag{14}$$

For $0 \le a, b \le 1$, denote $a * b = a(1-b) + b(1-a)$. Let $\boldsymbol{\mathcal{R}}_T^{(s)}$ be the closure of the set consisting of all T-tuples $(R_1, R_2, \ldots, R_T)$ satisfying the following condition: There exist $0 \le p_1, \ldots, p_T \le 1$ such that

$$0 \le R_t \le \overline{q}_{t-1}\left[h(p_t * \alpha) - h(\alpha)\right], \qquad t = 1, \ldots, T$$

where

$$\overline{q}_0 = 1-\beta, \quad \overline{q}_t = 1 - (p_t * \alpha)\overline{q}_{t-1}, \qquad t = 1, \ldots, T.$$

We now state our main results for the WUMs with symmetric input noise as the following theorems.

Theorem 4: For the WUMs with symmetric input noise, the ε-error capacity region is given by $\boldsymbol{\mathcal{C}}_T^{(s)} = \boldsymbol{\mathcal{R}}_T^{(s)}$.

Theorem 5: For the WUMs with symmetric input noise, the maximum total number of information bits stored in one storage cell of the memories during the T updating cycles is given by

$$C_s(T) = (1-\beta)\log\theta_{T+1} + \beta\log\theta_T$$

where the sequences $\{\theta_i\}_{i=1}^{\infty}$ satisfy the recursive equation

$$\theta_i = 2^{-h(\alpha)}[\theta_{i-1} + \theta_{i-2}], \quad \theta_1 = 1, \qquad \theta_2 = 2^{1-h(\alpha)}.$$

As a direct consequence of Theorem 5, we have the following.

Corollary 2: For the WUMs with symmetric input noise, the average capacity is given by

$$C_s = \log\left[\frac{1 + \sqrt{1 + 4 \cdot 2^{h(\alpha)}}}{2}\right] - h(\alpha).$$

Let $\boldsymbol{\mathcal{R}}_T^{(a)}$ be the closure of the set consisting of all T-tuples $(R_1, R_2, \ldots, R_T)$ satisfying the followng condition: There exist $0 \le p_1, \ldots, p_T \le 1$ such that

$$0 \le R_t \le \hat{q}_{t-1}[h(p_t(1-\alpha)) - p_t h(\alpha)], \qquad t = 1, \ldots, T$$

where

$$\hat{q}_0 = 1-\beta, \quad \hat{q}_t = 1 - [1 - p_t(1-\alpha)]\hat{q}_{t-1}, \qquad t = 1, \ldots, T.$$

We now state our main results for the WUMs with asymmetric input noise as the following theorems.

Theorem 6: For the WUMs with asymmetric input noise, the ε-error capacity region is given by $\boldsymbol{\mathcal{C}}_T^{(a)} = \boldsymbol{\mathcal{R}}_T^{(a)}$.

Theorem 7: For the WUMs with asymmetric input noise, the maximum total number of information bits stored in one storage cell of the memories during the T updating cycles is given by

$$C_a(T) = (1-\beta)\log\lambda_{T+1} + \beta\log\lambda_T$$

where the sequences $\{\lambda_i\}_{i=1}^{\infty}$ satisfy the recursive equation

$$\lambda_i = 2^{-h(\alpha)/(1-\alpha)}\lambda_{i-1} + \lambda_{i-2}, \quad \lambda_1 = 1$$
$$\lambda_2 = 1 + 2^{-h(\alpha)/(1-\alpha)}.$$

As a direct consequence of Theorem 7, we have the following.

Corollary 3: For the WUMs with asymmetric input noise, the average capacity is given by

$$C_a = \log\left[\frac{1 + \sqrt{1 + 4 \cdot 2^{2h(\alpha)/(1-\alpha)}}}{2}\right] - \frac{h(\alpha)}{1-\alpha}.$$

Remark: Let $\beta = \alpha = 0$ in the models of WUMs with symmetric or asymmetric input noise, we get the model of WUMs. This fact also implies that for the WUMs, the ε-error capacity region is identical to the zero-error capacity region.

The Proofs of Theorems 4 and 6 will be given in Section VII. Similar to the proof of Lemma 2 in Section IV, we can prove the following two lemmas by induction. Theorems 5 and 7 follow from these two lemmas directly. The proofs of Lemmas 3 and 4 are omitted here.

