ELSEVIER

Theoretical Computer Science 270 (2002) 739–760

Theoretical Computer Science

www.elsevier.com/locate/tcs

Probabilistic rebound Turing machines

Lan Zhang, Katsushi Inoue *, Akira Ito, Yue Wang

Department of Computer Science and System Engineering, Faculty of Engineering, Yamaguchi University, Ube, Yamaguchi 755-8611 Japan

Received March 2000; revised November 2000; accepted January 2001
Communicated by M. Ito

Abstract

This paper introduces a probabilistic rebound Turing machine (PRTM), and investigates the fundamental property of the machine. We first prove a sublogarithmic lower space bound on the space complexity of this model with bounded errors for recognizing specific languages. This lower bound strengthens a previous lower bound for conventional probabilistic Turing machines with bounded errors. We then show, by using our lower space bound and an idea in the proof of it, that

(i) $\mathcal{L}[PRTM(\mathrm{o}(\log n))]$ is incomparable with the class of context-free languages,

(ii) there is a language accepted by a two-way deterministic one counter automaton, but not in $\mathcal{L}[\mathrm{PRTM}(\mathrm{o}(\log n))]$, and

(iii) there is a language accepted by a deterministic one-marker rebound automaton, but not in $\mathcal{L}[\mathrm{PRTM}(\mathrm{o}(\log n))]$,

where $\mathcal{L}[\mathrm{PRTM}(\mathrm{o}(\log n))]$ denotes the class of languages recognized by $\mathrm{o}(\log n)$ space-bounded PRTMs with error probability less than $\frac{1}{2}$. Furthermore, we show that there is an infinite space hierarchy for $\mathcal{L}[\mathrm{PRTM}(\mathrm{o}(\log n))]$. We finally show that $\mathcal{L}[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is not closed under concatenation, Kleene+, and length-preserving homomorphism. This paper answers two open problems in a previous paper. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Probabilistic rebound Turing machine; Rebound automaton; Space hierarchy; Closure property

1. Introduction

The rebound automaton (RA) introduced by Sugata et al. [17] has the same structure as a two-dimensional finite automaton [1, 8, 14], but an input to it is a square tape whose top row is a word to be recognized, and whose other symbols are all blank. It was demonstrated in [17] that many nonregular languages (e.g., the languages

* Corresponding author.
E-mail address: inoue@csse.yamaguchi-u.ac.jp (K. Inoue).

PII: S0304-3975(01)00098-6

$\{a^nb^nc^n \mid n \geqslant 1\}$ and $\{ww \mid w \in \{0,1\}^*\}$) are accepted by RAs. In papers [9, 11, 12, 15], investigations of RAs have been continued. Furthermore, in [20], alternating rebound Turing machines were introduced, and their fundamental properties were investigated. It should be noted that Petersen [11] (resp., [12]) gave a language accepted by a two-way (resp., one-way) deterministic one-counter automaton, but not accepted by any nondeterministic RA, and Petersen [12] gave a separation of the classes of languages accepted by deterministic and nondeterministic RAs, solving the long standing open problems.

Recently, in [19], we introduced a probabilistic rebound automaton (PRA), and showed that

(1) the class of languages recognized by PRAs with error probability less than $\frac{1}{2}$, $\mathcal{L}$[PRA], is incomparable with the class of context-free languages,
(2) there is a language accepted by a two-way nondeterministic one counter automaton, but not in $\mathcal{L}$[PRA],
(3) there is a language accepted by a deterministic one-marker RA, but not in $\mathcal{L}$[PRA], and
(4) $\mathcal{L}$[PRA] is not closed under concatenation and Kleene $+$.

It is quite natural to introduce a *probabilistic rebound Turing machine* (PRTM) which is a PRA equipped with one semi-infinite read-write storage tape. In this paper, we investigate recognizing powers, space hierarchy, and closure property of $o(\log n)$ space-bounded PRTMs with error probability less than $\frac{1}{2}$.

Section 2 of the paper presents some definitions and notations necessary for this paper.

Dwork and Stockmeyer [2] proved an impossibility result for probabilistic finite automata with bounded errors. By using an idea similar to the proof of this result, Freivalds and Karpinski [4] proved, for the first time, a sublogarithmic lower space bound for probabilistic Turing machines with bounded errors. In Section 3, by extending their proof techniques to our model, we first prove a sublogarithmic lower space bound on the space complexity of our model with bounded errors for recognizing specific languages. We believe that our sublogarithmic space lower bound is the strongest for probabilistic machines (with bounded errors) so far, because our $o(\log n)$ space-bounded models are more powerful than conventional $o(\log n)$ space-bounded probabilistic Turing machines, which is shown below.

Freivalds and Karpinski [4] showed, by using their lower space bound theorem, that there is a context-free language not accepted by any $o(\log n)$ space-bounded probabilistic Turing machine with bounded error. *Pal*, i.e., the set of all the palindromes, is such a language. Dwork and Stockmeyer [2] also showed that there is another context-free language '$\mathit{Center} = \{u1v \mid u,\ v \in \{0,1\}^*$ and the lengths of u and v are the same$\}$' not accepted by any $o(\log n)$ space-bounded probabilistic Turing machine with bounded error. It is easy to see that the languages *Pal* and *Center* are recognized by deterministic RAs, and thus by PRAs with bounded errors. Thus, $L(n)$ space-bounded PRTMs with bounded errors are more powerful than $L(n)$ space-bounded probabilistic Turing machines with bounded errors for any $L(n) = o(\log n)$.

In Section 4, by using our lower space bound theorem and an idea in the proof of it, we investigate relationships between $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$, which is the class of languages recognized by $\mathrm{o}(\log n)$ space-bounded PRTMs with error probability less than $\frac{1}{2}$, and other classes of languages. We first show that $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is incomparable with the class of context-free languages. We next show that there is a language accepted by a two-way deterministic one counter automaton [6], but not in $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$. This result solves an open problem in [19]. Further, we show that there is a language accepted by a deterministic one-marker RA, but not in $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

Section 5 investigates a space hierarchy for $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$, and shows that if $L(n)$ is space constructible by a deterministic rebound Turing machine (DRTM) [20], $\log\log n \leqslant L(n) = \mathrm{o}(\log n)$, and $L'(n) = \mathrm{o}(L(n))$, then there is a language accepted by $L(n)$ space-bounded DRTM, but not recognized by any $L'(n)$ space-bounded PRTM with error probability less than $\frac{1}{2}$.

Section 6 investigates closure properties of $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$. We first show that $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is not closed under concatenation and Kleene $+$. We next show that $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is not closed under length-preserving homomorphism. This result solves an open problem in [19]. We again note that Petersen [11] showed that the class of languages accepted by nondeterministic RAs is not closed under length-preserving homomorphism.

Section 7 presents several open problems.

2. Preliminaries

Let Σ be a finite set of symbols. A *two-dimensional tape* over Σ is a two-dimensional rectangular array of elements of Σ. The set of all two-dimensional tapes over Σ is denoted by $\Sigma^{(2)}$. Given a tape $x \in \Sigma^{(2)}$, we let $l_1(x)$ be the number of rows of x, and $l_2(x)$ be the number of columns of x. If $1 \leqslant i \leqslant l_1(x)$ and $1 \leqslant j \leqslant l_2(x)$, we let $x(i,j)$ denote the symbol in x with coordinates (i,j). Furthermore, we define

$$x[(i,j),(i',j')],$$

only when $1 \leqslant i \leqslant i' \leqslant l_1(x)$ and $1 \leqslant j \leqslant j' \leqslant l_2(x)$, as the two-dimensional tape z satisfying the following:

(i) $l_1(z) = i' - i + 1$ and $l_2(z) = j' - j + 1$;
(ii) for each k, $r[1 \leqslant k \leqslant l_1(z), 1 \leqslant r \leqslant l_2(z)]$, $z(k,r) = x(k+i-1, r+j-1)$.

For each $m,n \geqslant 1$, let $\Sigma^{m\times n} = \{x \in \Sigma^{(2)} \mid l_1(x) = m \ \&\ l_2(x) = n\}$. For each $n \geqslant 1$, let Σ^n be the set of words in Σ^+ of length n. For any word w, $|w|$ denotes the length of w, and for any set A, $|A|$ denotes the number of elements of A.

We now introduce a probabilistic rebound Turing machine which is a probabilistic rebound automaton [19] equipped with one semi-infinite read–write storage tape. Let S be a finite set. A *coin-tossing distribution on S* is a mapping Ψ from S to $\{0, \frac{1}{2}, 1\}$ such that $\Sigma_{a\in S}\Psi(a) = 1$. The mapping means "choose a with probability $\Psi(a)$".

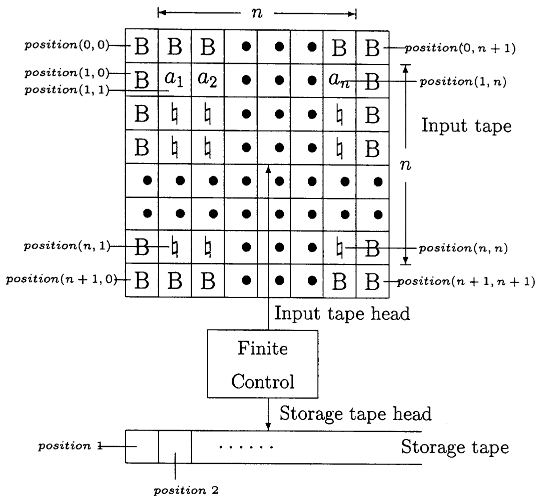

Fig. 1. Probabilistic rebound Turing machine.

