PU. M. A. Vol. **23** (2012), No.3, pp. 291 – 315

Minimal Overlapping Embeddings and Exact Matches in Words

Brian K. Miceli
Department of Mathematics
Trinity University
San Antonio, TX 78212-7200 USA
email: bmiceli@trinity.edu

and
Jeffrey Remmel
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112 USA
email: remmel@math.ucsd.edu

(Received: October 17, 2011, and in revised form November 1, 2012)

Abstract. Let $\mathbb{N}$ denote the natural numbers $\{0,1,2,\ldots\}$. Duane and Remmel [5] found the generating function of the number of consecutive occurrences of a word $u \in \mathbb{N}^*$ over all words $w \in \mathbb{N}^*$ assuming that either u has the non-overlapping property, i.e., no two consecutive occurrences of u can share a letter, or u has the minimal overlapping property, i.e., two consecutive occurrences of u can share only one letter. We show how the techniques of Duane and Remmel can be extended to find generating functions for the number of consecutive occurrences of a certain class of words $u \in \{0,1\}^*$ over all words $w \in \{0,1\}^*$ where u has neither the non-overlapping nor the minimal overlapping property. For example, we find the generating function for number of consecutive occurrences of u in the set of words $\{0,1\}^*$ for any u of the form $0^{k_1}10^{k_2}1\cdots 0^{k_n}10^{k_{n+1}}$ where $n \geq 2$ and $\min(k_1,k_{n+1}) > \max(k_2,\ldots,k_n)$.

Mathematics Subject Classification(2010). 05A15, 05E05.

Keywords: pattern avoidance, binary word, generating function, embedding.

1 Introduction

For any alphabet A, we let A^* denote the set of all words over A and we let ϵ denote the empty word. Given words u and v in A^*, we say that u is a factor of v if there are words w_1 and w_2 such that $v = w_1uw_2$. In such a situation, we say u is a prefix of v if $w_1 = \epsilon$ and u is a suffix of v if $w_2 = \epsilon$.

Let $\mathbb{N} = \{0,1,2,\ldots\}$ and for any $m \in \mathbb{N} - \{0\}$, let $[m] = \{0,\ldots,m-1\}$. Given any word $w = w_1\ldots w_n \in \mathbb{N}^*$, we let $\sum w = w_1 + \cdots + w_n$, $|w| = n$, $z(w) = \prod_{i=1}^n z_{w_i}$, and $\mathrm{red}(w)$ be the word that results by replacing the i-th smallest integer that appears in w by $i-1$. For example, $\mathrm{red}(13443551) = 01221330$. For any word $w = w_1\ldots w_n \in \mathbb{N}^*$ and number $a \in \mathbb{N}$, we let $n_a(w)$ denote the number of occurrences of a in w.

Let $u \in \mathbb{N}^j$ and $w = w_1\ldots w_n \in \mathbb{N}^n$. We say that w has

1. a *u-match starting at position i* if $\mathrm{red}(u) = u$ and $\mathrm{red}(w_iw_{i+1}\ldots w_{i+j-1}) = u$,

2. an *exact u-match starting at position i* if $w_i w_{i+1} \dots w_{i+j-1} = u$, and
3. an *end point embedding of u starting at position i* if $w_i \geq u_1$, $w_{i+j-1} \geq u_j$ and $w_{i+r} = u_{r+1}$ for $r = 1, \dots, j-2$.

Let $u\text{-mch}(w)$, $E\text{-}u\text{-mch}(w)$, and $ep\text{-}u\text{-mch}(w)$ denote the number of u-matches in w, exact u-matches in w, and end point embeddings of u in w, respectively. For example, if $u = 201$ and $w = 201413122302$, then w has u-matches starting at positions 1, 4, 6, and 10, has an exact u-matches starting at position 1, and has end point embeddings of u starting at positions 1 and 10. Thus $u\text{-mch}(w) = 4$, $E\text{-}u\text{-mch}(w) = 1$, and $ep\text{-}u\text{-mch}(w) = 2$.

For any $k \geq 1$ and any $u \in [k]^n$ where $n \geq 1$, let

$$A_{u,k}(t, x, z_0, \dots, z_{k-1}) = \sum_{n \geq 0} t^n \sum_{w \in [k]^n} x^{E\text{-}u\text{-mch}(w)} z(w), \tag{1}$$

$$B_{u,k}(t, x, z_0, \dots, z_{k-1}) = \sum_{n \geq 0} t^n \sum_{w \in [k]^n} x^{u\text{-mch}(w)} z(w), \tag{2}$$

and

$$C_{u,k}(t, x, z_0, \dots, z_{k-1}) = \sum_{n \geq 0} t^n \sum_{w \in [k]^n} x^{ep\text{-}u\text{-mch}(w)} z(w). \tag{3}$$

The main goal of the this paper is find generating functions of the form $A_{u,2}(t, x, z_0, z_1)$ for certain classes of words $u \in \{0,1\}^*$. There has been considerable work on this type of problem in the literature. See the recent book by Heubach and Mansour [8] for many examples. Indeed there is a relatively standard method for find such generating functions. First, suppose we are given a word $u = u_1 \dots u_n \in \{0,1\}^*$. The basic idea is to classify words according to their first $n-1$ letters. That is, for any $v = v_1 \dots v_{n-1}$, we let I_v denote the set of all words $a \in \{0,1\}^*$ that start with v, i.e., $I_v = \{vy : y \in \{0,1\}^*\}$. Then we let

$$\phi_v(x,t) = \sum_{w \in I_v} x^{E\text{-}u\text{-mch}(w)} t^{|w|}.$$

The $\phi_v(x,t)$'s satisfy some simple recurrences. To begin, we notice that the words $w \in I_u$ fall into three categories: (i) $w = v$, (ii) $w = v1z$ where $z \in \{0,1\}^*$, or (iii) $w = v0z$ where $z \in \{0,1\}^*$. This leads to the recursion

$$\phi_{v_1 \dots v_{n-1}}(x,t) = t^{n-1} + t x^{\chi(v_1 \dots v_{n-1} 0 = u)} \phi_{v_2 \dots v_{n-1} 0}(x,t) + t x^{\chi(v_1 \dots v_{n-1} 1 = u)} \phi_{v_2 \dots v_{n-1} 1}(x,t), \tag{4}$$

where for any statement A, $\chi(A) = 0$ if A is false and $\chi(A) = 1$ if A is true. Putting together all the equations of the form (4), we can obtain a matrix equation $A_u[\phi] = [-t^{n-1}]$ where $[\phi]$ is a column vector of height 2^{n-1} whose entries are the $\phi_v(x,z)$'s for $v \in \{0,1\}^{n-1}$ and $[-t^{n-1}]$ is a column vector of height 2^{n-1} whose entries are equal to $-t^{n-1}$. A_u is then a $2^{n-1} \times 2^{n-1}$ matrix which will depend on u. One can prove that A_u is always invertible so that $[\phi] = A_u^{-1}[-t^{n-1}]$. If one can find an explicit expression for A_u^{-1}, then one can find explicit expressions for all the $\phi_v(x,t)$'s with $v \in \{0,1\}^{n-1}$. Then our desired generating function is

$$\sum_{w \in \{0,1\}^*} x^{E\text{-}u\text{-mch}(w)} t^{|w|} = 1 + \sum_{i=1}^{n-2} \frac{1}{(1-t)^i} + \sum_{v \in \{0,1\}^{n-1}} \phi_v(x,t). \tag{5}$$

Here the term $1+\sum_{i=1}^{n-2}\frac{1}{(1-t)^i}$ accounts for all words w whose lengths are less than or equal to $n-2$. One can easily modify the method to keep track of the number of 1's in a word as well. See Goulden and Jackson [7] or Heubach and Mansour [8] for more details. Similar methods can be used to compute $B_{u,2}(t,x,z_0,z_1)$ and $C_{u,2}(t,x,z_0,z_1)$.

The above method is easy to apply when the u is short so that the matrix A_u is not too big, however, it can be very difficult to apply when u becomes too long. For example, if $u = 0^4 1010^2 1010^5$, then u has length 17 so that the above method would require one to invert a $2^{16}\times 2^{16}$ matrix and simplify a sum with $15+2^{16}$ terms. However, the methods that we shall develop in this paper will allow us to easily compute $A_{u,2}(t,x,z_0,z_1)$ when $u = 0^4 1010^2 1010^5$.

Duane and Remmel [5] developed a different method for computing the generating functions for the number of u-matches and exact u-matches over words w in either $[k]^*$ or $\mathbb{N}^*$ for words u that have either the minimal overlapping or non-overlapping property. That is, suppose $k \geq 2$, and $u \in [k]^j$.

1. We say that u has the *k-minimal overlapping property* (*$\mathbb{N}$-minimal overlapping property*) if $\mathrm{red}(u) = u$ and the smallest i such that there exists a $w \in [k]^i$ ($w \in \mathbb{N}^i$) with u-mch$(w) = 2$ is $2j-1$. This means that in a word w, two u-matches in w can share at most one letter which must occur at the end of the first u-match and at the start of the second u-match.

2. We say that u has the *k-non-overlapping property* if $\mathrm{red}(u) = u$ and the smallest i such that there exists a $w \in [k]^i$ with u-mch$(w) = 2$ is $2j$. This means that in a word w, no two u-matches in w can share a letter.

3. We say that u has the *k-exact match minimal overlapping property* (*$\mathbb{N}$-exact match minimal overlapping property*) if the smallest i such that there exists a $w \in [k]^i$ ($w \in \mathbb{N}^i$) with E-u-mch$(w) = 2$ is $2j-1$. This means that in a word $w \in \{0,1,\ldots,k-1\}^*$, two exact u-matches in w can share at most one letter which must occur at the end of the first exact u-match and at the start of the second exact u-match.

4. We say that u has the *k-exact match non-overlapping property* (*$\mathbb{N}$-exact match non-overlapping property*) if the smallest i such that there exists a $w \in [k]^i$ ($w \in \mathbb{N}^i$) with u-mch$(w) = 2$ is $2j$. This means that in a word w, no two exact u-matches in w can share a letter.