Lemma 3: For $0 \le p_1, \ldots, p_T \le 1$, let $\overline{q}_0 = 1-\beta$ and

$$\overline{q}_t = 1 - (p_t * \alpha)\overline{q}_{t-1}, \qquad t = 1, \ldots, T.$$

Then

$$\sum_{t=1}^{T} \overline{q}_{t-1}\left[h(p_t * \alpha) - h(\alpha)\right] \le (1-\beta)\log\theta_{T+1} + \beta\log\theta_T$$

where the sequences $\{\theta_i\}_{i=1}^{\infty}$ satisfy the recursive equation

$$\theta_i = 2^{-h(\alpha)}[\theta_{i-1} + \theta_{i-2}], \quad \theta_1 = 1, \quad \theta_2 = 2^{1-h(\alpha)}.$$

The equality holds only when

$$p_{T-i} = \frac{1}{1-2\alpha}\left[\frac{2^{-h(\alpha)}\theta_i}{\theta_{i+2}} - \alpha\right], \qquad i = 0, 1, \ldots, T-1.$$

Lemma 4: For $0 \le p_1, \ldots, p_T \le 1$, let $\hat{q}_0 = 1-\beta$ and

$$\hat{q}_t = 1 - \left[1 - p_t(1-\alpha)\right]\hat{q}_{t-1}, \qquad t = 1, \ldots, T.$$

Then

$$\sum_{t=1}^{T} \hat{q}_{t-1}\left[h(p_t(1-\alpha)) - p_t h(\alpha)\right] \le (1-\beta)\log\lambda_{T+1} + \beta\log\lambda_T$$

where the sequences $\{\lambda_i\}_{i=1}^{\infty}$ satisfy the recursive equation

$$\lambda_i = 2^{-h(\alpha)/(1-\alpha)}\lambda_{i-1} + \lambda_{i-2}, \qquad \lambda_1 = 1$$
$$\lambda_2 = 1 + 2^{-h(\alpha)/(1-\alpha)}.$$

The equality holds only when

$$p_{T-i} = \frac{2^{-h(\alpha)/(1-\alpha)}\lambda_{i+1}}{(1-\alpha)\lambda_{i+2}}, \qquad i = 0, 1, \ldots, T-1.$$

VII. Proofs of Theorems 4 and 6

Heegard [16] showed that the inner bound on the capacity region of reusable memories in Theorem 3 is tight for WOMs with symmetric input noise. By modifying his methods, we subsequently show that this inner bound is also tight for the WUMs with symmetric or asymmetric input noise. In the sequel, we present the proof of Theorem 4 in details. Since Theorem 6 can be proved in the same way, we only present a sketch for the proof of Theorem 6.

A. Proofs of the Direct Parts of Theorems 4 and 6

The direct parts of Theorems 4 and 6 follow from Theorem 3 by setting $U_t = X_t, t = 1, 2, \ldots, T$.

Proof of the Direct Part of Theorem 4: The conditional probability distribution $\boldsymbol{W}$ for this specific class of reusable memories, WUMs with symmetric input noise, is given by (11) and (12). In Theorem 3, we set $U_t = X_t$ and

$$P_t(0|0) = \Pr\{X_t = 0|Y_{t-1} = 0\} = 1 - p_t, \tag{15}$$

$$P_t(1|0) = \Pr\{X_t = 1|Y_{t-1} = 0\} = p_t \tag{16}$$

$$P_t(0|1) = \Pr\{X_t = 0|Y_{t-1} = 1\} = \frac{(1-p_t)(1-\alpha)}{(1-p_t)(1-\alpha) + p_t\alpha} \tag{17}$$

$$P_t(1|1) = \Pr\{X_t = 1|Y_{t-1} = 1\} = \frac{p_t\alpha}{(1-p_t)(1-\alpha) + p_t\alpha}. \tag{18}$$

Denote

$$\overline{q}_t = \Pr\{Y_t = 0\}, \qquad t = 0, 1, \ldots, T. \tag{19}$$

By direct computation, we have $\overline{q}_0 = 1 - \beta$ and for $t = 1, \ldots, T$

$$\overline{q}_t = 1 - (p_t * \alpha)\overline{q}_{t-1} \tag{20}$$

$$I(X_t; Y_t) - I(X_t; Y_{t-1}) = \overline{q}_{t-1}[h(p_t * \alpha) - h(\alpha)]. \tag{21}$$