A *probabilistic rebound Turing machine* (PRTM) is a 10-tuple

$$M = (Q, \Sigma, \Gamma, \#, \natural, \mathrm{B}, \delta, q_0, q_\mathrm{a}, q_\mathrm{r}),$$

where

(1) Q is a finite set of *states*,
(2) Σ is a finite *input alphabet*,
(3) Γ is a finite *storage tape alphabet*,
(4) $\# \in \Gamma$ is the *blank symbol on the storage tape*,
(5) $\natural \notin \Sigma$ is the *blank symbol on the input tape*,
(6) $\mathrm{B} \notin \Sigma$ is the *boundary symbol*,
(7) δ is a *transition function*,
(8) $q_0 \in Q$ is the *initial state*,
(9) $q_\mathrm{a} \in Q$ is the *accepting state*, and
(10) q_r is the *rejecting state*.

For each word $w \in \Sigma^+$, let $w(\natural)$ denote the two-dimensional square tape over $\Sigma \cup \{\natural\}$, whose top row is w and whose other symbols are all $\natural$'s. As shown in Fig. 1, an input tape for M is a two-dimensional square tape $w(\natural)$ for some $w = a_1 a_2 \ldots a_n \in \Sigma^+$, $n \geqslant 1$, surrounded by the boundary symbols B. Of course, M has one semi-infinite storage tape (initially blank), a finite control, an input tape head, and a storage tape head. For convenience sake, a position is also assigned to each cell of the read-only input tape and the storage tape as shown in Fig. 1.

The transition function δ is defined on $(Q - \{q_\mathrm{a}, q_\mathrm{r}\}) \times (\Sigma \cup \{\natural, \mathrm{B}\}) \times \Gamma$ such that for each $q \in Q - \{q_\mathrm{a}, q_\mathrm{r}\}$, each $\sigma \in \Sigma \cup \{\natural, \mathrm{B}\}$, and each $\gamma \in \Gamma$, $\delta[q, \sigma, \gamma]$ is a coin-tossing distribution on $Q \times (\Gamma - \{\#\}) \times \{\mathrm{Left}, \mathrm{Right}, \mathrm{Up}, \mathrm{Down}, \mathrm{Stay}\} \times \{\mathrm{Left}, \mathrm{Right}, \mathrm{Stay}\}$, where Left means "moving left", Right "moving right", Up "moving up", Down "moving down", and Stay "staying there". The meaning of δ is that if M is in state q with

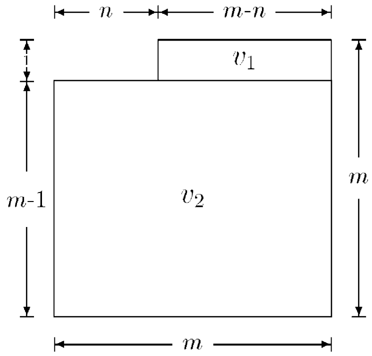

Fig. 2. (m,n)-chunk.

the input head scanning the symbol σ and the storage tape head scanning the symbol γ, then with probability $\delta[q,\sigma,\gamma](q',\gamma',d_1,d_2)$ the machine enters state q', rewrites the symbol γ by the symbol γ', either moves the input head one cell in direction d_1 if $d_1 \in \{\text{Left}, \text{Right}, \text{Up}, \text{Down}\}$ or does not move the input head if $d_1 = \text{Stay}$, and either moves the storage head one cell in direction d_2 if $d_2 \in \{\text{Left}, \text{Right}\}$ or does not move the storage head if $d_2 = \text{Stay}$.

Suppose that an input tape $w(\natural)$ with $w \in \Sigma^n$ $(n \geqslant 1)$ is presented to M. M starts in the initial state q_0 with the input head on the upper left-hand corner of $w(\natural)$, with all the cells of the storage tape blank and with the storage tape head on the left end of the storage tape. The computation of M on $w(\natural)$ is then governed (probabilistically) by the transition function δ until M either accepts by entering the accepting state q_a or rejects by entering the rejecting state q_r. We assume that δ is defined so that the input head never falls off an input tape out of the boundary symbols B, the storage tape head cannot write the blank symbol, and fall off the storage tape by moving left. M halts when it enters state q_a or q_r.

Let $T \subseteq \Sigma^+$ and $0 \leqslant \varepsilon < \frac{1}{2}$. A PRTM M recognizes T with error probability ε if for all $w \in T$, M accepts $w(\natural)$ with probability at least $1-\varepsilon$, and for all $w \notin T$, M rejects $w(\natural)$ with probability at least $1-\varepsilon$.

Let $L: N \to N \cup \{0\}$ be a function. We say that a PRTM M is $L(n)$ *space-bounded* if for each $n \geqslant 1$ and for each input tape $w(\natural)$ with $|w| = n$, M on $w(\natural)$ uses at most $L(n)$ cells of the storage tape. By $\pounds[\text{PRTM}(L(n))]$, we denote the class of sets of words recognized by $L(n)$ space-bounded PRTMs with error probability less than $\frac{1}{2}$.

The reader is referred to [7] for undefined terms.

3. A lower space bound for PRTMs

In this section, we prove a lower space bound for PRTMs, which is used in the subsequent sections. We first give some preliminaries necessary for getting the lower space bound.

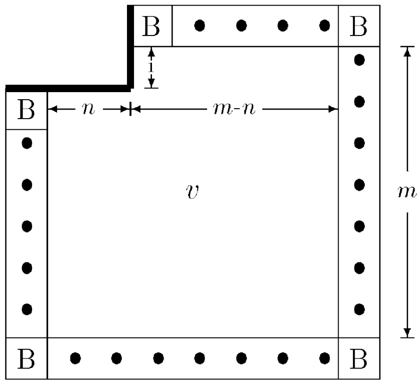

Fig. 3. v(B).

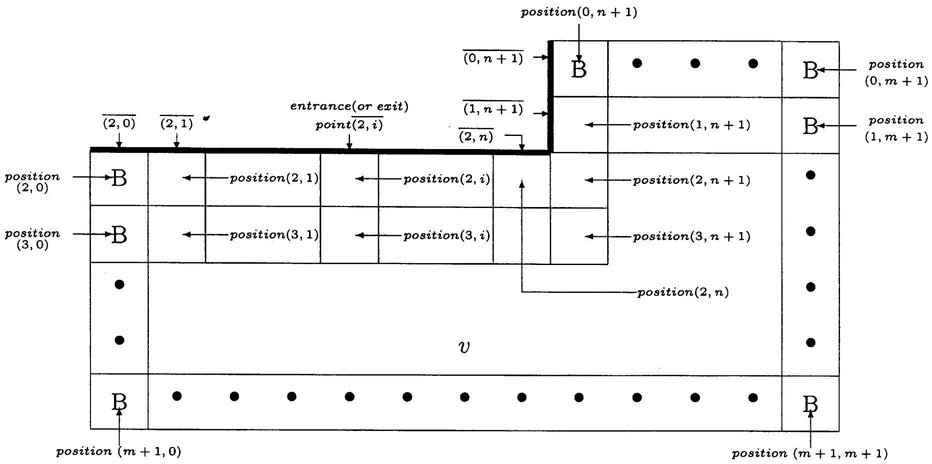

Fig. 4. Illustration for v(B) ($v:(m,n)$-chunk.

Let Σ be an alphabet. For each $m \geqslant 2$ and each $1 \leqslant n \leqslant m-1$, an (m,n)-*chunk* over Σ is a pattern over $\Sigma \cup \{\natural\}$ as shown in Fig. 2, where $v_1 \in \Sigma^{1\times(m-n)}$ and $v_2 \in \{\natural\}^{(m-1)\times m}$. By $ch_{(m,n)}(v_1)$, we denote the (m,n)-chunk as shown in Fig. 2.

Let M be a PRTM whose input alphabet is Σ, and $\natural$ and B be the blank symbol and boundary symbol of M, respectively. For any (m,n)-chunk v, we denote by v(B) the pattern obtained from v by attaching the boundary symbols B to v as shown in Fig. 3. Below, we assume without loss of generality that M enters or exits the pattern v(B) only at the face designated by the bold line in Fig. 3. Thus, the number of the entrance points to v(B) (or the exit points from v(B)) for M is $n+3$. We suppose that these entrance points (or exit points) are named $\overline{(2,0)}$, $\overline{(2,1)},\ldots,\overline{(2,n)}$, $\overline{(1,n+1)}$, $\overline{(0,n+1)}$ as shown in Fig. 4. Let $PT(v(\mathrm{B}))$ be the set of these entrance points (or exit points). To each cell of v(B), we assign a position as shown in Fig. 4.

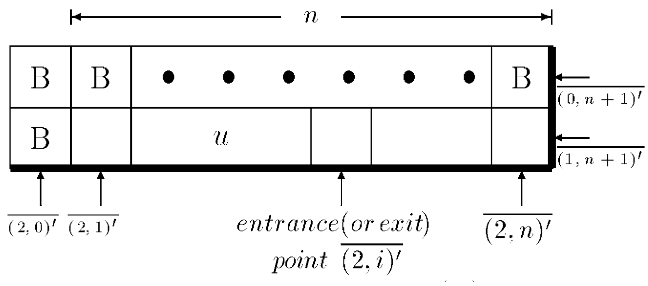

Fig. 5. Illustration for u(B).