For example, it is easy to see that $u = 010$ has the k-minimal overlapping property, the $\mathbb{N}$-minimal overlapping property, the k-exact match minimal overlapping property, and the $\mathbb{N}$-exact match minimal overlapping property for all $k \geq 2$. Now $u = 0011$ has the 2-non-overlapping property but does not have the k-non-overlapping property or the k-minimal overlapping property for $k \geq 3$ since u-mch$(001122) = 2$. We note that there is an obvious definition of the $\mathbb{N}$-non-overlapping property, but there are no such words with the property. That is, if $u = u_1 \ldots u_j$ where $u_1 = u_j$, then clearly there are consecutive matches of u which overlap. If $u_1 < u_j$ and $a = u_j - u_1$, then $u_1 \ldots u_j(u_2+a)\ldots(u_j+a)$ has u-matches which overlap while if $u_1 > u_j$ and $b = u_1 - u_j$, then $(b+u_1)\ldots(b+u_j)u_2\ldots u_j$ has u-matches which overlap.

Duane and Remmel [5] found the generating functions for the distribution of u-matches over $[k]^*$ where u has either the k-minimal overlapping property or the k-non-overlapping property and the distribution of exact u-matches over $[k]^*$ where u has either the k-exact match-minimal overlapping property or the k-exact match non-overlapping property in terms of the number of maximum packings. That is, suppose that $\mathrm{red}(u) = u$ and u has the k-minimal overlapping property. Then the shortest words $w \in [k]^*$ such that u-mch$(w) = n$ have length $j + (n-1)(j-1) = n(j-1)+1$ so we let $\mathcal{MP}^k_{u,n(j-1)+1}$ denote the set of words $w \in [k]^{n(j-1)+1}$ such that u-mch$(w) = n$. We will refer to elements of $\mathcal{MP}^k_{u,n(j-1)+1}$ as *maximum packings* for u, and we define

$$\begin{aligned} mp^k_{u,n(j-1)+1} &= |\mathcal{MP}^k_{u,n(j-1)+1}|, \\ mp^k_{u,n(j-1)+1}(z) &= \sum_{w \in \mathcal{MP}^k_{u,n(j-1)+1}} z^{\sum w}, \text{ and} \\ mp^k_{u,n(j-1)+1}(z_0,\ldots,z_{k-1}) &= \sum_{w \in \mathcal{MP}^k_{u,n(j-1)+1}} z(w). \end{aligned}$$

Similarly if red$(u) = u$ and u has the $\mathbb{N}$-minimal overlapping property, then the shortest words $w \in \mathbb{N}^*$ such that u-mch$(w) = n$ have length $n(j-1)+1$ so we let $\mathcal{MP}^{\mathbb{N}}_{u,n(j-1)+1}$ denote the set of words $w \in \mathbb{N}^{n(j-1)+1}$ such that u-mch$(w) = n$. In this case, $\mathcal{MP}^{\mathbb{N}}_{u,n(j-1)+1}$ is an infinite set of words. We will refer to elements of $\mathcal{MP}^{\mathbb{N}}_{u,n(j-1)+1}$ as *maximum packings* for u, and we define

$$\begin{aligned} mp^{\mathbb{N}}_{u,n(j-1)+1}(z) &= \sum_{w \in \mathcal{MP}^k_{u,n(j-1)+1}} z^{\sum w}, \text{ and} \\ mp^{\mathbb{N}}_{u,n(j-1)+1}(z_0,z_1,\ldots) &= \sum_{w \in \mathcal{MP}^k_{u,n(j-1)+1}} z(w). \end{aligned}$$

If u has the k-exact match minimal overlapping property, then the shortest words $w \in [k]^*$ such that E-u-mch$(w) = n$ have length $n(j-1)+1$ so we let $\mathcal{EMP}^k_{u,n(j-1)+1}$ denote the set of words $w \in [k]^{n(j-1)+1}$ such that E-u-mch$(w) = n$. We will refer to elements of $\mathcal{EMP}^k_{u,n(j-1)+1}$ as *exact match maximum packings* for u, and we define

$$\begin{aligned} emp^k_{u,n(j-1)+1} &= |\mathcal{EMP}^k_{u,n(j-1)+1}|, \\ emp^k_{u,n(j-1)+1}(z) &= \sum_{w \in \mathcal{EMP}^k_{u,n(j-1)+1}} z^{\sum w}, \text{ and} \\ emp^k_{u,n(j-1)+1}(z_0,\ldots,z_{k-1}) &= \sum_{w \in \mathcal{EMP}^k_{u,n(j-1)+1}} z(w). \end{aligned}$$

If u has the $\mathbb{N}$-exact match minimal overlapping property and $u \in [k]^*$, then we define

$$\begin{aligned} \mathcal{EMP}^{\mathbb{N}}_{u,n(j-1)+1} &= \mathcal{EMP}^k_{u,n(j-1)+1}, \\ emp^{\mathbb{N}}_{u,n(j-1)+1} &= emp^k_{u,n(j-1)+1}, \\ emp^{\mathbb{N}}_{u,n(j-1)+1}(z) &= emp^k_{u,n(j-1)+1}(z), \text{ and} \\ emp^{\mathbb{N}}_{u,n(j-1)+1}(z_0,z_1,\ldots) &= emp^k_{u,n(j-1)+1}(z_0,\ldots,z_{k-1}). \end{aligned}$$

For example, suppose that $u = 010$ and $k \geq 2$. Then the only word $w \in [k]^{2n+1}$ such that E-u-mch$(u) = n$ is $w = 0(10)^n$ so that for all $k \geq 2$ and $n \geq 1$,

$$\begin{aligned}
\mathcal{EMP}^{\mathbb{N}}_{010,2n+1} &= \mathcal{EMP}^{k}_{010,2n+1} = \{0(10)^n\},\\
emp^{\mathbb{N}}_{010,2n+1} &= emp^{k}_{010,2n+1} = 1,\\
emp^{\mathbb{N}}_{010,2n+1}(z) &= emp^{k}_{010,2n+1}(z) = z^n, \text{ and}\\
emp^{\mathbb{N}}_{010,n(j-1)+1}(z_0,\ldots,z_{k-1}) &= emp^{k}_{010,n(j-1)+1}(z_0,z_1,\ldots) = z_0^{n+1}z_1^n.
\end{aligned}$$

However, if we are just considering u-matches instead of exact u-matches in $[k]^{2n+1}$, then the only words $w \in [k]^{2n+1}$ such that u-mch$(w) = n$ are of the form $si_1si_2\ldots si_ns$ where $s \in [k-1]$ and $i_1,\ldots,i_n \in \{s+1,\ldots,k-1\}$. Thus,

$$\begin{aligned}
mp^{k}_{010,2n+1} &= \sum_{s=0}^{k-2}(k-1-s)^n = \sum_{i=1}^{k-1} i^n,\\
mp^{k}_{010,2n+1}(z) &= \sum_{s=0}^{k-2} z^{(n+1)s}(z^{s+1}[k-1-s]_z)^n = \sum_{i=1}^{k-1} z^{2n(k-i)-n-1}[i]_z^n, \text{ and}\\
mp^{k}_{010,2n+1}(z_0,\ldots,z_{k-1}) &= \sum_{s=0}^{k-2} z_s^{n+1}\left(\sum_{t=s+1}^{k-1} z_t\right)^n,
\end{aligned}$$

where $[t]_q = 1 + q + q^2 + \cdots + q^{t-1}$. By expanding our alphabet, we also get,

$$mp^{\mathbb{N}}_{010,2n+1}(z_0,z_1,\ldots) = \sum_{s=0}^{\infty} z_s^{n+1}\left(\sum_{t\geq s+1} z_t\right)^n.$$

Note that if $u = u_1\ldots u_j$ has the k-non-overlapping property, the k-exact match non-overlapping property, or the $\mathbb{N}$-exact match non-overlapping property, then the only maximum packings of u are of length j.

Finally, Duane and Remmel [5] proved the following facts for $u = u_1\ldots u_j$ where $j \geq 3$.

(I) If red$(u) = u$ and u has the k-minimal overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in[k]^n} x^{u\text{-mch}(w)} z(w) = \frac{1}{1-((z_0+\cdots+z_{k-1})t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n mp^{k}_{u,n(j-1)+1}(z_0,\ldots,z_{k-1}))}. \quad (6)$$

(II) If u has the k-exact match minimal overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in[k]^n} x^{E\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((z_0+\cdots+z_{k-1})t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n emp^{k}_{u,n(j-1)+1}(z_0,\ldots,z_{k-1}))}. \quad (7)$$

(III) If $\mathrm{red}(u) = u$ and u has the k-non-overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in[k]^n} x^{u\text{-mch}(w)} z(w) = \frac{1}{1-((z_0+\cdots+z_{k-1})t+(x-1)mp^k_{u,j}(z_0,\ldots,z_{k-1})t^j)}. \tag{8}$$

(IV) If u has the k-exact match non-overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in[k]^n} x^{E\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((z_0+\cdots+z_{k-1})t+(x-1)emp^k_{u,j}(z_0,\ldots,z_{k-1})t^j)}. \tag{9}$$

(V) If $\mathrm{red}(u) = u$ and u has the $\mathbb{N}$-minimal overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t+\sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n mp^{\mathbb{N}}_{u,n(j-1)+1}(z_0,z_1,\ldots))}. \tag{10}$$

(VI) If u has the $\mathbb{N}$-exact match minimal overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{E\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t+\sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n emp^{\mathbb{N}}_{u,n(j-1)+1}(z_0,z_1,\ldots))}. \tag{11}$$

(VII) If u has the $\mathbb{N}$-exact match non-overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in[k]^n} x^{E\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t+(x-1)emp^{\mathbb{N}}_{u,j}(z_0,z_1,\ldots)t^j)}. \tag{12}$$

The main goal of this paper is use methods similar to Duane and Remmel to give generating functions for the number of exact u-matches where $u \in \{0,1\}^j$ for $j \geq 3$ over all words $w \in \{0,1\}^*$ for certain classes of words u which do not have the 2-minimal overlapping property or the 2-non-overlapping property.