The direct part of Theorem 4 follows from Theorem 3, (20) and (21). □

Proof of the Direct Part of Theorem 6: The conditional probability distribution $\boldsymbol{W}$ for this specific class of reusable memories, WUMs with asymmetric input noise, is given by (13) and (14). In Theorem 3, we set $U_t = X_t$ and

$$P_t(0|0) = \Pr\{X_t = 0|Y_{t-1} = 0\} = p_t \tag{22}$$

$$P_t(1|0) = \Pr\{X_t = 1|Y_{t-1} = 0\} = 1 - p_t \tag{23}$$

$$P_t(0|1) = \Pr\{X_t = 0|Y_{t-1} = 1\} = 1 \tag{24}$$

$$P_t(1|1) = \Pr\{X_t = 1|Y_{t-1} = 1\} = 0. \tag{25}$$

Denote

$$\hat{q}_t = \Pr\{Y_t = 0\}, \qquad t = 0, 1, \ldots, T. \tag{26}$$

By direct computation, we have $\hat{q}_0 = 1 - \beta$ and for $t = 1, \ldots, T$

$$\hat{q}_t = 1 - [1 - p_t(1-\alpha)]\hat{q}_{t-1} \tag{27}$$

$$I(X_t; Y_t) - I(X_t; Y_{t-1}) = \hat{q}_{t-1}[h(p_t(1-\alpha)) - p_t h(\alpha)]. \tag{28}$$

The direct part of Theorem 6 follows from Theorem 3, (27) and (28). □

We note that in the preceding proofs, both (15)–(18) and (22)–(25) actually imply that for every $t = 1, 2, \ldots, T$, the random variables Y_{t-1}, Y_t, X_t form a Markov chain.

B. Proofs of the Converse Parts of Theorems 4 and 6

We only present a proof for the converse part of Theorem 4 here. Since we can prove the converse part of Theorem 6 in the same way, we omit the proof of it here.

Consider the WUMs with symmetric input noise described by (11) and (12). Assume that the vector $(R_1, R_2, \ldots, R_T)$ is ε-achievable. From the definition in Section VI, we know that for any $\varepsilon > 0$ there exists an $(n, M_1, \ldots, M_T, \delta)$ code for some n, such that $R_t = (1/n)\log M_t$ for $t = 1, 2, \ldots, T$ and the error probability $\delta < \varepsilon$. Let W_t, $X^n(t)$, $Y^n(t)$, $\bar{W}_t$, $t = 1, 2, \ldots, T$ be the random variables defined in Section VI. It is not hard to see from the definition that if $Y^n(t-1)$ is given, then $W_t, X^n(t), Y^n(t), \bar{W}_t$ form a Markov chain. From the Data Processing Theorem, we have

$$I\left(W_t; \bar{W}_t | Y^n(t-1)\right) \le I\left(X^n(t); Y^n(t) | Y^n(t-1)\right). \tag{29}$$

Note that W_t is independent of $Y^n(t-1)$

$$H\left(W_t | Y^n(t-1)\right) = H(W_t) = \log M_t. \tag{30}$$

From the Fano inequality and the fact that $\Pr\{W_t \ne \bar{W}_t\} < \varepsilon$, we have

$$H\left(W_t | \bar{W}_t, Y^n(t-1)\right) \le H\left(W_t|\bar{W}_t\right) \le \varepsilon \log M_t + h(\varepsilon). \tag{31}$$

By (30) and (31), we have

$$\begin{aligned} I\left(W_t; \bar{W}_t | Y^n(t-1)\right) &= H\left(W_t | Y^n(t-1)\right) - H\left(W_t | \bar{W}_t, Y^n(t-1)\right) \\ &\ge (1-\varepsilon)\log M_t - h(\varepsilon). \end{aligned} \tag{32}$$

On the other hand, we know that

$$H(Y^n(t)|Y^n(t-1)) - H(Y^n(t)|X^n(t), Y^n(t-1)) = I(X^n(t); Y^n(t)|Y^n(t-1)) \tag{33}$$

and

$$H(Y^n(t)|Y^n(t-1)) \le \sum_{i=1}^{n} H(Y_i(t)|Y_i(t-1)). \tag{34}$$

Since $Y^n(t)$ is generated by the conditional distribution $\boldsymbol{W}^n$ through $X^n(t)$ and $Y^n(t-1)$, we have

$$H(Y^n(t)|X^n(t), Y^n(t-1)) = \sum_{i=1}^{n} H(Y_i(t)|X_i(t), Y_i(t-1)). \tag{35}$$

From (33)–(35), we obtain that

$$I(X^n(t); Y^n(t)|Y^n(t-1)) \le \sum_{i=1}^{n}\Big[H(Y_i(t)|Y_i(t-1)) - H(Y_i(t)|X_i(t), Y_i(t-1))\Big]. \tag{36}$$