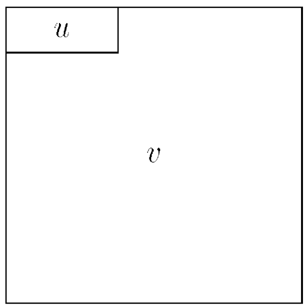

Fig. 6. $[u]v$.

Let $PS(v(\mathrm{B}))$ be the set of all the positions of $v(\mathrm{B})$. For each $n \geqslant 1$, an *n-chunk* over Σ is a pattern in $\Sigma^{1\times n}$. For any n-chunk u, we denote by $u(\mathrm{B})$ the pattern obtained from u by attaching the boundary symbols B to u as shown in Fig. 5. We again assume without loss of generality that M enters or exits the pattern $u(\mathrm{B})$ only at the face designated by the bold line in Fig. 5. The number of the entrance points to $u(\mathrm{B})$ (or the exit points from $u(\mathrm{B})$) for M is again $n+3$, and these entrance points (or exit points) are named $\overline{(2,0)'}, \overline{(2,1)'}, \ldots, \overline{(2,n)'}, \overline{(1,n+1)'}, \overline{(0,n+1)'}$ as shown in Fig. 5. Let $PT(u(\mathrm{B}))$ be the set of these entrance points (or exit points). (Note that the entrance points of an n-chunk are distinguished from the entrance points of an (m,n)-chunk only by "dash".)

For any (m,n)-chunk v over Σ and any n-chunk u over Σ, let $[u]v$ be the tape in $(\Sigma \cup \{\natural\})^{m\times m}$ consisting of v and u as shown in Fig. 6.

The result in this section is based on an idea firstly used by Rabin [13], and then adapted in different contexts by Greenberg and Weiss [5], Dwork and Stockmeyer [2] and Freivalds and Karpinski [4].

Let M be a PRTM. A *storage state* of M is a combination of the state of the finite control, the nonblank contents of the storage tape, and the storage tape head position. Let q_{a} and q_{r} be the accepting and the rejecting states of M, respectively, and x be an (m,n)-chunk (or an n-chunk) over Σ $(m,n \geqslant 1)$. We define the *chunk probabilities* of M on x as follows. A *starting condition* for the chunk probability is a pair (s,l), where s is a storage state of M and $l \in PT(x(\mathrm{B}))$; its intuitive meaning is "M has just entered $x(\mathrm{B})$ in storage state s from entrance point l of $x(\mathrm{B})$". A starting condition for the chunk probability of M on $x(\mathrm{B})$ is either:

(i) "Initial" meaning that M has just started in its initial storage state with the input head on the upper left-hand corner of x, where the initial storage state of M is the storage state of M such that the state of the finite control is the initial state of M, all the cells of the storage tape are blank, and the storage tape head is on the left end of the storage tape, or

(ii) a pair (s, l), where s is a storage state of M and $l \in PT(x(\mathrm{B}))$; its intuitive meaning is "M has just entered $x(\mathrm{B})$ in storage state s from entrance point l of $x(\mathrm{B})$".

A *stopping condition* for the chunk probability is either:

(i) a pair (s, l) as above, meaning that M exits from $x(\mathrm{B})$ in storage state s at exit point l,

(ii) "Loop" meaning that the computation of M loops forever within $x(\mathrm{B})$,

(iii) "Accept" meaning that M halts in the accepting state q_{a} before exiting from $x(\mathrm{B})$ at an exit point of $x(\mathrm{B})$, or

(iv) "Reject" meaning that M halts in the rejecting state q_{r} before exiting from $x(\mathrm{B})$ at an exit point of $x(\mathrm{B})$.

For each starting condition σ and each stopping condition τ, let $p(x, \sigma, \tau)$ be the probability that stopping condition τ occurs given that M is started in starting condition σ on an (m, n)-chunk (or n-chunk) x. If W is a large set of words, then a pigeonhole argument shows that there must be two chunks $ch_{(m,n)}(w)$ and $ch_{(m,n)}(w')$, where $w, w' \in W$ and $|w| = |w'| = m - n$ for some $m > n \geqslant 1$, which M cannot distinguish, that is, the probabilities $p(ch_{(m,n)}(w), \sigma, \tau)$ and $p(ch_{(m,n)}(w'), \sigma, \tau)$ are very close. If L is a language such that for any two different words w and w' in W with the same length, there is a word u such that $wu \in L$ iff $w'u \notin L$, then M does not recognize L.

Computations of a PRTM are modeled by Markov chains [16] with finite state space, say $\{1, 2, \ldots, s\}$ for some s. A particular Markov chain is completely specified by its matrix $R = \{r_{ij}\}_{1 \leqslant i,j \leqslant s}$ of transition probabilities. If the Markov chain is in state i, then it next moves to state j with probability r_{ij}. The chains we consider have the designated starting state, say, state 1, and some set T_R of trapping states, so $r_{tt} = 1$ for all $t \in T_R$. For $t \in T_R$, let $p^*[t, R]$ denote the probability that Markov chain R is trapped in state t when started in state 1. The following lemma which bounds the effect of small changes in the transition probabilities of a Markov chain is used below.

Let $\beta \geqslant 1$. Say that two numbers r and r' are *β-close* if either (i) $r = r' = 0$ or (ii) $r > 0$, $r' > 0$, and $\beta^{-1} \leqslant r/r' \leqslant \beta$. Two Markov chains $R = \{r_{ij}\}_{i,j=1}^{s}$ and $R' = \{r'_{ij}\}_{i,j=1}^{s}$ are *β-close* if r_{ij} and r'_{ij} are β-close for all pairs i, j.

Lemma 3.1 (Dwork and Stockmeyer [2]). *Let R and R' be two s-state Markov chains which are β-close, and let t be a trapping state of both R and R'. Then $p^*[t, R]$ and $p^*[t, R']$ are β^{2s}-close.*

We are now ready to prove our lower space bound.

Theorem 3.1. *Let $A, B \subseteq \Sigma^*$ with $A \cap B = \phi$. Suppose that there is an infinite set I of positive integers and a function $G(n)$ such that $G(n)$ is a fixed function bounded by some exponential in n, and for each $n \in I$ there is a set $W(n)$ of words in Σ^* such that*:

(1) $|w| = G(n)$ *for all* $w \in W(n)$,
(2) *there are constants* $c > 1$ *and* $r > 0$ *such that* $|W(n)| \geqslant 2^{c^{n^r}}$ *for all* $n \in I$,
(3) *for every* $n \in I$ *and every* $w, w' \in W(n)$ *with* $w \neq w'$, *there is a word* $u \in \Sigma^n$ *such that*:

$$\text{either} \begin{cases} uw \in A \\ uw' \in B \end{cases} \quad \text{or} \quad \begin{cases} uw \in B, \\ uw' \in A. \end{cases}$$

Then, if an $L(n)$ space-bounded PRTM with error probability $\varepsilon < \frac{1}{2}$ separates A and B, then $L(H(n))$ cannot be $\mathrm{o}(n^r)$, *where* $H(n) = n + G(n)$.

Proof. Suppose that there is a PRTM($L(n)$) M separating A and B with error probability $\varepsilon < \frac{1}{2}$. By $C(n)$ we denote the set of possible storage states of M on input tapes of side-length n ($n \geqslant 1$), and let $c(n) = |C(n)|$. It is obvious that $c(n) \leqslant \mathrm{O}(exp(L(n)))$. For any integer $n \in I$, let $V(n) \triangleq \{ch_{(H(n),n)}(w) \mid w \in W(n)\}$, where $H(n) = n + G(n)$ as described in the theorem.

Suppose to the contrary that $L(H(n)) = \mathrm{o}(n^r)$ and $c(H(n)) = 2^{\mathrm{o}(n^r)}$. We shall below consider the computations of M on input tapes of side-length $H(n)$.

Consider the chunk probabilities $p(v, \sigma, \tau)$ defined above. For each $(H(n), n)$-chunk v in $V(n)$, there are a total of

$$d(n) = c(H(n)) \times |PT(v(\mathrm{B}))| \times (c(H(n)) \times |PT(v(\mathrm{B}))| + 3) = \mathrm{O}(n^2\{c(H(n))\}^2)$$

chunk probabilities. Fix some ordering of the pairs (σ, τ) of starting and stopping conditions and let $P(v)$ be the vector of these $d(n)$ probabilities according to this ordering.