Our method will consist of two steps. First we shall prove a theorem which is analogous to (VI) above for end-point embeddings. That is, we say that u has the *end point embedding minimal overlapping property* if the smallest i such that there exists a $w \in \mathbb{N}^i$ such that $ep\text{-}u\text{-}\mathrm{mch}(w) = 2$ is $2j-1$. This means that in a word $w \in \mathbb{N}^*$, two end point embeddings of u in w can share at most one letter which must occur at the end of the first end point embedding of u and at the start of the second end point embedding of u. For example, if $u = u_1 \ldots u_j$ where $j \geq 3$ and $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$, then u will have the end point embedding minimal overlapping property. If u has the end point embedding minimal overlapping property, then the shortest words $w \in \mathbb{N}^*$ such that ep-u-mch$(w) = n$ have length $n(j-1)+1$ so we let $\mathcal{EPMP}^{\mathbb{N}}_{u,n(j-1)+1}$ denote the set of words $w \in \mathbb{N}^{n(j-1)+1}$ such that ep-u-mch$(w) = n$. We will refer to elements of $\mathcal{EPMP}^{\mathbb{N}}_{u,n(j-1)+1}$ as *end point embedding maximum packings* for u, and we define

$$\begin{aligned} epmp_{u,n(j-1)+1}(z) &= \sum_{w \in \mathcal{EPMP}^{\mathbb{N}}_{u,n(j-1)+1}} z^{\sum w}, \text{ and} \\ epmp_{u,n(j-1)+1}(z_0, z_1, \ldots) &= \sum_{w \in \mathcal{EPMP}^{\mathbb{N}}_{u,n(j-1)+1}} z(w). \end{aligned}$$

For example, suppose that $u = 5324$. Then the only words $w \in \mathbb{N}^{3n+1}$ such that ep-u-mch$(u) = n$ are of the form $w = w_1 32 w_4 32 \ldots w_{3(n-1)+1} 32 w_{3n+1}$ where $w_1 \geq 5$, $w_{3n+1} \geq 4$, and $w_{3k+1} \geq max(4,5)$ for $k = 1, \ldots, n-1$. Thus,

$$\begin{aligned} epmp_{5324,3n+1}(z) &= \frac{z^5}{1-z}\frac{z^4}{1-z}\left(\frac{z^{max(4,5)}}{1-z}\right)^{n-1} (z^2z^3)^n, \text{ and} \\ epmp_{5324,3n+1}(z_0, z_1, \ldots) &= \left(\sum_{i \geq 5} z_i\right)\left(\sum_{i \geq 4} z_i\right)\left(\sum_{i \geq max(4,5)} z_i\right)^{n-1} (z_2z_3)^n. \end{aligned}$$

We note that in general, if $u = u_1 \ldots u_j \in \mathbb{N}^*$ with $j \geq 3$ and $min(u_1, u_j) \geq max(u_2, \ldots, u_{j-1})$, then

$$\begin{aligned} epmp_{u,n(j-1)+1}(z) &= \frac{z^{u_1}}{1-z}\frac{z^{u_j}}{1-z}\left(\frac{z^{max(u_1,u_j)}}{1-z}\right)^{n-1} (z^{u_2+\cdots+u_{j-1}})^n \\ &= \frac{z^{u_1+u_j-max(u_1,u_j)}}{1-z}\left(\frac{z^{max(u_1,u_j)+\sum_{i=2}^{j-1} u_i}}{1-z}\right)^n. \end{aligned} \tag{13}$$

As in Duane and Remmel [5], we can extend our notions to sets of words $\Upsilon \subseteq \mathbb{N}^j$. That is, we say that $w = w_1 \ldots w_n$ has an *end point embedding of* Υ *starting at position* i if there exists a $u = u_1 \ldots u_j \in \Upsilon$ such that $w_i \geq u_1$, $w_{i+j-1} \geq u_j$ and $w_{i+r} = u_{r+1}$ for $r = 1, \ldots, j-2$. We let ep-Υ-mch(w) denote the number of end point embeddings of Υ in w. We say that Υ has the *end point embedding minimal overlapping property* if the smallest i such that there exists a $w \in \mathbb{N}^i$ such that ep-u-mch$(w) = 2$ is $2j-1$. This means that in a word $w \in \mathbb{N}^*$, two end point embeddings of Υ in w can share at most one letter which must occur at the end of the first end point embedding of u and at the start of the second end point embedding of u. For example, if $u = u_1 \ldots u_j$ and $v = v_1 \ldots v_j$ where $j \geq 3$ and $min(u_1, u_j, v_1, v_j) > max(u_2, \ldots, u_{j-1}, v_2, \ldots, v_{j-1})$, then $\Upsilon = \{u, v\}$ will have the end point embedding minimal overlapping property. If Υ has the end point embedding minimal overlapping property, then the shortest words $w \in \mathbb{N}^*$ such that ep-Υ-mch$(w) = n$ have length $n(j-1)+1$ so we let $\mathcal{EPMP}^{\mathbb{N}}_{\Upsilon,n(j-1)+1}$ denote the set of words $w \in \mathbb{N}^{n(j-1)+1}$ such that *ep*-Υ-mch$(w) = n$. We will refer to elements of $\mathcal{EPMP}^k_{\Upsilon,n(j-1)+1}$ as *end point embedding maximum packings* for Υ. We let

$$\begin{aligned} epmp_{\Upsilon,n(j-1)+1}(z) &= \sum_{w \in \mathcal{EPMP}^{\mathbb{N}}_{\Upsilon,n(j-1)+1}} z^{\sum w}, \text{ and} \\ epmp_{\Upsilon,n(j-1)+1}(z_0,z_1,\ldots) &= \sum_{w \in \mathcal{EPMP}^{\mathbb{N}}_{\Upsilon,n(j-1)+1}} z(w). \end{aligned}$$

We will show that one can modify the proof of (VII) to prove the following.

THEOREM 1.1 *If $\Upsilon \subseteq \mathbb{N}^j$ has the end point embedding minimal overlapping property, then*

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}\Upsilon\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n epmp_{\Upsilon,n(j-1)+1}(z_0,z_1,\ldots))}, \tag{14}$$

and

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}\Upsilon\text{-mch}(w)} z^{\sum w} = \frac{1}{1-(\frac{1}{1-z}t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n epmp_{\Upsilon,n(j-1)+1}(z))}. \tag{15}$$

The second step of our method uses a bijection $\Theta : \mathbb{N}^+ \to \{0,1\}^*$ where $\mathbb{N}^+ = \mathbb{N}^* - \{\epsilon\}$ which will allow use the generating functions of Theorem 1.1 to obtain new generating functions of the form $A_{u,2}(t,x,z_0,z_1)$ for certain classes of words u. The bijection Θ is defined as follows. If $n \in \mathbb{N}^+$ is a single letter, we define $\Theta(0) = \epsilon$ and $\Theta(n) = 0^n$, for $n \geq 1$. If $w = w_1 w_2 \ldots w_n \in \mathbb{N}^+$ where $n \geq 2$, then we let $\Theta(w_1 w_2 \ldots w_n) = 0^{w_1}10^{w_2}1\ldots 0^{w_{n-1}}10^{w_n}$. For example, if $w = 52301$, then $\Theta(w) = 0^5 10^2 10^3 110$. Since every word $u \in \{0,1\}^*$ can be written uniquely in the form $0^{w_1}10^{w_2}1\ldots 0^{w_{n-1}}10^{w_n}$ where $w_i \geq 0$, it is easy to see that Θ^{-1} is well defined. Now suppose that we consider words of the form $\Theta(u_1\ldots u_j) = 0^{u_1}10^{u_2}1\ldots 0^{u_{j-1}}10^{u_j}$ where $j \geq 3$ and $u_1, u_j > max(u_2,\ldots,u_{j-1})$. We claim that if w is a word in $\mathbb{N}^+$, then every end point embedding of $u_1\ldots u_j$ in w gives rise to an exact matching of $\Theta(u_1\ldots u_j)$ in $\Theta(w)$. That is, if $w_i \ldots w_{i+j-1}$ is an end point embedding in w, then in $\Theta(w)$, we have a subsequence $0^{w_i}10^{w_{i+1}}1\ldots 0^{w_{i+j-2}}10^{w_{i+j-1}}$ where $w_1 \geq u_1$, $w_{i+j-1} \geq u_j$ and $u_s = w_{i+s-1}$ for $s = 2,\ldots,j-1$ and, hence, this sequence will contain a unique exact match of $\Theta(u_1\ldots u_j)$. Vice versa, suppose that we have an exact match of $\Theta(u_1\ldots u_j)$ in $\Theta(w)$. Then there must exist $s,t \geq 0$ such that there is a factor $v = 0^s 0^{u_1}10^{u_2}1\ldots 0^{u_{j-1}}10^{u_j}0^t$ of $\Theta(w)$ where either v is a prefix of $\Theta(w)$ or v is preceded by a 1 in $\Theta(w)$ and either v is a suffix of $\Theta(w)$ or v is followed by a 1 in $\Theta(w)$ which means that w contains a factor $(s+u_1)u_2\ldots u_{j-1}(u_j+t)$ which is an end point embedding of u in w.

We can then use Θ and Theorem 1.1 to prove the following theorem.

THEOREM 1.2 *Suppose that* $u = u_1 \dots u_j \in \mathbb{N}^*$ *and* $\overline{u} = max(u_1, u_j)u_2 \dots u_{j-1}$ *where* $j \geq 3$ *and* $min(u_1, u_j) > max(u_2, \dots, u_{j-1})$*. Let* $\Theta(u) = 0^{u_1}10^{u_2}1 \dots 0^{u_{j-1}}10^{u_j}$*,* $a = \sum \overline{u}$*, and* $b = \sum u$*. Then*

$$\sum_{v \in \{0,1\}^*} x^{E\text{-}\Theta(u_1 \dots u_j)\text{-mch}(v)} z^{n_1(v)} t^{|v|}$$

$$= \frac{(1-t)(1-t-(tz)^{j-1}(x-1)t^a)}{tz((1-t-tz)(1-t-(tz)^{j-1}(x-1)t^a) - (tz)^j(x-1)t^b)} - \frac{1}{tz}$$

$$= \frac{1-t-z^{j-1}(x-1)(t^{a+j-1}-t^{b+j-1})}{(1-t-tz)(1-t-z^{j-1}(x-1)t^{a+j-1}) - z^j(x-1)t^{b+j}}. \quad (16)$$

Setting $x = 0$ and $z = 1$, we obtain the following generating functions for the number of words w which have no $\Theta(u_1 \dots u_j)$-matches.