For every $t = 1, 2, \ldots, T$, set

$$\Pr\{X_i(t) = 0|Y_i(t-1) = 0\} = 1 - p_i(t) \tag{37}$$

$$\Pr\{X_i(t) = 1|Y_i(t-1) = 0\} = p_i(t) \tag{38}$$

$$\overline{q}_i(t) = \Pr\{Y_i(t) = 0\}. \tag{39}$$

By direct computation, we obtain that

$$\overline{q}_i(t) = 1 - (p_i(t) * \alpha)\overline{q}_i(t-1) \tag{40}$$

$$H(Y_i(t)|Y_i(t-1)) = \overline{q}_i(t-1)h(p_i(t) * \alpha) \tag{41}$$

$$H(Y_i(t)|X_i(t), Y_i(t-1)) = \overline{q}_i(t-1)h(\alpha). \tag{42}$$

It follows from (36), (41), and (42) that

$$I(X^n(t); Y^n(t)|Y^n(t-1)) \le \sum_{i=1}^{n} \overline{q}_i(t-1)\Big[h(p_i(t) * \alpha) - h(\alpha)\Big]. \tag{43}$$

From (29), (32), and (43), we obtain that

$$(1-\varepsilon)R_t - \frac{1}{n}h(\varepsilon) \le \frac{1}{n}\sum_{i=1}^{n} \overline{q}_i(t-1)[h(p_i(t) * \alpha) - h(\alpha)]. \tag{44}$$

Denote

$$\overline{q}_t = \frac{1}{n}\sum_{i=1}^{n} \overline{q}_i(t), \qquad t = 0, 1, \ldots, T. \tag{45}$$

For $t = 1, 2, \ldots, T$, denote

$$p_t = \frac{\sum_{i=1}^{n} p_i(t)\overline{q}_i(t-1)}{\sum_{i=1}^{n} \overline{q}_i(t-1)} = \frac{\sum_{i=1}^{n} p_i(t)\overline{q}_i(t-1)}{n\overline{q}_{t-1}}. \quad (46)$$

Since for all $i = 1, 2, \ldots, n$, $\overline{q}_i(0) = 1 - \beta$, we know from (45) that $\overline{q}_0 = 1 - \beta$. It follows from (40), (45), and (46) that for $t = 1, 2, \ldots, T$

$$\begin{aligned}\overline{q}_t &= \frac{1}{n}\sum_{i=1}^{n} \overline{q}_i(t) \\ &= \frac{1}{n}\sum_{i=1}^{n} [1 - (p_i(t) * \alpha)\overline{q}_i(t-1)] \\ &= 1 - (p_t * \alpha)\overline{q}_{t-1}. \end{aligned} \quad (47)$$

Since the binary entropy function $h(x)$ is convex, we have

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n} \overline{q}_i(t-1)[h(p_i(t) * \alpha) - h(\alpha)] \\ &= \overline{q}_{t-1}\left[\left[\sum_{i=1}^{n} (\overline{q}_i(t-1)/n\overline{q}_{t-1}) h(p_i(t) * \alpha)\right] - h(\alpha)\right] \\ &\le \overline{q}_{t-1}\left[h\left(\sum_{i=1}^{n} \overline{q}_i(t-1)(p_i(t) * \alpha)/n\overline{q}_{t-1}\right) - h(\alpha)\right] \\ &= \overline{q}_{t-1}[h(p_t * \alpha) - h(\alpha)]. \end{aligned} \quad (48)$$

Combining (44) with (48), we know that for every $t = 1, 2, \ldots, T$

$$(1 - \varepsilon)R_t - \frac{1}{n}h(\varepsilon) \le \overline{q}_{t-1}[h(p_t * \alpha) - h(\alpha)]. \quad (49)$$

From (47), (49), and letting $\varepsilon \to 0$, we obtain that

$$(R_1, R_2, \ldots, R_T) \in \mathcal{R}_T^{(s)}.$$

This completes the proof of the converse part of Theorem 4.

VIII. Conclusion

In this correspondence, from an information-theoretic point of view, we study the capacity of WUMs with nonperiodic codes. For the situation where the encoder knows and the decoder does not know the previous content of the memory, we determine the zero-error capacity region, the average capacity, and the maximum total number of information bits stored in the WUMs for fixed T successive cycles. Furthermore, two models of WUMs with symmetric or asymmetric input noise are introduced and studied. We determine the ε-error capacity region, the average capacity, and the maximum total number of information bits stored in the WUMs with symmetric or asymmetric input noise for fixed T successive cycles.

Acknowledgment

The authors would like to thank the anonymous reviewers and Associate Editor Prakash Narayan for their valuable suggestions and comments that helped to improve the correspondence.