We first show that if $v \in V(n)$ and if p is a nonzero element of $P(v)$, then $p \geqslant 2^{-c(H(n))a(n)}$, where $a(n) = |PS(v(\mathrm{B}))| = \mathrm{O}(\{H(n)\}^2)$. Form a Markov chain $K(v)$ with states of the form (s, l), where s is a storage state of M and $l \in PS(v(\mathrm{B})) \cup PT(v(\mathrm{B}))$. The chain state (s, l) with $l \in PS(v(\mathrm{B}))$ corresponds to M being in storage state s scanning the symbol at position l of $v(\mathrm{B})$. Transition probabilities from such states are obtained from the transition probabilities of M in the obvious way. For example, if the symbol at position (i, j) of $v(\mathrm{B})$ is e, and if M in storage state s reading the symbol e can move its input head left and enter storage state s' with probability $\frac{1}{2}$, then the transition probability from state $(s, (i, j))$ to state $(s', (i, j-1))$ is $\frac{1}{2}$. Chain states of the form $(s, \overline{(i, j)})$ with $\overline{(i, j)} \in PT(v(\mathrm{B}))$ are trap states of $K(v)$ and correspond to M just having exited from $v(\mathrm{B})$ at exit point $\overline{(i, j)}$ of $v(\mathrm{B})$. Now consider, for example, $p = p(v, \sigma, \tau)$, where $\sigma = (s, \overline{(i, j)})$ and $\tau = (s', \overline{(k, l)})$ with $\overline{(i, j)}, \overline{(k, l)} \in PT(v(\mathrm{B}))$. If $p > 0$, then there must be some paths of nonzero probability in $K(v)$ from $(s, (i, j))$

to $(s', \overline{(k,l)})$, and since $K(v)$ has at most $c(H(n))a(n)$ nontrapping states, the length of the shortest path among such paths is at most $c(H(n))a(n)$. Since $\frac{1}{2}$ is the smallest nonzero transition probability of M, it follows that $p \geqslant 2^{-c(H(n))a(n)}$. If $\sigma = (s, \overline{(i,j)})$ with $\overline{(i,j)} \in PT(v(\mathrm{B}))$ and $\tau = Loop$, there must be a path Pa of nonzero probability in $K(v)$ from state $(s,(i,j))$ to some state $(s',(i',j'))$ such that there is no path of nonzero probability from $(s',(i',j'))$ to any trap state of the form $(s'', \overline{(k,l)})$ with $\overline{(k,l)} \in PT(v(\mathrm{B}))$. Again, if there is such a path Pa, there is one of length at most $c(H(n))a(n)$. The remaining cases are similar.

Fix an arbitrary $n \in I$. Divide $W(n)$ into M-equivalence classes by making w and w' *M-equivalent* if $P(ch(w))$ and $P(ch(w'))$ are zero in exactly the same coordinates, where for each $x \in W(n)$, $ch(x)$ denotes $ch_{(H(n),n)}(x) \in V(n)$.

Let $E(n)$ be a largest M-equivalence class. Then we have

$$|E(n)| \geqslant |W(n)|/2^{d(n)}.$$

Let $d'(n)$ be the number of nonzero coordinates of $P(ch(w))$ for $w \in E(n)$. Let $\hat{P}(ch(w))$ be the $d'(n)$-dimensional vector of nonzero coordinates of $P(ch(w))$. Note that $\hat{P}(ch(w)) \in [2^{-c(H(n))a(n)}, 1]^{d'(n)}$ for all $w \in E(n)$. Let $\log \hat{P}(ch(w))$ be the componentwise log of $\hat{P}(ch(w))$. Then $\log \hat{P}(ch(w)) \in [-c(H(n))a(n), 0]^{d'(n)}$. By dividing each coordinate interval $[-c(H(n))a(n), 0]$ into subintervals of length μ, we divide the space $[-c(H(n))a(n), 0]^{d'(n)}$ into at most $(c(H(n))a(n)/\mu)^{d(n)}$ cells, each of size $\mu \times \mu \times \cdots \times \mu$. We want to choose μ large enough, such that the number of cells is smaller than the size of $E(n)$, that is

$$\left(\frac{c(H(n))a(n)}{\mu}\right)^{d(n)} < \frac{|W(n)|}{2^{d(n)}}. \tag{1}$$

Concretely, we choose $\mu = 2^{-n}$. From the assumption on the rate of growth of $|W(n)|$, from the assumption that $H(n) = n + G(n)$ is bounded by some exponential function in n, and since, by assumption from the contrary, $c(H(n)) = 2^{o(n^r)}$, it follows that (1) holds for $\mu = 2^{-n}$ with large $n \in I$. Assuming (1), there must be two different $w, w' \in E(n)$ such that $\log \hat{P}(ch(w))$ and $\log \hat{P}(ch(w'))$ belong to the same cell. Therefore, if p and p' are two nonzero probabilities in the same coordinate of $P(ch(w))$ and $P(ch(w'))$, respectively, then

$$|\log p - \log p'| \leqslant \mu.$$

It follows that p and p' are 2^{μ}-close. Therefore $P(ch(w))$ and $P(ch(w'))$ are componentwise 2^{μ}-close.

For this w and w', let $u \in \Sigma^n$ be the word in Assumption (3) in the statement in the theorem (Note that u is an n-chunk over Σ). We describe two Markov chains R and R', which model the computations of M on $[u]ch(w)$ and $[u]ch(w')$, respectively. The state space of R is

$$C(H(n)) \times (PT(ch(w)(\mathrm{B})) \cup PT(u(\mathrm{B}))) \cup \{\text{Accept}, \text{Reject}, \text{Loop}\}.$$

Thus, the number of states of R is

$$z = c(H(n))(n+3+n+3)+3 = 2c(H(n))(n+3)+3.$$

The state $(s, \overline{(i,j)}) \in C(H(n)) \times PT(ch(w)(\mathrm{B}))$ of R corresponds to M just having entered $ch(w)(\mathrm{B})$ in storage state s from entrance point $\overline{(i,j)}$ of $ch(w)(\mathrm{B})$, and the state $(s', \overline{(k,l)'}) \in C(H(n)) \times PT(u(\mathrm{B}))$ of R corresponds to M just having entered $u(\mathrm{B})$ in storage state s' from entrance point $\overline{(k,l)'}$ of $u(\mathrm{B})$. For convenience sake, we assume that M begins to read any input tape x in the initial storage state $s_0 = (q_0, \lambda, 1)$, where q_0 is the initial state of M, by entering $x(1,1)$ from the lower edge of the cell on which $x(1,1)$ is written. Thus, the starting state of R is Initial $\triangleq (s_0, \overline{(2,1)'})$. The states Accept and Reject correspond to the computations halting in the accepting state and the rejecting state, respectively, and Loop means that M has entered an infinite loop. The transition probabilities of R are obtained from the chunk probabilities of M on $u(\mathrm{B})$ and $ch(w)\mathrm{B}$. For example, the transition probability from $(s, \overline{(i,j)})$ to $(s', \overline{(k,l)'})$ with $\overline{(i,j)} \in PT(ch(w)(\mathrm{B}))$ and $\overline{(k,l)'} \in PT(u(\mathrm{B}))$ is just $p(ch(w), (s, \overline{(i,j)}), (s', \overline{(k,l)}))$, the transition probability from $(s', \overline{(k,l)'})$ to $(s, \overline{(i,j)})$ with $\overline{(i,j)} \in PT(ch(w)(\mathrm{B}))$ and $\overline{(k,l)'} \in PT(u(\mathrm{B}))$ is $p(u, (s', \overline{(k,l)'}), (s, \overline{(i,j)'}))$, the transition probability from $(s, \overline{(i,j)})$ to Accept is $p(ch(w), (s, \overline{(i,j)}), \mathrm{Accept})$, and the transition probability from $(s', \overline{(k,l)'})$ to Accept is $p(u, (s', \overline{(k,l)'}), \mathrm{Accept})$. The states Accept, Reject and Loop are trap states. The chain R' is defined similarly, but using $[u]ch(w')$ in place of $[u]ch(w)$.

Suppose that $uw \in A$ and $uw' \in B$, the other case being symmetric. Let $acc(uw)$ (resp., $acc(uw')$) be the probability that M accepts input $[u]ch(w)$ (resp., $[u]ch(w')$). Then, $acc(uw)$ (resp., $acc(uw')$) is exactly the probability that the Markov chain R (resp., R') is trapped in state Accept when started in state Initial. Now $uw \in A$ implies $acc(uw) \geqslant 1-\varepsilon$. Since R and R' are 2^{μ}-close, Lemma 3.1 implies that

$$\frac{acc(uw')}{acc(uw)} \geqslant 2^{-2\mu z}.$$

$2^{-2\mu z}$ approaches 1 as n increases. Therefore, for large $n \in I$, we have

$$acc(uw') \geqslant 2^{-2\mu z}(1-\varepsilon) > \tfrac{1}{2},$$

because $\varepsilon < \frac{1}{2}$. But since $uw' \in B$, this contradicts the assumption that M separates A and B. $\square$

Remark 3.1. In the proof of Theorem 3.1, we assumed that the randomized decisions correspond to the choice from two possibilities with the probability $\frac{1}{2}$ each. In fact, we can do without this assumption. Let all the notations in this remark be defined as in the proof of Theorem 3.1, and let $p_{\min}$ be the smallest nonzero transition probability of M. Then, in the same way as in the proof of Theorem 3.1, we can show that if $v \in V(n)$ and if p is a nonzero element of $P(v)$, then $p \geqslant (p_{\min})^{c(H(n))a(n)}$. Therefore, it follows that for all $w \in E(n)$, $\hat{P}(ch(w)) \in [(p_{\min})^{c(H(n))a(n)}, 1]^{d'(n)}$, and thus $\log \hat{P}(ch(w)) \in [-(\log 1/p_{\min})c(H(n))a(n), 0]^{d'(n)}$. We again choose $\mu = 2^{-n}$. It follows that for large $n \in I$,

$$\left(\frac{(\log 1/p_{\min})c(H(n))a(n)}{\mu}\right)^{d(n)} < \frac{|W(n)|}{2^{d(n)}},$$

which corresponds to Eq. (1) in the proof of Theorem 3.1. Then, in the same way as in the proof of Theorem 3.1, we can derive a contradiction.