COROLLARY 1.3 *Suppose that* $u = u_1 \dots u_j \in \mathbb{N}^*$ *and* $\overline{u} = max(u_1, u_j)u_2 \dots u_{j-1}$ *where* $j \geq 3$ *and* $min(u_1, u_j) > max(u_2, \dots, u_{j-1})$*. Let* $\Theta(u) = 0^{u_1}10^{u_2}1 \dots 0^{u_{j-1}}10^{u_j}$*,* $a = \sum \overline{u}$*, and* $b = \sum u$*. Then*

$$\sum_{v \in \{0,1\}^*,\ E\text{-}\Theta(u_1 \dots u_j)\text{-mch}(v)=0} t^{|v|} = \frac{1-t+t^{a+j-1}-t^{b+j-1}}{(1-2t)(1-t+t^{a+j-1})+t^{b+j}}. \quad (17)$$

For example, suppose that $u = 312$. Then $\Theta(u) = 0^31010^2$ which certainly does not have the 2-minimal overlapping property nor the 2-non-overlapping property. In this example, $\overline{u} = 31$, so that $\sum u = 6$ and $\sum \overline{u} = 4$. This yields

$$\sum_{v \in \{0,1\}^*} x^{E\text{-}0^31010^2\text{-mch}(v)} z^{n_1(v)} t^{|v|} = \frac{1-z^2(x-1)t^6(1+t)}{1-t(1+z)-z^2(x-1)t^6+z^3(x-1)t^7(1+t)}. \quad (18)$$

Setting $z = 1$ and $x = 0$ in (18) then gives the generating function for words in $\{0,1\}^*$ that have no exact 0^31010^2-matches, that is,

$$\sum_{v \in \{0,1\}^*,\ E\text{-}0^31010^2\text{-mch}(v)=0} t^{|v|} = \frac{1+t^6+t^7}{1-2t+t^6-t^7-t^8}. \quad (19)$$

One can use (19) to compute the initial sequence of the number of words $w \in \{0,1\}^n$ which have no 0^31010^2-matches, which is

$$1, 2, 4, 8, 16, 32, 64, 128, 255, 508, 1012, 2016, 4016, \dots.$$

This sequence does not appear in the On-Line Encyclopedia of Integers Sequences.

The outline of this paper is as follows. In section 2, we shall to derive Theorem 1.2 from Theorem 1.1. In addition, we shall show how Theorem 1.1 can be used to compute a variety of other generating functions of the form $A_{u,2}(t, x, z_0, z_1)$ where u has neither the the 2-minimal overlapping property or the 2-non-overlapping property. In section 3, we shall give our proof of Theorem 1.1. Finally, in section 4, we shall briefly discuss the problem of extending our results to find generating functions of the form $A_{u,k}(t, x, z_0, \dots, z_{k-1})$ where $k \geq 3$ and u has neither the the k-minimal overlapping property nor the k-non-overlapping property.

2 The proof of Theorem 1.2

In this section, we shall give several examples of how one can use Theorem 1.1 plus the bijection Θ to compute generating functions of the form $A_{u,2}(t, x, z_0, z_1)$. In particular, we will prove Theorem 1.2.

Suppose that $u = u_1 \cdots u_j$ where $j \geq 3$ and $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$. We then define

$$\overline{u} = max(u_1, u_j)u_2 \cdots u_{j-1} \text{ and } \hat{u} = u_2 + \cdots + u_{j-1}.$$

Note that $\sum u - \sum \overline{u} = u_1 + u_j - max(u_1, u_j)$. Then we know that u has the end point embedding minimal overlapping property and for $n \geq 1$,

$$\begin{aligned} t^{n(j-1)+1}(x-1)^n epmp_{u,n(j-1)+1}(z) &= t^{n(j-1)+1}(x-1)^n \frac{z^{u_1+u_j-max(u_1,u_j)}}{(1-z)} \left(\frac{z^{max(u_1,u_j)+\hat{u}}}{1-z}\right)^n \\ &= \frac{tz^{u_1+u_j-max(u_1,u_j)}}{(1-z)} \left(\frac{t^{j-1}(x-1)z^{\sum \overline{u}}}{(1-z)}\right)^n. \end{aligned}$$

Thus by (15), we have that

$$\begin{aligned} \sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} &= \frac{1}{1-\left(\frac{1}{1-z}t + \sum_{n\geq 1} \frac{tz^{u_1+u_j-max(u_1,u_j)}}{(1-z)} \left(\frac{t^{j-1}(x-1)z^{\sum \overline{u}}}{(1-z)}\right)^n\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+z^{u_1+u_j-max(u_1,u_j)} \sum_{n\geq 1} \left(\frac{t^{j-1}(x-1)z^{\sum \overline{u}}}{(1-z)}\right)^n\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+z^{u_1+u_j-max(u_1,u_j)} \frac{t^{j-1}(x-1)z^{\sum \overline{u}}}{(1-z)} \frac{1}{1-\frac{t^{j-1}(x-1)z^{\sum \overline{u}}}{(1-z)}}\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+\frac{t^{j-1}(x-1)z^{\sum u}}{(1-z-t^{j-1}(x-1)z^{\sum \overline{u}})}\right)} \\ &= \frac{1}{\frac{1-z}{(1-z)} - \frac{t}{(1-z)} - \frac{t^j(x-1)z^{\sum u}}{(1-z)(1-z-t^{j-1}(x-1)z^{\sum \overline{u}})}} \\ &= \frac{(1-z)(1-z-t^{j-1}(x-1)z^{\sum \overline{u}})}{(1-z-t)(1-z-t^{j-1}(x-1)z^{\sum \overline{u}}) - t^j(x-1)z^{\sum u}}. \end{aligned}$$

Hence, we have the following corollary of Theorem 1.1.

COROLLARY 2.1 *Suppose that $u = u_1 \ldots u_j$ and $\overline{u} = max(u_1, u_j)u_2 \ldots u_{j-1}$ where $j \geq 3$ and $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$. Then*

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \frac{(1-z)(1-z-t^{j-1}(x-1)z^{\sum \overline{u}})}{(1-z-t)(1-z-t^{j-1}(x-1)z^{\sum \overline{u}}) - t^j(x-1)z^{\sum u}}. \tag{20}$$

Note that if $|w| \geq 2$ and $\Theta(w) = v$, then v will contain $|w| - 1$ 1's so that $|v| = |w| - 1 + \sum w$. It follows that

$$\sum_{n\geq 2}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \sum_{v\in\{0,1\}^*, n_1(v)\geq 1} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-}\mathrm{mch}(v)} t^{n_1(v)+1} z^{|v|+1}. \tag{21}$$

Now,

$$\sum_{n\leq 1}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = 1 + \frac{zt}{(1-z)},$$

so that

$$\sum_{v\in\{0,1\}^*, n_1(v)\geq 1} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|}$$
$$= \frac{1}{zt}\left(\sum_{n\geq 0}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} - \left(1 + \frac{zt}{(1-z)}\right)\right). \tag{22}$$

Finally the words with no 1's in $\{0,1\}^*$ contribute $\frac{1}{1-z}$ to $\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|}$. Thus Θ shows that

$$\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|}$$
$$= \frac{1}{1-z} + \frac{1}{zt}\left(\sum_{n\geq 0}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} - \left(1 + \frac{zt}{(1-z)}\right)\right)$$
$$= \frac{1}{zt}\left(\sum_{n\geq 0}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w}\right) - \frac{1}{zt}. \tag{23}$$

Combining (20) and (23), we have the following theorem.

THEOREM 2.2 *Suppose that* $u = u_1 \ldots u_j \in \mathbb{N}^*$ *and* $\overline{u} = max(u_1, u_j)u_2 \ldots u_{j-1}$ *where* $j \geq 3$ *and* $min(u_1, u_j) > max(u_2, \ldots, u_{j-1})$. *Let* $\Theta(u) = 0^{u_1}10^{u_2}1\ldots 0^{u_{j-1}}10^{u_j}$, $a = \sum \overline{u}$, *and* $b = \sum u$. *Then*

$$\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|}$$
$$= \frac{(1-z)(1-z-(zt)^{j-1}(x-1)z^a)}{zt((1-z-zt)(1-z-(zt)^{j-1}(x-1)z^a) - (zt)^j(x-1)z^b)} - \frac{1}{zt}$$
$$= \frac{1-z-t^{j-1}(x-1)(z^{a+j-1}-z^{b+j-1})}{(1-z-zt)(1-z-t^{j-1}(x-1)z^{a+j-1}) - t^j(x-1)z^{b+j}}. \tag{24}$$

Interchanging the role of t and z in (24) yields Theorem 1.2. Note finally that

$$\begin{aligned}\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-mch}(v)}\left(\frac{z_1}{z_0}\right)^{n_1(v)}(tz_0)^{|v|} &= \sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-mch}(v)} z_1^{n_1(v)} z_0^{|v|-n_1(v)} t^{|v|} \\ &= \sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(u_1\ldots u_j)\text{-mch}(v)} z_1^{n_1(v)} z_0^{n_0(v)} t^{|v|} \\ &= = A_{\Theta(u_1\ldots u_j),2}(t,x,z_0,z_1).\end{aligned}$$

Thus we can easily obtain a closed expression for $A_{\Theta(u_1\ldots u_j),2}(t,x,z_0,z_1)$ from Theorem 1.2.

For example, suppose that $u = 41215$ so that $v = \Theta(u) = 0^4 1010^2 1010^5$. Thus v has length 17 so that the standard method described in the introduction would require us to invert $2^{16}\times 2^{16}$ matrix A_v and simplify a sum with $15+2^{16}$ terms. However in our case, we can apply Theorem 1.2 and Corollary 1.3 to prove that

$$\begin{aligned}\sum_{w\in\{0,1\}^*} x^{E\text{-}v\text{-mch}(w)} z^{n_1(w)} t^{|w|} &= \frac{1-t-z^4(x-1)(t^{13}-t^{17})}{(1-t-tz)(1-t-z^4(x-1)t^{13})-z^5(x-1)z^{18}} \text{ and,} \\ \sum_{w\in\{0,1\}^*, E\text{-}v\text{-mch}(w)=0} t^{|w|} &= \frac{1-t+t^{13}-t^{17}}{(1-2t)(1-t-t^{13})+t^{18}}.\end{aligned}$$

Next we give some further applications of Theorem 1.1 to enumerating the number of exact matches for subsets of words $\Upsilon \subseteq \{0,1\}^*$. That is, suppose that

$$\Upsilon = \{u_1 u_{2,i}\ldots u_{j-1,i}u_j : i = 1,\ldots k\},$$

where $u_1, u_j > \max(\{u_{s,i} : s = 2,\ldots,j-1, i = 1,\ldots,k\})$. An example of such a set would be $\{3004, 3014, 3104, 3114, 3124\}$. For $i = 1,\ldots,k$, we let

$$\begin{aligned} a_i &= u_{2,i}+\cdots+u_{j-1,i}, \\ b_i &= \max(u_1,u_j)+u_{2,i}+\cdots+u_{j-1,i}, \text{ and} \\ c_i &= u_1+u_j+u_{2,i}+\cdots+u_{j-1,i}.\end{aligned}$$

In this situation, we see that $v = v_1\ldots v_{n(j-1)+1}$ will be an end point maximal packing for Υ if

1. $v_1 \geq u_1$ and $v_{n(j-1)+1} \geq u_j$,
2. $v_{i(j-1)+1} \geq max(u_1,u_j)$ for $i = 1,\ldots,n-1$, and
3. $v_{i(j-1)+2},\ldots,v_{i(j-1)+j-1} = u_{2,s}\ldots u_{j-1,s}$ for some $1\leq s\leq k$.