References

[1] F. M. J. Willems and A. J. H. Vinck, "Repeated recording for an optical disk," in *Proc. 7th Symp. Information Theory in the Benelux*, Delft, The Netherlands, May 22–23, 1986, pp. 49–53.

[2] J. M. Borden, "Coding for write-unidirectional memories," preprint, 1986.

[3] G. Simonyi, "On write-unidirectional memory codes," *IEEE Trans. Inform. Theory*, vol. 35, pp. 663–669, May 1989.

[4] W. M. C. J. van Overveld, "The four cases of write-unidirectional memory codes over arbitrary alphabets," *IEEE Trans. Inform. Theory*, vol. 37, pp. 872–878, May 1991.

[5] G. Cohen, "On the capacity of write-unidirectional memories," *Bull. Inst. Math., Academia Sinica*, vol. 16, pp. 285–293, 1988.

[6] G. Cohen and G. Simonyi, "Coding for write-unidirectional memories and conflict resolution," *Discr. Appl. Math.*, vol. 24, pp. 103–114, 1989.

[7] G. Cohen, " $\mathrm{W}^*\mathrm{M}$'s: A survey writing on some binary memories with constraints," in *Geometries, Codes and Cryptography*, G. Longo, Ed. New York: Springer-Verlag, 1990, pp. 213–227.

[8] A. V. Kuznetsov and A. J. H. Vinck, "On the general defective channel with informed encoder and capacities of some constrained memories," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1866–1871, Nov. 1994.

[9] A. V. Kuznetsov, "A lower bound on the capacity of write-unidirectional memory in the presence of random error," in *IEEE Int. Symp. Information Theory, Abstract of Papers*, Kobe, Japan, June 19–24, 1988, p. 4.

[10] F. M. J. Willems, "Converses for Write-Unidirectional Memories," EUT Rep. 89-E-220, Eindhoven, The Netherlands, 1989.

[11] K. U. Koschnick, "Good code design with combinatorial approximation algorithms," Ph.D. dissertation, Univ. Bielefeld, Bielefeld, Germany, 1990.

[12] W. M. C. J. van Overveld, "On the capacity region for deterministic two-way channels and write-unidirectional memories," Ph.D. dissertation, Eindhoven Univ. Technol., Eindhoven, The Netherlands, 1991.

[13] R. Ahlswede and Z. Zhang, "On multiuser write-efficient memories," *IEEE Trans. Inform. Theory*, vol. 40, pp. 674–686, May 1994.

[14] R. Ahlswede and G. Simonyi, "Reusable memories in the light of the old varying and a new outputwise varying channel theory," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1143–1150, July 1991.

[15] R. Ahlswede and Z. Zhang, "Coding for write-efficient memory," *Inform. Comput.*, vol. 83, pp. 80–97, 1994.

[16] C. Heegard, "On the capacity of permanent memory," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 34–42, Jan. 1985.

[17] C. Heegard and A. El Gamal, "On the capacity of a computer memory with defects," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 731–739, Sept. 1983.

[18] A. V. Kuznetsov and B. S. Tsybakov, "Coding in a memory with defective cell," *Probl. Inform. Transm.*, vol. 10, pp. 132–138, 1974.

[19] S. I. Gel'fand and M. S. Pinsker, "Coding for a channel with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, pp. 19–31, 1980.

[20] R. L. Rivest and A. Shamir, "How to reuse a write-once memory," *Inform. Contr.*, vol. 55, pp. 1–19, 1982.

[21] A. Fiat and A. Shamir, "Generalized write-once memories," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 470–480, May 1984.

[22] J. K. Wolf, A. D. Wyner, J. Ziv, and J. Körner, "Coding for write-once memory," *AT&T Bell Labs. Tech. J*, vol. 63, pp. 1089–1112, 1984.

[23] G. Cohen, P. Godlewski, and F. Merkx, "Linear block codes for write-once memories," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 697–700, Sept. 1986.

[24] G. Zémor and G. Cohen, "Error-correcting WOM-codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 730–734, May 1991.

[25] F.-W. Fu and A. J. H. Vinck, "On the capacity of generalized write-once memory with state transitions described by an arbitrary directed acyclic graph," *IEEE Trans. Inform. Theory*, vol. 45, pp. 308–314, Jan. 1999.

[26] F.-W. Fu and R. W. Yeung, "On the capacity and error-correcting codes of write-efficient memories," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2299–2314, Nov. 2000.

[27] M. H. Bennett, "3M rewritable media," in *New Advances in Optical Recording Technology, IEE Coll. Dig.*, vol. 1989/45, Mar. 1989.