4. Comparison with other classes of languages

This section investigates relationships between $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ and other classes of languages.

Freivalds [3] proved a surprising fact that the language $\{a^n b^n \mid n \geqslant 1\}$ is recognized by a two-way probabilistic finite automaton with bounded error. Wang [18] proved, by using this result, that the non context-free language $\{a^n b^n c^n \mid n \geqslant 1\}$ is also recognized by a two-way probabilistic finite automaton with bounded error. On the other hand, Freivalds and Karpinski [4] proved that *Pal* (= the set of all the palindromes) is not recognized by any $\mathrm{o}(\log n)$ space-bounded probabilistic Turing machine with bounded error. From this observation, it follows that BPSPACE($\mathrm{o}(\log n)$), i.e., the class of languages recognized by $\mathrm{o}(\log n)$ space-bounded probabilistic Turing machines with bounded errors, is incomparable with the class of context-free languages.

It is easy to see that the language *Pal* is recognized by a deterministic rebound automaton, and thus in $\pounds[\mathrm{PRTM}(0)]$. From this, it follows that for any $L(n) = \mathrm{o}(\log n)$, $\pounds[\mathrm{PRTM}(L(n))]$ properly contains BPSPACE($L(n)$).

It is natural to ask what is the relationship between $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ and the class of context-free languages. We first answer this question.

Lemma 4.1. *There is a context-free language not in* $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

Proof. Let

$$L_1 = \{u2w_1 2w_2 2 \ldots 2w_k \mid k \geqslant 1 \,\&\, u \in \{0,1\}^+ \\ \&\, \forall i(1 \leqslant i \leqslant k)[w_i \in \{0,1\}^+] \,\&\, \exists j(1 \leqslant j \leqslant k)[u = {w_j}^R]\},$$

where for any word w, w^R denotes the reverse of w.

As is easily seen, L_1 is a context-free language (The proof is omitted here). By using Theorem 3.1 in the previous section, we below show that $L_1 \notin \pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

For any integer $n \geqslant 1$, let $V(n) = \{2w_1 2w_2 2 \ldots 2w_{2^n} \mid \forall i(1 \leqslant i \leqslant 2^n)[w_i \in \{0,1\}^n]\}$. For each $w = 2w_1 2w_2 2 \ldots 2w_{2^n} \in V(n)$, let $contents(w) = \{v \in \{0,1\}^n \mid v = w_i$ for some $i(1 \leqslant i \leqslant 2^n)\}$. Divide $V(n)$ into contents-equivalence classes by making w and w' *contents-equivalent* if $contents(w) = contents(w')$. There are

$$contents(n) = \binom{2^n}{1} + \binom{2^n}{2} + \cdots + \binom{2^n}{2^n} = 2^{2^n} - 1$$

contents-equivalence classes of words in $V(n)$. (Note that *contents*(n) corresponds to the number of all the nonempty subsets of $\{0,1\}^n$). We denote by $W(n)$ the set of all the representatives arbitrarily chosen from these *contents*(n) contents-equivalence classes. Let I be the set of all the positive integers. Thus, for any $n \in I$, $|W(n)| = \mathit{contents}(n) = 2^{2^n} - 1 \geqslant 2^{c^n}$ for some constant $c > 1$. It is easily seen that for every $n \in I$ and every $w, w' \in W(n)$ with $w \neq w'$, there is a word $u \in \{0,1\}^n$ such that

$$\text{either} \begin{cases} uw \in L_1 \\ uw' \in \overline{L_1} \end{cases} \quad \text{or} \quad \begin{cases} uw \in L_1, \\ uw' \in \overline{L_1}, \end{cases}$$

where for any language T, $\bar{T}$ denotes the complement of T.

Further, for each $n \in I$ and for each $w \in W(n)$, $|w| = (n+1)2^n \triangleq G(n)$, which is bounded by some exponential in n. Thus, by Theorem 3.1, if an $L(n)$ space-bounded PRTM with error probability $\varepsilon < \frac{1}{2}$ recognizes L_1, then $L(n+G(n))$ cannot be $o(n)$, and thus $L(n)$ cannot be $o(\log n)$. This completes the proof of "$L_1 \notin \pounds[\mathrm{PRTM}(o(\log n))]$". $\square$

It is well known [17] that there is a language accepted by a deterministic rebound automaton, but not generated by any context-free grammar. ($L_2 = \{a^n b^n c^n \mid n \geqslant 1\}$ is such a language.) Thus, there is a language in $\pounds[\mathrm{PRTM}(0)]$ which is not context-free. From this fact and Lemma 4.1, we have:

Theorem 4.1. $\pounds[\mathrm{PRTM}(o(\log n))]$ *is incomparable with the class of context-free languages.*

It is easy to see that the language *Pal* is accepted by a two-way deterministic one counter automaton. Thus, it follows that there is a language accepted by a two-way deterministic one counter automaton, but not in BPSPACE$(o(\log n))$. Does a similar fact hold for PRTMs? The following theorem answers the question.

Theorem 4.2. *There is a language accepted by a two-way deterministic one counter automaton, but not in* $\pounds[\mathrm{PRTM}(o(\log n))]$.

Proof. Let

$$\begin{aligned} H = \{ & u\$(e^n f)^{(n+1)^2} 0^{m_1} 1 0^{m_2} 1 \ldots 0^{m_k} \in \{a,b,c,d,e,f,\$,0,1\}^+ \mid \\ & (k, n \geqslant 0) \,\&\, u \in \{a,b,c,d\}^n \,\&\, \forall i (1 \leqslant i \leqslant k)[m_i \geqslant 0] \\ & \&\, \exists j (1 \leqslant j \leqslant k)[m_j = |u|_a (n+1)^2 + |u|_b (n+1) + |u|_c] \}, \end{aligned}$$

where $|u|_a$ (resp., $|u|_b, |u|_c$) denotes the number of a's (resp., b's, c's) occurring in u, and let

$$L_3 = H\{\$\}^*.$$

Petersen [11] showed that H is accepted by a two-way deterministic one counter automaton. Petersen's idea is straightforwardly applicable for accepting L_3 by a two-way deterministic one counter automaton.

We below show that $L_3 \notin \pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

Suppose to the contrary that there exists an $L(n)$ space-bounded PRTM M recognizing L_3 with error probability $\varepsilon < \frac{1}{2}$, where $L(n) = \mathrm{o}(\log n)$.

For any large integer n, let

- $U(n) \triangleq$ the set of all the n-chunks over $\{a, b, c, d\}$,
- $W(n) \triangleq \{\$(e^n f)^{(n+1)^2} 0^{m_1} 1 0^{m_2} 1 \ldots 0^{m_k} \$^s \mid \forall i(1 \leqslant i \leqslant p(n))[0 \leqslant m_i \leqslant p(n)-1]\,\&$ $|0^{m_1} 1 0^{m_2} 1 \ldots 0^{m_{\phi(n)}} \$^s| = p(n)^2\}$, where $p(n) = (n^3 + 6n^2 + 11n + 6)/6$, and
- $V(n) \triangleq \{ch_{(r(n)+n,n)}(w) \mid w \in W(n)\}$, where $r(n) = (n+1)^3 + p(n)^2 + 1 = \mathrm{O}(n^6)$ is the length of each word in $W(n)$.

We shall below consider the computations (using at most $L(r(n)+n)$ storage tape cells) of M on the input tapes $[u]v$ of side-length $r(n)+n$ with $v \in V(n)$ and $u \in U(n)$. For each $n \geqslant 1$, let $C(n)$ be the set of all the storage states of M using at most $L(r(n) + n)$ storage tape cells, and $c(n) = |C(n)|$. Then, $c(n) = b^{L(r(n)+n)}$ for some constant b. Consider the chunk probabilities $p(v, \sigma, \tau)$ defined before. For each $(r(n)+n, n)$-*chunk* v in $V(n)$, there are a total of

$$d(n) = c(n) \times |PT(v(\mathrm{B}))| \times (c(n) \times |PT(v(\mathrm{B}))| + 3) = \mathrm{O}(n^2 t^{L(r(n)+n)})$$

chunk probabilities for some constant t. Fix some ordering of the pairs (σ, τ) of starting and stopping conditions and let $P(v)$ be the vector of these $d(n)$ probabilities according to this ordering.

As in the proof of Theorem 3.1, it follows that if $v \in V(n)$ and if p is a nonzero element of $P(v)$, then $p \geqslant 2^{-c(n)a(n)}$, where $a(n) = |PS(v(\mathrm{B}))| = \mathrm{O}(n^{12})$.