It follows that

$$\begin{aligned} epmp_{\Upsilon,n(j-1)+1}(z) &= \frac{z^{u_1}}{1-z}\frac{z^{u_j}}{1-z}\left(\frac{z^{max(u_1,u_j)}}{1-z}\right)^{n-1}\left(\sum_{i=1}^k z^{a_i}\right)^n \\ &= \frac{z^{u_1+u_j-max(u_1,u_j)}}{1-z}\left(\frac{\sum_{i=1}^k z^{b_i}}{1-z}\right)^n.\end{aligned}$$

Thus, by (15), we have that

$$\begin{aligned}
\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}\Upsilon\text{-mch}(w)} z^{\sum w} &= \frac{1}{1-\left(\frac{1}{1-z}t+\sum_{n\geq 1}\frac{tz^{u_1+u_j-max(u_1,u_j)}}{(1-z)}\left(\frac{t^{j-1}(x-1)\sum_{i=1}^k z^{b_i}}{(1-z)}\right)^n\right)} \\
&= \frac{1}{1-\frac{t}{1-z}\left(1+z^{u_1+u_2-max(u_1,u_j)}\sum_{n\geq 1}\left(\frac{t^{j-1}(x-1)\sum_{i=1}^k z^{b_i}}{(1-z)}\right)^n\right)} \\
&= \frac{1}{1-\frac{t}{1-z}\left(1+z^{u_1+u_2-max(u_1,u_j)}\frac{\left(\frac{t^{j-1}(x-1)\sum_{i=1}^k z^{b_i}}{(1-z)}\right)}{\left(1-\frac{t^{j-1}(x-1)\sum_{i=1}^k z^{b_i}}{(1-z)}\right)}\right)} \\
&= \frac{1}{1-\frac{t}{1-z}\left(1+\frac{t^{j-1}(x-1)\sum_{i=1}^k z^{c_i}}{(1-z-t^{j-1}(x-1)\sum_{i=1}^k z^{b_i})}\right)} \\
&= \frac{1}{\frac{1-z}{(1-z)}-\frac{t}{(1-z)}-\frac{t^j(x-1)\sum_{i=1}^k z^{b_i}}{(1-z)(1-z-t^{j-1}(x-1)\sum_{i=1}^k z_i^b)}} \\
&= \frac{(1-z)(1-z-t^{j-1}(x-1)\sum_{i=1}^k z^{b_i})}{(1-z-t)(1-z-t^{j-1}(x-1)\sum_{i=1}^k z^{b_i})-t^j(x-1)\sum_{i=1}^k z^{c_i})}.
\end{aligned}$$

Using the same reasoning that we applied to get Equation (23), it follows that

$$\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(\Upsilon)\text{-mch}(v)} t^{n_1(v)} z^{|v|} = \frac{1}{zt}\left(\sum_{n\geq 0}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}\Upsilon\text{-mch}(w)} z^{\sum w}\right) - \frac{1}{zt}.$$

It follows that if $\Theta(\Upsilon) = \{\Theta(u) : u \in \Upsilon\}$, where $\Upsilon = \{u_1 u_{2,i} \ldots u_{j-1,i} u_j : i = 1, \ldots k\}$ with $u_1, u_j > \max(\{u_{s,i} : s = 2, \ldots, j-1, i = 1, \ldots, k\})$, and, for $i = 1, \ldots, k$,

$$\begin{aligned}
b_i &= \max(u_1, u_j) + u_{2,i} + \cdots + u_{j-1,i} \text{ and} \\
c_i &= u_1 + u_j + u_{2,i} + \cdots + u_{j-1,i},
\end{aligned}$$

then

$$\begin{aligned}
&\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(\Upsilon)\text{-mch}(v)} t^{n_1(v)} z^{|v|} \\
&\quad = \frac{(1-z)(1-z-(zt)^{j-1}(x-1)\sum_{i=1}^k z^{b_i})}{zt((1-z-zt)(1-z-(zt)^{j-1}(x-1)\sum_{i=1}^k z^{b_i})-(zt)^j(x-1)\sum_{i=1}^k z^{c_i})} - \frac{1}{zt} \\
&\qquad = \frac{(1-z-t^{j-1}(x-1)(\sum_{i=1}^k z^{b_i+j-1}-\sum_{i=1}^k z^{c_i+j-1})}{(1-z-zt)(1-z-t^{j-1}(x-1)\sum_{i=1}^k z^{b_i+j-1})-t^j(x-1)\sum_{i=1}^k z^{c_i+j}}. \qquad (25)
\end{aligned}$$

Interchanging the role of t and z in (25), we obtain the following theorem.

THEOREM 2.3 *Suppose that* $\Theta(\Upsilon) = \{\Theta(u) : u \in \Upsilon\}$, *where* $\Upsilon = \{u_1 u_{2,i} \ldots u_{j-1,i} u_j : i = 1, \ldots k\}$ *with* $u_1, u_j > \max(\{u_{s,i} : s = 2, \ldots, j-1, i = 1, \ldots, k\})$, *and, for* $i = 1, \ldots, k$,

$$\begin{aligned} b_i &= \max(u_1, u_j) + u_{2,i} + \cdots + u_{j-1,i} \text{ and} \\ c_i &= u_1 + u_j + u_{2,i} + \cdots + u_{j-1,i}. \end{aligned}$$

Then

$$\begin{aligned} &\sum_{v \in \{0,1\}^*} x^{E\text{-}\Theta(\Upsilon)\text{-mch}(v)} z^{n_1(v)} t^{|v|} \\ &= \frac{(1-t)(1-t-(tz)^{j-1}(x-1)\sum_{i=1}^k t^{b_i})}{tz((1-t-tz)(1-t-(tz)^{j-1}(x-1)\sum_{i=1}^k t^{b_i}) - (tz)^j (x-1)\sum_{i=1}^k t^{c_i})} - \frac{1}{tz} \\ &= \frac{(1-t-z^{j-1}(x-1)(\sum_{i=1}^k t^{b_i+j-1} - \sum_{i=1}^k t^{c_i+j-1})}{(1-t-tz)(1-t-z^{j-1}(x-1)\sum_{i=1}^k t^{b_i+j-1}) - z^j(x-1)\sum_{i=1}^k t^{c_i+j}}. \end{aligned} \tag{26}$$

Setting $x = 0$ and $t = 1$ in (26), we have the following corollary.

COROLLARY 2.4 *Suppose that* $\Theta(\Upsilon) = \{\Theta(u) : u \in \Upsilon\}$, *where* $\Upsilon = \{u_1 u_{2,i} \ldots u_{j-1,i} u_j : i = 1, \ldots k\}$ *with* $u_1, u_j > \max(\{u_{s,i} : s = 2, \ldots, j-1, i = 1, \ldots, k\})$, *and, for* $i = 1, \ldots, k$,

$$\begin{aligned} b_i &= \max(u_1, u_j) + u_{2,i} + \cdots + u_{j-1,i} \text{ and} \\ c_i &= u_1 + u_j + u_{2,i} + \cdots + u_{j-1,i}. \end{aligned}$$

Then

$$\sum_{v \in \{0,1\}^*,\ E\text{-}\Theta(\Upsilon)\text{-mch}(v)=0} t^{|v|} = \frac{1-t+(\sum_{i=1}^k t^{b_i+j-1} - \sum_{i=1}^k t^{c_i+j-1})}{(1-2t)(1-t+\sum_{i=1}^k t^{b_i+j-1}) + \sum_{i=1}^k t^{c_i+j}}. \tag{27}$$

For example, suppose $\Upsilon = \{303, 313, 323\}$. Then

$$\Theta(\Upsilon) = \{0^3 110^3, 0^3 1010^3, 0^3 10010^3\}.$$

Note that the words in $\Theta(\Upsilon)$ have different lengths. Then $\sum_{i=1}^3 t^{b_i} = t^3 + t^4 + t^5$ and $\sum_{i=1}^3 t^{c_i} = t^6 + t^7 + t^8$, and one can use Mathematica to compute that

$$\begin{aligned} \sum_{v \in \{0,1\}^*,\ E\text{-}\Theta(\Upsilon)\text{-mch}(v)=0} t^{|v|} &= \frac{1-t+t^5+t^6+t^7-t^8-t^9-t^{10}}{(1-2t)(1-t+t^5+t^6+t^7)+t^9+t^{10}+t^{11}} \\ &= \frac{1+t^5+2t^6+3t^7+2t^8+t^9}{1-2t+t^5-t^7-3t^8-2t^9-t^{10}}. \end{aligned} \tag{28}$$

Another simple example comes from considering words of the form $u = b0a1b$ where $b \geq 2$ and $a \geq 0$. If we fix a and b, then we see that u has the end point embedding minimal overlapping property. The endpoint maximum packings for u of length $4n+1$ are of the form $w = w_1 0a1 w_2 0a1 w_3 \ldots 0a1 w_n 0a1 w_{n+1}$ where $w_i \geq b$ for $i = 1, \ldots, n+1$. It follows that

$$\begin{aligned} epmp_{u,4n+1}(z) &= \left(\frac{z^b}{1-z}\right)^{n+1} (z^{1+a})^n \\ &= \frac{z^b}{1-z}\left(\frac{z^{1+a+b}}{1-z}\right)^n. \end{aligned}$$

Thus by (15), we have that

$$\begin{aligned}\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} &= \frac{1}{1-\left(\frac{1}{1-z}t + \sum_{n\geq 1}\frac{tz^b}{(1-z)}\left(\frac{t^4(x-1)z^{1+a+b}}{(1-z)}\right)^n\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+z^b\sum_{n\geq 1}\left(\frac{t^4(x-1)z^{1+a+b}}{(1-z)}\right)^n\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+z^b\frac{t^4(x-1)z^{1+a+b}}{(1-z)}\frac{1}{1-\frac{t^4(x-1)z^{1+a+b}}{(1-z)}}\right)} \\ &= \frac{1}{1-\frac{t}{1-z}\left(1+\frac{t^4(x-1)z^{1+a+2b}}{(1-z-t^4(x-1)z^{1+a+b})}\right)} \\ &= \frac{(1-z)(1-z-t^4(x-1)z^{1+a+b})}{(1-z-t)(1-z-t^4(x-1)z^{1+a+b})-t^5(x-1)z^{1+a+2b}}.\end{aligned}$$

Hence, we have the following corollary of Theorem 1.1.