For each $w = \$(e^n f)^{(n+1)^2} 0^{m_1} 1 0^{m_2} 1 \ldots 0^{m_{p(n)}} \$^s \in W(n)$, let $contents(w) = \{m \mid m = m_i$ for some $i(1 \leqslant i \leqslant p(n))\}$. Divide $W(n)$ into contents-equivalence classes by making w and w' *contents-equivalent* if $contents(w) = contents(w')$. There are

$$contents(n) = \binom{p(n)}{1} + \binom{p(n)}{2} + \cdots + \binom{p(n)}{p(n)} = 2^{p(n)} - 1 = 2^{\mathrm{O}(n^3)}$$

contents-equivalence classes of words in $W(n)$. We denote by $CONTENTS(n)$ the set of all the representatives arbitrarily chosen from these $contents(n)$ contents-equivalence classes. Of course, $|CONTENTS(n)| = contents(n)$. Divide $CONTENTS(n)$ into M-equivalence classes by making w and w' M-*equivalent* if $P(ch(w))$ and $P(ch(w'))$ are zero in exactly the same coordinates, where for each $x \in W(n)$, $ch(x)$ denotes $ch_{(r(n)+n,n)}(x)$. Let $E(n)$ be a largest M-equivalence class. Then we have

$$|E(n)| \geqslant contents(n)/2^{d(n)}.$$

As in the proof of Theorem 3.1, we choose μ such that

$$\left(\frac{c(n)a(n)}{\mu}\right)^{d(n)} < \frac{contents(n)}{2^{d(n)}} (\leqslant |E(n)|). \tag{2}$$

Concretely, we choose $\mu = 1/n^3$. (From the assumption that $L(n) = \mathrm{o}(\log n)$, we have $L(r(n)+n) = \mathrm{o}(\log n)$. By using this, we can easily show that for large n, (2) holds for $\mu = 1/n^3$.) Assuming (2), as in the proof of Theorem 3.1, it follows that there must be two different words w, $w' \in E(n)$ such that $P(ch(w))$ and $P(ch(w'))$ are componentwise 2^{μ}-close. For this w and w', we choose a number $m \in contents(w) - contents(w')$, and let $u \in U(n)$ be an n-chunk such that $m = |u|_a(n+1)^2 + |u|_b(n+1) + |u|_c$. (Note that for n-chunks y in $U(n)$, there are $p(n)$ different numbers $|y|_a(n+1)^2 + |y|_b(n+1) + |y|_c$.) As in the proof of Theorem 3.1, we consider two Markov chains, R and R', which model the computations of M on $[u]ch(w)$ and $[u]ch(w')$, respectively. The state space of R is

$$C(n) \times (PT(ch(w)(\mathrm{B})) \cup PT(u(\mathrm{B}))) \cup \{\mathrm{Accept}, \mathrm{Reject}, \mathrm{Loop}\}$$

and thus the number of states of R is

$$z = c(n)(n+3+n+3) + 3 = 2c(n)(n+3) + 3.$$

Similar also for R'.

Let $acc(uw)$ (resp., $acc(uw')$) be the probability that M accepts input $[u]ch(w)$ (resp., $[u]ch(w')$). Then, $acc(uw)$ (resp., $acc(uw')$) is exactly the probability that the Markov chain R (resp., R') is trapped in state Accept when started in state Initial $= (s_0, \overline{(2,1)'})$, where $s_0 = (q_0, \lambda, 1)$ and q_0 is the initial state of M. From the fact that $uw \in L_3$, it follows that $acc(uw) \geqslant 1 - \varepsilon$. Since R and R' are 2^{μ}-*close*, Lemma 3.1 implies that

$$\frac{acc(uw)}{acc(uw')} \geqslant 2^{-2\mu z}.$$

$2^{-2\mu z}$ approaches 1 as n increases. Therefore, for large n, we have

$$acc(uw') \geqslant 2^{-2\mu z}(1-\varepsilon) > \tfrac{1}{2},$$

because $\varepsilon < \frac{1}{2}$. This is a contradiction, because $uw' \notin L_3$. This completes the proof of "$L_3 \notin \pounds[\mathrm{PRTM}(L(n))]$". □

The following corollary answers an open problem in [19]:

Corollary 4.1. *There is a language accepted by a two-way deterministic one counter automaton, but not recognized by any probabilistic rebound automaton with error probability less than* $\frac{1}{2}$.

A *deterministic one-marker rebound automaton* [9] is a deterministic rebound automaton with one marker. It is shown in [19] that language L_1 in the proof of Lemma 4.1 is accepted by a deterministic one-marker rebound automaton. From this and the fact that "$L_1 \notin \pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$", we have:

Theorem 4.3. *There is a language accepted by a deterministic one-marker rebound automaton, but not in* $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

5. Space hierarchy

This section shows that there is an infinite space hierarchy for PRTMs with spaces below $\log n$.

Let $L : N \to N \cup \{0\}$ be a function. A *deterministic rebound Turing machine* (DRTM) [20] M is said to be *strongly* $L(n)$ *space-bounded* if for any $n \geqslant 1$ and for any input tape $w(\natural)$ with $|w| = n$, M on $w(\natural)$ uses at most $L(n)$ cells of the storage tape.

A function $L : N \to N \cup \{0\}$ is DRTM *space constructible* if there is a strongly $L(n)$ space-bounded DRTM M such that for each $n \geqslant 1$, there exists some input tape $w(\natural)$ with $|w| = n$ on which M halts after its storage tape head has marked off exactly $L(n)$ cells of the storage tape. (In this case, we say that M *constructs* the function L.)

Theorem 5.1. *If* $L : N \to N \cup \{0\}$ *is DRTM space constructible*, $\log\log n < L(n) = \mathrm{o}(\log n)$, *and* $L'(n) = \mathrm{o}(L(n))$, *then there exists a set accepted by some strongly* $L(n)$ *space-bounded DRTM*, *but not in* $\pounds[\mathrm{PRTM}(L'(n))]$.

Proof. Let M be a strongly $L(n)$ space-bounded DRTM which constructs the function L, and $T[L, M]$ be the following set, which depends on L and M:

$$T[L,M] = \{x \in (\Sigma \times \{0,1,2\})^+ \mid \exists n \geqslant 3[|x| = n \,\&\, \exists r \leqslant L(n)$$
$$[(\text{when the tape } h_1(x)(\natural) \text{ is presented to } M, \text{ it uses } r \text{ cells of}$$
$$\text{the storage tape and halts}) \,\&\, h_2(x) \in T(r)]]\},$$

where

(i) Σ is the input alphabet of M,
(ii) $h_1 : \Sigma \times \{0,1,2\} \to \Sigma$ is the length-preserving homomorphism such that $h_1(a,b) = a$ for any $(a,b) \in \Sigma \times \{0,1,2\}$,
(iii) $h_2 : \Sigma \times \{0,1,2\} \to \{0,1,2\}$ is the length-preserving homomorphism such that $h_2(a,b) = b$ for any $(a,b) \in \Sigma \times \{0,1,2\}$, and
(iv) for any integer $r \geqslant 1$, $T(r) = \{w_0 2 w_1 2 w_2 2 \ldots 2 w_k 2^l \mid k \geqslant 1 \,\&\, \forall i (0 \leqslant i \leqslant k)[w_i \in \{0,1\}^r] \,\&\, l \geqslant 0 \,\&\, \exists j (1 \leqslant j \leqslant k)[w_0 = w_j]\}$.

It is shown in [20] that $T[L,M]$ is accepted by a strongly $L(n)$ space-bounded DRTM. We below show that $T[L,M] \notin \pounds[\mathrm{PRTM}(L'(n))]$, where $L'(n) = \mathrm{o}(L(n))$.

The proof is similar to that of Theorem 3.1. For each integer $n \geqslant 1$, let $x(n) \in \Sigma^+$ be a fixed word such that

(i) $|x(n)| = n$ and
(ii) when $x(n)(\natural)$ is presented to M, M marks off exactly $L(n)$ cells of the storage tape and halts. (Note that for each $n \geqslant 1$, there exists such a word $x(n)$, because M constructs the function L.)

Suppose to the contrary that there exists an $L'(n)$ space-bounded PRTM M_2 recognizing $T[L,M]$ with error probability $\varepsilon < \frac{1}{2}$. For any large integer n, let

- $U(n) \triangleq \{u \in (\Sigma \times \{0,1\})^{L(n)} \mid h_1(u) = x(n)(1 : L(n))\}$,
- $W(n) \triangleq \{w \in (\Sigma \times \{0,1,2\})^{n-L(n)} \mid h_1(w) = x(n)(L(n)+1 : n)\ \&\ (h_2(w) = 2w_1 2w_2 2 \ldots 2w_{2^{L(n)}} 2^{l(n)}$ for some $w_1, w_2, \ldots, w_{2^{L(n)}}$ such that $w_i \in \{0,1\}^{L(n)}$ for each $1 \leqslant i \leqslant 2^{L(n)})$, where $l(n) = n - \{L(n) + (L(n)+1)2^{L(n)}\}$, and
- $V(n) \triangleq \{ch_{(n,L(n))}(w) \mid w \in W(n)\}$.

Note that $l(n) \geqslant 0$ and $W(n)$ is well defined for large n, because $L(n) = \mathrm{o}(\log n)$. We shall below consider the computations (using at most $L'(n)$ storage tape cells) of M_2 on the input tapes $[u]v$ of side-length n with $v \in V(n)$ and $u \in U(n)$. For each $n \geqslant 1$, let $C(n)$ be the set of all the storage states of M_2 using at most $L'(n)$ storage tape cells, and $c(n) = |C(n)|$. Then, $c(n) = b^{L'(n)}$ for some constant b. Consider the chunk probabilities $p(v,\sigma,\tau)$ defined before. For each $(n,L(n))$-*chunk* v in $V(n)$, there are a total of

$$d(n) = c(n) \times |PT(v(\mathrm{B}))| \times (c(n) \times |PT(v(\mathrm{B}))| + 3) = \mathrm{O}(L(n)^2 t^{L'(n)})$$

chunk probabilities for some constant t. Fix some ordering of the pairs (σ,τ) of starting and stopping conditions and let $P(v)$ be the vector of these $d(n)$ probabilities according to this ordering.

As in the proof of Theorem 3.1, it follows that if $v \in V(n)$ and if p is a nonzero element of $P(v)$, then $p \geqslant 2^{-c(n)a(n)}$, where $a(n) = |PS(v(\mathrm{B}))| = \mathrm{O}(n^2)$.

For each $w = W(n)$, where $h_2(w) = 2w_1 2w_2 2 \ldots 2w_{2^{L(n)}} 2^{l(n)}$, let *contents*$(w) = \{u \in U(n) \mid h_2(u) = w_i$ for some $i(1 \leqslant i \leqslant 2^{L(n)})\}$. Divide $W(n)$ into contents-equivalence classes by making w and w' *contents-equivalent* if *contents*(w) = *contents*(w'). There are

$$contents(n) = \binom{2^{L(n)}}{1} + \binom{2^{L(n)}}{2} + \cdots + \binom{2^{L(n)}}{2^{L(n)}} = 2^{2^{L(n)}} - 1$$

contents-equivalence classes of words in $W(n)$. We denote by *CONTENTS*(n) the set of all the representatives arbitrarily chosen from these *contents*(n) contents-equivalence classes. Of course, $|CONTENTS(n)| = contents(n)$. Divide *CONTENTS*$(n)$ into M_2-equivalence classes by making w and w' M_2-*equivalent* if $P(ch(w))$ and $P(ch(w'))$ are zero in exactly the same coordinates, where for each $x \in W(n)$, $ch(x)$ denotes $ch_{(n,L(n))}(x)$. Let $E(n)$ be a largest M_2-equivalence class. Then we have

$$|E(n)| \geqslant contents(n)/2^{d(n)}.$$

As in the proof of Theorem 3.1, we choose μ such that

$$\left(\frac{c(n)a(n)}{\mu}\right)^{d(n)} < \frac{contents(n)}{2^{d(n)}} (\leqslant |E(n)|). \tag{3}$$

Concretely, we choose $\mu = 2^{-L(n)}$. (From the assumption that $L'(n) = \mathrm{o}(L(n))$, by a simple calculation, it follows that for large n, (3) holds for $\mu = 2^{-L(n)}$.) Assuming (3), as in the proof of Theorem 3.1, it follows that there must be two different words w, $w' \in E(n)$ such that $P(ch(w))$ and $P(ch(w'))$ are componentwise 2^{μ}-close. For this w and w', we consider an $L(n)$-*chunk* $u \in contents(w) - contents(w')$. As in the proof of Theorem 3.1, we consider two Markov chains, R and R', which model the computations of M_2 on $[u]ch(w)$ and $[u]ch(w')$, respectively. The state space of R is

$$C(n) \times (PT(ch(w)(\mathrm{B})) \cup PT(u(\mathrm{B}))) \cup \{\text{Accept, Reject, Loop}\}$$

and thus the number of states of R is

$$z = c(n)(L(n) + 3 + L(n) + 3) + 3 = 2c(n)(L(n) + 3) + 3.$$

Similar also for R'.

Let $acc(uw)$ (resp., $acc(uw')$) be the probability that M_2 accepts input $[u]ch(w)$ (resp., $[u]ch(w')$). Then, $acc(uw)$ (resp., $acc(uw')$) is exactly the probability that the Markov chain R (resp., R') is trapped in state Accept when started in state Initial $= (s_0, \overline{(2,1)'})$, where $s_0 = (q_0, \lambda, 1)$ and q_0 is the initial state of M_2. From the fact that $uw \in T[L,M]$, it follows that $acc(uw) \geqslant 1 - \varepsilon$. Since R and R' are 2^{μ}-*close*, Lemma 3.1 implies that

$$\frac{acc(uw)}{acc(uw')} \geqslant 2^{-2\mu z}.$$

$2^{-2\mu z}$ approaches 1 as n increases. Therefore, for large n, we have

$$acc(uw') \geqslant 2^{-2\mu z}(1 - \varepsilon) > \tfrac{1}{2},$$

because $\varepsilon < \frac{1}{2}$. This is a contradiction, because $uw' \notin T[L,M]$. This completes the proof of "$T[L,M] \notin \pounds[\mathrm{PRTM}(L'(n))]$". $\square$

Since $(\log\log n)^k$, $k \geqslant 1$, is DRTM space constructible [20], it follows from Theorem 5.1 that the following corollary holds.

Corollary 5.1. *For any integer* $k \geqslant 1$, $\pounds[\mathrm{PRTM}(\log\log n)^k] \subset \pounds[\mathrm{PRTM}(\log\log n)^{k+1}]$.

Remark 5.1. A function $L : N \to N \cup \{0\}$ is DRTM *fully space constructible* if there is a strongly $L(n)$ space-bounded DRTM which, for each $n \geqslant 1$ and each input tape $w(\natural)$ with $|w| = n$, makes use of exactly $L(n)$ cells of the storage tape and halts.

We consider the following functions:

- $\log^{(1)} n = \begin{cases} 0, & (n = 0), \\ \lceil \log n \rceil, & (n \geqslant 1) \end{cases}$
  and for each $k \geqslant 1$,
- $\log^{(k+1)} n = \log^{(1)}(\log^{(k)} n)$,
- $f_k(n) = \begin{cases} \log^{(k)}(\sqrt{n/4}) & \text{if } n = 2^{2m} \text{ for some } m \geqslant 0, \\ 1 & \text{otherwise.} \end{cases}$

Recently, Petersen [10] showed that for each $k \geqslant 1$, $f_k(n)$ is fully space constructible by a two-dimensional deterministic Turing machine. By using this fact, we showed in [20] that

(i) for each $k \geqslant 1$, $f_k(n)$ is DRTM fully space constructible, and
(ii) for each $k \geqslant 2$, $f_k(n)$ space-bounded deterministic (nondeterministic) rebound Turing machines are more powerful than $f_{k+1}(n)$ space-bounded deterministic (nondeterministic) rebound Turing machines.

We conjecture that for each $k \geqslant 2$, a fact similar to (ii) above holds also for PRTMs, but we have no proof of this conjecture.

6. Closure properties

This section investigates closure properties under concatenation, Kleene + and length-preserving homomorphism.

We first consider concatenation and Kleene + operations.

Theorem 6.1. *£[PRTM(o(log n))] is not closed under concatenation and* Kleene +.

Proof. Let

$$L_3 = \{cw2w_1 2w_2 2 \ldots 2w_k \mid k \geqslant 1 \ \& \ w \in \{0,1\}^+ \\ \& \ \forall i (1 \leqslant i \leqslant k)[w_i \in \{0,1\}^+] \ \& \ w = {w_k}^R\}$$

and

$$L_4 = \{2w_1 2w_2 2 \ldots 2w_k c \mid k \geqslant 1 \ \& \ \forall i (1 \leqslant i \leqslant k)[w_i \in \{0,1\}^+]\}.$$

As is easily seen, both L_3 and L_4 are accepted by deterministic rebound automata, and thus in £[PRTM(o(log n))]. On the other hand, it is proved by using the same techniques as in the proof of Lemma 4.1 that

$$L_3 L_4 = \{cw2w_1 2w_2 2 \ldots 2w_k 2w_{k+1} c \mid k \geqslant 1 \ \& \ w \in \{0,1\}^+ \\ \& \ \forall i (1 \leqslant i \leqslant k+1)[w_i \in \{0,1\}^+] \ \& \ \exists j (1 \leqslant j \leqslant k)[w = {w_j}^R]\}$$

is not in £[PRTM(o(log n))]. Therefore, it follows that £[PRTM(o(log n))] is not closed under concatenation.

Let

$$L_5 = L_3 \cup L_4.$$

As easily seen, L_5 is in £[PRTM(o(log n))]. On the other hand,

$$L_5^+ \cap \{c\}(\{0,1\}^+\{2\})^+\{c\} = L_3 L_4 \notin £[\mathrm{PRTM}(\mathrm{o}(\log n))].$$

From the fact that $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is closed under intersection with regular languages, it follows that $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ is not closed under Kleene +. □

We next show witness languages for getting the nonclosure result under length-preserving homomorphism. Fix each integer $n \geqslant 0$, $m(n) = 2^{2n}$ and let $\equiv_{m(n)}$ be an equivalence relation defined on words in $\{a,b,\$\}^{m(n)}$. Two words $v = v_0v_1 \dots v_{m(n)-1}$ and $w = w_0w_1 \dots w_{m(n)-1}$ (with v_i, $w_i \in \{a,b,\$\}$) are equivalent ($v \equiv_{m(n)} w$) if $v_i = w_i$ for each $i \in \mathrm{POS}(m(n))$, where $\mathrm{POS}(m(n)) = \{i \in \{0,1,\dots,m(n)-1\} | i = (2^{2n-1}-2)/3 + \sum_{j=0}^{n-1} c_j \cdot 2^{2j}$ with $c_j \in \{0,1\}$ for $0 \leqslant j \leqslant n-1\}$. Intuitively two words are equivalent if they agree on symbols with offsets that have 1's in the $n-1$ least significant digits of their binary expansions that appear at odd positions. (Note that (i) $2^1 + 2^3 + 2^5 + \cdots + 2^{2(n-1)-1} = (2^{2n-1}-2)/3$, and (ii) $|\mathrm{POS}(m(n))| = 2^n$.) We use the following languages K and K' which were introduced in [11].