COROLLARY 2.5 *Suppose that $u = b0a1b$ where $b \geq 2$ and $a \geq 0$. Then*

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \frac{(1-z)(1-z-t^4(x-1)z^{1+a+b})}{(1-z-t)(1-z-t^4(x-1)z^{1+a+b})-t^5(x-1)z^{1+a+2b}}. \tag{29}$$

Then as before, Θ shows that

$$\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(b0a1b)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|} = \frac{1}{zt}\left(\sum_{n\geq 0}(zt)^n \sum_{w\in\mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w}\right) - \frac{1}{zt}. \tag{30}$$

Combining (29) and (30), we see that

$$\begin{aligned}&\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(b0a1b)\text{-}\mathrm{mch}(v)} t^{n_1(v)} z^{|v|} = \\ &\frac{(1-z)(1-z-t^4(x-1)z^{5+a+b})}{zt((1-z-zt)(1-z-t^4(x-1)z^{5+a+b})-t^5(x-1)z^{6+a+2b})} - \frac{1}{zt} = \\ &\frac{1-z-t^4(x-1)z^{5+a+b}(1-z^b)}{(1-z-tz)(1-z-t^4(x-1)z^{5+a+b})-t^5(x-1)z^{6+a+2b}}.\end{aligned} \tag{31}$$

Interchanging the role of t and z in (31), we then obtain the following theorem.

THEOREM 2.6 *Suppose that $u = b0a1b$ where $b \geq 2$ and $a \geq 0$. Let $\Theta(u) = 0^b110^a1010^b$. Then*

$$\begin{aligned}&\sum_{v\in\{0,1\}^*} x^{E\text{-}\Theta(b0a1b)\text{-}\mathrm{mch}(v)} z^{n_1(v)} t^{|v|} \\ &= \frac{(1-t)(1-t-z^4(x-1)t^{5+a+b})}{tz((1-t-tz)(1-t-z^4(x-1)t^{5+a+b})-z^5(x-1)t^{6+a+2b})} - \frac{1}{tz} \\ &= \frac{1-t-z^4(x-1)t^{5+a+b}(1-t^b)}{(1-t-zt)(1-t-z^4(x-1)t^{5+a+b})-z^5(x-1)t^{6+a+2b}}.\end{aligned} \tag{32}$$

Setting $x = 0$ and $t = 1$, we obtain the following generating functions for the number of words w which have no $\Theta(u_1 \dots u_j)$-matches.

COROLLARY 2.7 *Suppose that $u = b0a1b$ where $b \geq 2$ and $a \geq 0$, so that $\Theta(u) = 0^b110^a1010^b$. Then*

$$\sum_{v \in \{0,1\}^*,\ E\text{-}\Theta(u_1 \dots u_j)\text{-mch}(v)=0} t^{|v|} = \frac{1 - t + t^{5+a+b}(1-t^b)}{(1-2t)(1-t+t^{5+a+b}) + t^{6+a+2b}}. \tag{33}$$

If $NM_{a,b,n}$ is equal to the number of words w in $\{0,1\}^n$ such that w has no $0^b110^a1010^b$-matches, then one can use Mathematica to show that the sequence $(NM_{0,2,n})_{n\geq 0}$ starts out

$$1, 2, 4, 8, 16, 32, 64, 128, 256, 511, 1020, 2036, 4064, 8112, 16192, 32320, 65513, \dots,$$

the sequence $(NM_{1,2,n})_{n\geq 0}$ starts out

$$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1023, 2044, 4084, 8160, 16304, 32576, 65088, \dots$$

and the sequence $(NM_{2,2,n})_{n\geq 0}$ starts out

$$1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2047, 4092, 8180, 16352, 32688, 64344, \dots.$$

None of these sequences appear in the On-Line Encyclopedia of Integer Sequences.

3 The proof of Theorem 1.1

In this section, we will prove Theorem 1.1 and two closely related results by modifying the proof of (VI) due to Duane and Remmel [5]. We will prove Theorem 1.1 by applying a ring homomorphism, defined on the ring Λ of symmetric functions over infinitely many variables, $x_1, x_2, \dots$, with coefficients in the field, $\mathbb{C}$, of complex numbers to a simple symmetric function identity. There has been a long line of research (see Beck, Remmel, and Whitehead [1], Brenti [2, 3], Dotsenko and Khoroshkin [4], Kitaev, Niedermaier, Remmel, and Riehl [9], Langley [10], Langley and Remmel [11], Mendes and Remmel [12, 13, 14], Mendes, Remmel, and Riehl [15], Ram, Remmel, and Whitehead [16], Wagner [17]) which shows that a large number of generating functions for permutation statistics can be obtained by applying homomorphisms defined on Λ to simple symmetric function identities. For example, the n-th elementary symmetric function, e_n and the n-th homogeneous symmetric function, h_n, are defined by the generating functions

$$E(t) = \sum_{n\geq 0} e_n t^n = \prod_{i\geq 1}(1 + x_i t) \tag{34}$$

and

$$H(t) = \sum_{n\geq 0} h_n t^n = \prod_{i\geq 1} \frac{1}{1 - x_i t}. \tag{35}$$

Thus,

$$H(t) = 1/E(-t). \tag{36}$$

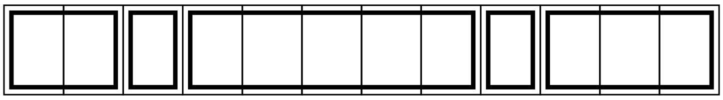

Figure 1: A brick tabloid of shape (12) and type $(1, 1, 2, 3, 5)$.

It is well known that $\{e_0, e_1, \ldots\}$ is an algebraically independent set of generators for Λ and hence we can define a ring homomorphism $\xi : \Lambda \to R$, where R is a ring, by simply specifying $\xi(e_n)$ for all $n \geq 0$. We will prove Theorem 1.1 and its extensions by applying appropriate ring homomorphisms to (36).

Let $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ be an integer partition, that is, λ is a finite sequence of weakly increasing nonnegative integers. Let $\ell(\lambda)$ denote the number of nonzero integers in λ. If the sum of these integers is n, we say that λ is a partition of n and write $\lambda \vdash n$. For any partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$, let $e_\lambda = e_{\lambda_1} \cdots e_{\lambda_\ell}$. The well-known fundamental theorem of symmetric functions says that $\{e_\lambda : \lambda \text{ is a partition}\}$ is a basis for Λ or that $\{e_0, e_1, \ldots\}$ is an algebraically independent set of generators for Λ. Similarly, if we define $h_\lambda = h_{\lambda_1} \cdots h_{\lambda_\ell}$, then $\{h_\lambda : \lambda \text{ is a partition}\}$ is also a basis for Λ.

A *brick tabloid* B of shape (n) and type $\lambda = (\lambda_1, \ldots, \lambda_k)$ is a filling of a row of n squares of cells with bricks of lengths $\lambda_1, \ldots, \lambda_k$ such that bricks do not overlap. We write $B = (b_1, \ldots, b_k)$ if the sizes of the bricks in B are $b_1, \ldots, b_k$, reading from left to right. The brick tabloid $B = (2, 1, 5, 1, 3)$ of size (12) and type $(1, 1, 2, 3, 5)$ is displayed in Figure 1.

Let $\mathcal{B}_{\lambda,n}$ denote the set of all brick tabloids of shape (n) and shape λ, and let $B_{\lambda,n} = |\mathcal{B}_{\lambda,n}|$. Eğecioğlu and Remmel [6] proved that

$$h_n = \sum_{\lambda \vdash n} (-1)^{n-\ell(\lambda)} B_{\lambda,n} e_\lambda. \tag{37}$$

We say that $w = w_1 \ldots w_n \in \mathbb{N}^n$ has a *right end point embedding of* $u = u_1 \ldots u_j$ *starting at position* i if $w_{i+j-1} \geq u_j$ and $w_{i+r} = u_{r+1}$ for $r = 0, \ldots, j-2$. Similarly, we say that $w = w_1 \ldots w_n \in \mathbb{N}^n$ has a *left end point embedding of* $u = u_1 \ldots u_j$ *starting at position* i if $w_1 \geq u_1$ and $w_{i+r} = u_{r+1}$ for $r = 1, \ldots, j-1$. Let *rep-u*-mch(w) denote the number of right end point embeddings of u in w and *lep-u*-mch(w) denote the number of left end point embeddings of u in w. We say that $u = u_1 \ldots u_j \in \mathbb{N}^j$ has the *right end point embedding minimal overlapping property* (*left end point embedding minimal overlapping property*) if the smallest i such that there exists a $w \in \mathbb{N}^i$ such that *rep-u*-mch$(w) = 2$ (*lep-u*-mch$(w) = 2$) is $2j-1$. If u has the right end point embedding minimal overlapping property, then the shortest words $w \in \mathbb{N}^*$ such that *rep-u*-mch$(w) = n$ have length $n(j-1)+1$ so we let $\mathcal{REPMP}_{u,n(j-1)+1}$ denote the set of words $w \in \mathbb{N}^{n(j-1)+1}$ such that *rep-u*-mch$(w) = n$. We will refer to elements of $\mathcal{REPMP}^k_{u,n(j-1)+1}$ as *right end point embedding maximum packings* for u. We let

$$\begin{aligned} repmp_{u,n(j-1)+1}(z) &= \sum_{w \in \mathcal{REPMP}_{u,n(j-1)+1}} z^{\sum w}, \text{ and} \\ repmp_{u,n(j-1)+1}(z_0, z_1, \ldots) &= \sum_{w \in \mathcal{REPMP}_{u,n(j-1)+1}} z(w). \end{aligned}$$

Similarly if u has the left end point embedding minimal overlapping property, then the shortest words $w \in \mathbb{N}^*$ such that *lep-u*-mch$(w) = n$ have length $n(j-1)+1$, so we let $\mathcal{LEPMP}_{u,n(j-1)+1}$ denote the set of words $w \in \mathbb{N}^{n(j-1)+1}$ such that *lep*-u-mch$(w) = n$. We will refer to elements of $\mathcal{LEPMP}^k_{u,n(j-1)+1}$ as *left end point embedding maximum packings* for u. We let

$$lepmp_{u,n(j-1)+1}(z) = \sum_{w \in \mathcal{REPMP}_{u,n(j-1)+1}} z^{\sum w}, \text{ and}$$
$$lepmp_{u,n(j-1)+1}(z_0, z_1, \ldots) = \sum_{w \in \mathcal{REPMP}_{u,n(j-1)+1}} z(w).$$

Then we have the following theorem.