Let

$$K = \{u\$w_1w_2\dots w_{2^k-1} | (n,k \geqslant 1) \;\&\; u \in \{a,b\}^{m(n)-1} \;\&\; \forall i(1 \leqslant i \leqslant 2^k-1)$$
$$[w_i \in \{a,b\}^{m(n)}] \;\&\; \exists i(1 \leqslant i \leqslant 2^k-1)[u\$ \equiv_{m(n)} w_i]\},$$

where $m(n) = 2^{2n}$, and K' be like K, with the exception that exactly one of the w_i's contains marked symbols a', b' at the positions that are relevant for equivalence.

Theorem 6.2. $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$ *is not closed under length-preserving homomorphism.*

Proof. It is shown in [11] that the above language K' is accepted by a deterministic rebound automaton. Thus, $K' \in \pounds[\mathrm{PRTM}(0)]$. Let h be the length-preserving homomorphism such that $h(a) = h(a') = a$, $h(b) = h(b') = b$, $h(\$) = \$$. To prove the theorem, we below show that $K = h(K')$ is not in $\pounds[\mathrm{PRTM}(\mathrm{o}(\log n))]$.

For any integer $m(n)$, $n \geqslant 1$, let $U(m(n)) = \{u_0u_1 \dots u_{m(n)-1} \in \{a,b\}^{m(n)} | u_j = b$ for each $j \notin \mathrm{POS}(m(n))\}$, where $m(n) = 2^{2n}$, and thus $\sqrt{m(n)} = 2^n$. Note that

(i) $|\mathrm{POS}(m(n))| = 2^n = \sqrt{m(n)}$,
(ii) $|U(m(n))| = 2^{|\mathrm{POS}(m(n))|} = 2^{\sqrt{m(n)}}$, and
(iii) $u \not\equiv_{m(n)} u'$ for each u, u' ($u \neq u'$) in $U(m(n))$.

For each integer $m(n)$, $n \geqslant 1$, let $V(m(n)) = \{v_1v_2 \dots v_{2^{\sqrt{m(n)}}} | \forall i(1 \leqslant i \leqslant 2^{\sqrt{m(n)}})[v_i \in U(m(n))]\}$. For each $v = v_1v_2 \dots v_{2^{\sqrt{m(n)}}} \in V(m(n))$, let $\mathit{contents}(v) = \{u \in U(m(n)) | u = v_i$ for some $i(1 \leqslant i \leqslant 2^{\sqrt{m(n)}})\}$. Divide $V(m(n))$ into contents-equivalence classes by making v and v' *contents-equivalent* if $\mathit{contents}(v) = \mathit{contents}(v')$. There are

$$\mathit{contents}(m(n)) = \binom{2^{\sqrt{m(n)}}}{1} + \binom{2^{\sqrt{m(n)}}}{2} + \cdots + \binom{2^{\sqrt{m(n)}}}{2^{\sqrt{m(n)}}} = 2^{2^{\sqrt{m(n)}}} - 1$$

contents-equivalence classes of words in $V(m(n))$. We denote by $W(m(n))$ the set of all the representatives arbitrarily chosen from these $\mathit{contents}(m(n))$ contents-equivalence classes. Let $I=\{m(n)\mid n\geqslant 1\}$. Then, for any $m(n)\in I$, $|W(m(n))|=contents(m(n))=2^{2^{\sqrt{m(n)}}}-1\geqslant 2^{c^{\sqrt{m(n)}}}$ for some constant $c>1$. It is easily seen that for every $m(n)\in I$ and every $w,w'\in W(m(n))$ with $w\neq w'$, there is a word $u\$\in\{a,b\}^{m(n)-1}\{\$\}$ such that

$$\text{either}\begin{cases} u\$w\in K\\ u\$w'\in\bar{K}\end{cases}\quad\text{or}\quad\begin{cases} u\$w\in\bar{K},\\ u\$w'\in K.\end{cases}$$

Further, for each $m(n)\in I$ and for each $w\in W(m(n))$, $|w|=m(n)\times 2^{\sqrt{m(n)}}\triangleq G(m(n))$, which is bounded by some exponential in $m(n)$. Thus, by Theorem 3.1, if an $L(n)$ space-bounded PRTM with error probability $\varepsilon<\frac{1}{2}$ recognizes K, then $L(m(n)+G(m(n)))$ cannot be $o(\sqrt{m(n)})$, and thus $L(n)$ cannot be $o(\log n)$. This completes the proof of "$K\notin \pounds[\text{PRTM}(o(\log n))]$". $\square$

The following corollary answers an open problem in [19]:

Corollary 6.1. *The class of languages recognized by probabilistic rebound automata with error probability less than* $\frac{1}{2}$ *is not closed under length-preserving homomorphism.*

7. Conclusions

We conclude this paper by giving the following open problems.

(1) Does $\pounds[\text{PRTM}(L(n))]$ properly contain the class of languages accepted by $L(n)$ space-bounded deterministic (or nondeterministic) rebound Turing machines [20] for any $L(n)$?
(2) Is there a language in $\pounds[\text{PRTM}(o(\log n))]$ which cannot be accepted by any nondeterministic (or deterministic) two-way one counter automaton?
(3) Is there an infinite space hierarchy for PRTMs with error probability $\varepsilon<\frac{1}{2}$ whose space are below $\log\log n$?

Acknowledgements

We are very grateful to Professor H. Petersen for providing us with useful languages to write this paper. We thank the referees for suggestions for improving the exposition.

References

[1] M. Blum, C. Hewitt, Automata on a two-dimensional tape, IEEE Symp. on Switching and Automata Theory, 1967, pp. 155–160.
[2] C. Dwork, L. Stockmeyer, Finite state verifiers I: the power of interaction, J. ACM 39 (4) (1992) 800–828.
[3] R. Freivalds, Probabilistic two-way machines, Proc. Internat. Symp. on Foundations of Computer Science, Lecture Notes in Computer Science, vol. 118, Springer, Berlin, 1981, pp. 33–45.

[4] R. Freivalds, M. Karpinski, Lower space bounds for randomized computation, ICALP'94, Lecture Notes in Computer Science, vol. 820, Springer, Berlin, 1994, pp. 580–592.
[5] A.G. Greenberg, A. Weiss, A lower bound for probabilistic algorithms for finite state machines, J. Comput. Systems Sci. 33 (1986) 88–105.
[6] S.A. Greibach, Remarks on the complexity of nondeterministic counter languages, Theoret. Comput. Sci. 1 (1976) 269–288.
[7] J.D. Hopcroft, J.D. Ullman, Introduction to Automata Theory, Languages, and Computation, Addison-Wesley, Reading, MA, 1979.
[8] K. Inoue, I. Takanami, A. Nakamura, A note on two-dimensional finite automata, Inform. Process. Lett. 7 (1) (1978) 49–52.
[9] K. Inoue, I. Takanami, H. Taniguchi, A note on rebound automata, Inform. Sci. 26 (1982) 87–93.
[10] H. Petersen, Some results concerning two-dimensional Turing machines and finite automata, Proc. FCT'95, Lecture Notes in Computer Science, vol. 965, Springer, Berlin, 1995, pp. 374–382.
[11] H. Petersen, Fooling rebound automata, Proc. 24th Symp. on Mathematical Foundations of Computer Science (MFCS), Lecture Notes in Computer Science, vol. 1672, Springer, Berlin, 1999, pp. 241–250.
[12] H. Petersen, Separation results for rebound automata, Proc. 25th Symp. on Mathematical Foundations of Computer Science (MFCS), Lecture Notes in Computer Science, vol. 1893, Springer, Berlin, 2000, pp. 589–598.
[13] M.O. Rabin, Probabilistic automata, Inform and Control 6 (1963) pp. 230–245.
[14] A. Rosenfeld, Picture Languages (Formal Models for Picture Recognition), Academic Press, New York, 1979.
[15] M. Sakamoto, K. Inoue, I. Takanami, A two-way nondeterministic one-counter language not accepted by nondeterministic rebound automata, Trans. IEICE 73 (6) (1990) 879–881.
[16] E. Seneta, Non-Negative Matrices and Markov Chains, 2nd ed., Springer, New York, 1981.
[17] K. Sugata, H. Umeo, K. Morita, On the computing abilities of rebound automata, IEICE Japan Trans. (A) (1977) 367–374.
[18] J. Wang, A note on two-way probabilistic automata, Informat. Process. Lett. 43 (1992) 321–326.
[19] L. Zhang, T. Okazaki, K. Inoue, A. Ito, Y. Wang, A note on probabilistic rebound automata, IEICE Trans. Inform Systems E81-D (10) (1998) 1045–1052.
[20] L. Zhang, J. Xu, K. Inoue, A. Ito, Y. wang, Alternating rebound Turing machines, IEICE Trans. Fund. E82-A (5) (1999) 745–755.