THEOREM 3.1 *Let $u \in \mathbb{N}^j$ where $j \geq 3$.*

1. *If u has the end point embedding minimal overlapping property, then*

$$\sum_{n\geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{ep\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n epmp_{u,n(j-1)+1}(z_0,z_1,\ldots))}. \quad (38)$$

2. *If u has the right end point embedding minimal overlapping property, then*

$$\sum_{n\geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{rep\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n repmp_{u,n(j-1)+1}(z_0,z_1,\ldots))}. \quad (39)$$

3. *If u has the left end point embedding minimal overlapping property, then*

$$\sum_{n\geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{lep\text{-}u\text{-mch}(w)} z(w) = \frac{1}{1-((\sum_{i\geq 0} z_i)t + \sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n lepmp_{u,n(j-1)+1}(z_0,z_1,\ldots))}. \quad (40)$$

Note: We will only prove the first part of this theorem, as the proofs of the subsequent parts of the theorem are the same except that we use right end point maximal packings and left end point maximal packings in place of end point maximal packing throughout the proof.

Proof. Suppose that $u \in \mathbb{N}^j$ has the endpoint embedding minimal overlapping property. We define a ring homomorphism Γ on Λ by letting

1. $\Gamma(e_0) = 1$,
2. $\Gamma(e_1) = \sum_{i\geq 0} z_i$,
3. $\Gamma(e_{s(j-1)+1}) = (-1)^{s(j-1)}(x-1)^s epmp_{u,s(j-1)+1}(z_0, z_1, \ldots)$ for all $s \geq 1$, and

4. $\Gamma(e_n) = 0$ if $n \notin \{1\} \cup \{s(j-1)+1 : s \geq 1\}$.

Note that if $\Gamma(e_n) \neq 0$ and $n \geq 1$, then the sign associated with $\Gamma(e_n)$ is just $(-1)^{n-1}$.

We claim that for all $n \geq 1$,

$$\Gamma(h_n) = \sum_{w \in \mathbb{N}^n} x^{ep\text{-}u\text{-mch}(w)} z(w). \tag{41}$$

By (37), we have that

$$\Gamma(h_n) = \sum_{\mu \vdash n} (-1)^{n-\ell(\mu)} B_{\mu,n} \Gamma(e_\mu). \tag{42}$$

Now if μ is not a partition whose parts come from $\{1\} \cup \{s(j-1)+1 : s \geq 1\}$, then $\Gamma(e_\mu) = 0$, so we let $P_{j,n}$ denote the set of all partitions of n whose parts come from $\{1\} \cup \{s(j-1)+1 : s \geq 1\}$. It follows that if we set

$$A(i) = \left(\sum_{i \geq 0} z_i\right) \chi(i=1) + (x-1)^{(i-1)/(j-1)} epmp_{u,i}(z_0, z_1, \ldots) \chi(i = n(j-1)+1 > 1),$$

then

$$\begin{aligned}
\Gamma(h_n) &= \sum_{\mu \in P_{j,n}} (-1)^{n-\ell(\mu)} \sum_{(b_1,\ldots,b_{\ell(\mu)}) \in \mathcal{B}_{\mu,n}} \prod_{i=1}^{\ell(\mu)} \Gamma(e_{b_i}) \\
&= \sum_{\mu \in P_{j,n}} (-1)^{n-\ell(\mu)} \sum_{(b_1,\ldots,b_{\ell(\mu)}) \in \mathcal{B}_{\mu,n}} \prod_{i=1}^{\ell(\mu)} (-1)^{b_i - 1} A(b_i) \\
&= \sum_{\mu \in P_{j,n}} \sum_{(b_1,\ldots,b_{\ell(\mu)}) \in \mathcal{B}_{\mu,n}} \prod_{i=1}^{\ell(\mu)} A(b_i).
\end{aligned} \tag{43}$$

Next we want to give a combinatorial interpretation to the right hand side of (43). Suppose that we have a brick tabloid $B = (b_1, \ldots, b_\ell)$ of shape (n) such that $b_i \in \{1\} \cup \{s(j-1)+1 : s \geq 1\}$ for all i. Then if $b_i = 1$, we will interpret $\sum_{i \geq 0} z_i$ as allowing us to fill the cell corresponding to b_i with any letter from $\mathbb{N}$. If $b_i = s(j-1)+1 > 1$, then we will interpret the term $(x-1)^{(b_i-1)/(j-1)} epmp_{u,b_i}(z_0, z_1, \ldots) = (x-1)^s epmp_{u,s(j-1)+1}(z_0, \ldots, z_{k-1})$ as allowing us to fill the cells corresponding to b_i with a word $v_i \in \mathcal{EPMP}^{\mathbb{N}}_{u,s(j-1)+1}$ and then labeling each cell in b_i which is the start of an end point embedding of u in v_i with either x or -1. Let $\mathcal{O}_{u,n}$ denote the set of all labeled brick tabloids that can be constructed in this way. Thus an $O \in \mathcal{O}_{u,n}$ will consist of a triple $T = (B, w, L)$ where

1. $B = (b_1, \ldots, b_\ell)$ is brick tabloid of shape (n) such that $b_i \in \{1\} \cup \{s(j-1)+1 : s \geq 1\}$ for all i,
2. $w = w_1 \ldots w_n \in \mathbb{N}^n$ is a word such that w_i lies in i-th cell of B for $i = 1, \ldots, n$,
3. if $b_i = 1$, the letter in the cell corresponding to b_i can be any letter in $\mathbb{N}$, and

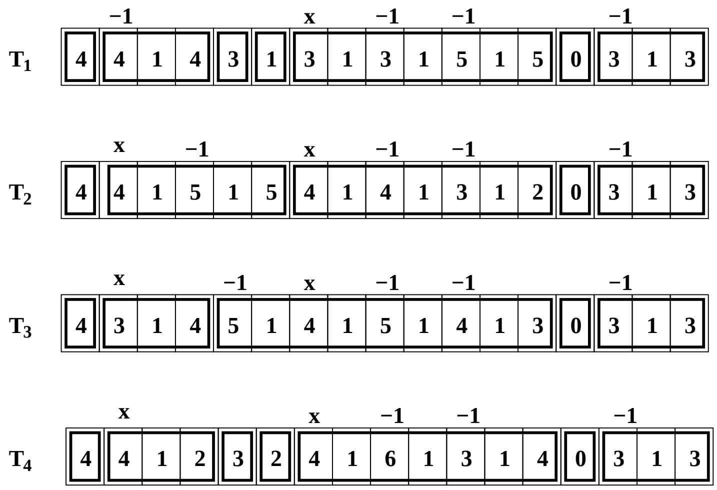

Figure 2: Elements of $\mathcal{O}_{312,17}$.

4. if $b_i = s(j-1)+1 > 1$, then cells of b_i are filled with a word v_i such that v_i is an end point embedding maximum packing for u of size $s(j-1)+1$, each cell of b_i which corresponds to the start of an endpoint embedding of u in v_i is labeled with either -1 or x, and where the collection of all -1's and x's is L.

We then define the *weight of* T, $wt(T)$, to be $z(w)$ times the product of the x labels in T and the *sign of* T, $sign(T)$, to be the product of the -1 labels in T. For example, if $u = 312$, then in Figure 2, we have pictured four elements of $\mathcal{O}_{213,17}$, $T_i = (B^{(i)}, w^{(i)}, L^{(i)})$ for $i = 1, \ldots, 4$. Then $wt(T^{(1)}) = xz_0z_1^6z_3^5z_4^3z_5^2$, $sign(T^{(1)}) = 1$, $wt(T^{(2)}) = x^2z_0z_1^6z_2z_3^3z_4^4z_5^2$, $sign(T^{(2)}) = 1$, $wt(T^{(3)}) = x^2z_0z_1^6z_3^4z_4^4z_5^2$, $sign(T^{(3)}) = 1$, $wt(T^{(4)}) = x^2z_0z_1^5z_2^2z_3^3z_4^4z_6$, and $sign(T^{(4)}) = -1$. It follows that

$$\Gamma(h_n) = \sum_{T \in \mathcal{O}_{u,n}} sign(T)wt(T). \tag{44}$$

Next we define a sign-reversing, weight-preserving involution $I : \mathcal{O}_{u,n} \to \mathcal{O}_{u,n}$. If $T = (B, w, L)$, then to define $I(T)$, we scan the cells of $B = (b_1, \ldots, b_\ell)$ from left to right looking for the first time where one of the following cases hold.

Case 1. There is a brick b_i of size j whose first cell is labeled with -1. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the brick b_i by j bricks of size 1, $w^* = w$, and L^* arises from L by removing the label -1 from the first cell of b_i.

Case 2. There are j consecutive bricks of size 1 in B, $b_i, b_{i+1}, \ldots, b_{i+j-1}$ such that the letters in these cells form an end point embedding of u. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the bricks $b_i, b_{i+1}, \ldots, b_{i+j-1}$ by a single brick b of size j, $w^* = w$, and L^* arises from L by adding a -1 label on the first cell of b.

Case 3. There is a brick b_i of size $(s+1)(j-1)+1$ where $s \geq 1$ such that all the labels on b_i are x's except for the cell which is j cells from the right which is labeled with -1. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the brick b_i by a brick of size $s(j-1)+1$ followed by $j-1$ bricks of size 1, $w^* = w$, and L^* arises from L by removing the -1 label that was in b_i.

Case 4. There are j consecutive bricks in B, $b_i, b_{i+1}, \ldots, b_{i+j-1}$ such that $b_i = s(j-1)+1 > 1$ and $b_{i+1}, \ldots b_{i+j-1}$ are of size 1, all the labels on b_i are x's, and the letters in these bricks form an end point maximum packing for u of size $(s+1)(j-1)+1$. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the bricks $b_i, b_{i+1}, \ldots, b_{i+j-1}$ by a single brick b of size $(s+1)(j-1)+1$, $w^* = w$, and L^* arises from L by adding a -1 label on the last cell of b_i.

Case 5. There is a brick b_i of size $(s+1)(j-1)+1$ where $s \geq 1$ such that the first cell of b_i is labeled with -1. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the brick b_i by $j-1$ bricks of size 1 followed by a brick of size $s(j-1)+1$, $w^* = w$, and L^* arises from L by removing the -1 label that was on the first cell of b_i.

Case 6. There are j consecutive bricks in B, $b_i, b_{i+1}, \ldots, b_{i+j-1}$ such that $b_i, \ldots b_{i+j-2}$ are bricks of size 1 and $b_{i+j-1} = s(j-1)+1 > 1$ and the letters in these bricks form an end point maximum packing for u of size $(s+1)(j-1)+1$. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the bricks $b_i, b_{i+1}, \ldots, b_{i+j-1}$ by a single brick b of size $(s+1)(j-1)+1$, $w^* = w$, and L^* arises from L by adding a -1 label on the first cell of b.

Case 7. There is a brick b_i of size $s(j-1)+1$ where $s \geq 3$ such that the first cell is labeled with an x and there is a cell which has a label -1 which is not the j-th cell from the right. Let the t-th cell of b_i be the leftmost cell of b_i which is labeled with -1. Then in this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the brick b_i by j consecutive bricks $c_1, c_2, \ldots, c_{j-1}, c_j$ where c_1 contains all the cells of b_i up to and including cell t, $c_2, \ldots, c_{j-1}$ are bricks of size 1, and c_j contains the remaining cells of b_i, $w = w^*$, and L^* is the labeling that results from L by removing the -1 label from cell t.

Case 8. There are j consecutive bricks $b_i, b_{i+1}, \ldots, b_{i+j-1}$ such that $b_i = c(s-1)+1 > 1$ and has only x labels, $b_{i+1}, \ldots, b_{i+j-2}$ are bricks of size 1, $b_{i+j-1} = d(j-1)+1 > 1$, and the letters in these three bricks form an end point maximum packing for u of size $(c+d+1)(j-1)+1$. In this case, $I(T) = (B^*, w^*, L^*)$ where B^* results from B by replacing the j bricks $b_i, b_{i+1}, \ldots, b_{i+j-1}$ by a single brick b, $w^* = w$, and L^* results from L by adding a label -1 on the last cell of b_i.

If none of Cases 1-8 apply, then we set $I(T) = T$.

For example, consider the images of T_1, T_2, T_3, and T_4 which are pictured in Figure 3. We see that T_1 is in Case 1 so that $I(T_1)$ results by replacing the second brick by three bricks of size 1 and removing the -1 label. This results in $I(T_1)$ pictured in the first row of Figure 3. It is then easy to see that $I(T_1)$ is in Case 2 so that $I^2(T_1) = T_1$. T_2 is in Case 3 where b_i is the second brick. Thus we obtain $I(T_2)$ by replacing b_2 with a brick of size 3 followed by two bricks of size 1 and removing the -1 label from cell 4. $I(T_2)$ is pictured in the second row of Figure 3.

t is then easy to see that $I(T_2)$ will be in Case 4 so that $I^2(T_2) = T_2$. T_3 is in Case 5 where b_i is

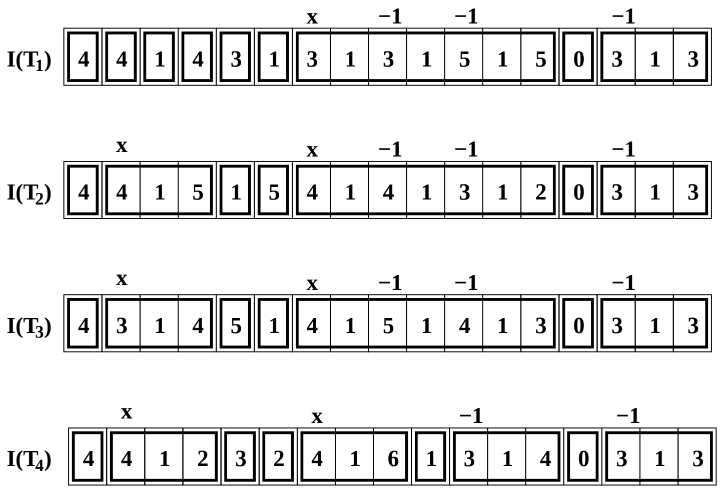

Figure 3: The images under I of the elements in Figure 2.

the third brick. Thus we obtain $I(T_3)$ by replacing b_3 by two bricks of size 1 followed by a brick of size 7 and removing the -1 label on cell 5. $I(T_3)$ is pictured in the third row of Figure 3. It is then easy to see that $I(T_3)$ will be in Case 6 so that $I^2(T_3) = T_3$. Finally, T_4 is in Case 7 with $t = 4$ so that we replace the fifth brick by three consecutive bricks of sizes 3, 1, and 3, reading from left to right, and remove the -1 label for cell 4 of brick b_5. $I(T_4)$ is pictured in the forth row of Figure 3. It is then easy to see that $I(T_4)$ will be in Case 8 so that $I^2(T_4) = T_4$.

It is not difficult to show that I is an involution. The argument here is exactly the same as the argument used by Duane and Remmel [5] to prove (11), and by construction, if $I(T) \neq T$, then $sign(T)wt(T) = -sign(I(T))wt(I(T))$. Hence I shows that

$$\Gamma(h_n) = \sum_{T \in \mathcal{O}_{u,n},\ I(T)=T} sign(T)wt(T). \tag{45}$$

Thus we must examine the fixed points of I. Suppose that $T = (B, w, L)$ is a fixed point of I where $B = (b_1, \ldots, b_\ell)$. There cannot be any -1 labels on any of the bricks in B, since otherwise we could apply one of Cases 1, 3, 5, or 7. Thus if $I(T) = T$, $sign(T) = 1$. It follows that $wt(T) = x^c z(w)$ where c is the number of endpoint embeddings of u in w that lie entirely with in some brick b_i in B. We claim that any end point embedding in w must lie entirely within some brick. That is, suppose that $w = w_1 \ldots w_n$ where $w_s w_{s+1} \ldots w_{s+j-1}$ is an end point embedding of u that does not lie in a single brick. Because u has the minimal overlapping property, there are only four possibilities, namely,

(i) cells $s, s+1, \ldots, s+j-1$ are covered by bricks of size 1,

(ii) cell s is part of brick b_i of size > 1 and cells $s+1, \ldots, s+j-1$ are covered by bricks of size 1,

(iii) cell $s+j-1$ is part of brick b_i of size > 1 and cells $s, \ldots, s+j-2$ are covered by bricks of size 1, or

(iv) cell s is part of a brick b_i of size > 1, $b_{i+1}, \ldots, b_{i+j-2}$ are bricks of size 1 covering cells $s+1, \ldots, s+j-2$, and cell $s+j-1$ is part of brick b_{i+j-1} which is of size > 1.

In case (i), we could apply Case 2 of the definition of I to cells $s, s+1, \ldots, s+j-1$; in case (ii), we could apply Case 4 of the definition of I to cells of b_i plus cells $s+1, \ldots, s+j-1$; in case (iii), we could apply Case 6 of the definition of I to cells $s, \ldots, s+j-2$ plus the cells of b_i; in case (iv), we can apply Case 8 of the definition of I to the cells of contained in the bricks $b_i, \ldots, b_{i+j-1}$. Thus in all the cases (i)-(iv), it would be the case that $I(T) \neq T$ which contradicts our choice of T. Thus we have shown that if $I(T) = T$, then $sign(T)wt(T) = x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z(w)$. Finally note that if $w \in [k]^n$, then we can construct a fixed point of I by placing bricks which cover the maximal length end point maximum packings in w, covering the remaining cells by bricks of size 1, and labeling the start of each end point embedding in w by x. It thus follows that

$$\Gamma(h_n) = \sum_{w \in \{0, \ldots, k-1\}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z(w).$$

But then,

$$\begin{aligned}
\Gamma(H(t)) &= \sum_{n \geq 0} t^n \sum_{w \in [k]^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z(w) \\
&= \Gamma(1/E(-t)) \\
&= \frac{1}{1 + \sum_{n \geq 1} (-t)^n \Gamma(e_n)} \\
&= \frac{1}{1 - ((\sum_{i \geq 0} z_i)t + \sum_{n \geq 1} t^{n(j-1)+1} (x-1)^n epmp_{u,n(j-1)+1}(z_0, z_1, \ldots))},
\end{aligned}$$

which proves (38). □

We can see that if $\Upsilon \subseteq \mathbb{N}^j$ has the end point minimal overlapping property, then the proof of Theorem 3.1 goes through if we replace u-matches with Υ-matches and ep-u-mch(w) with ep-Υ-mch(w). Thus, a minor modification of the proof of Theorem 3.1 proves Theorem 1.1, as desired. Further, replacing z_i by z^i for all $i \geq 0$ in Theorem 3.1 yields the following corollary.

Corollary 3.2 *Let $u \in \mathbb{N}^j$ where $j \geq 3$.*

1. *If u has the end point embedding minimal overlapping property, then*

$$\sum_{n \geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{ep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \frac{1}{1 - ((\frac{1}{(1-z)})t + \sum_{n \geq 1} t^{n(j-1)+1} (x-1)^n epmp_{u,n(j-1)+1}(z))}. \tag{46}$$

2. *If u has the right end point embedding minimal overlapping property, then*

$$\sum_{n \geq 0} t^n \sum_{w \in \mathbb{N}^n} x^{lep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \frac{1}{1 - ((\frac{1}{(1-z)})t + \sum_{n \geq 1} t^{n(j-1)+1} (x-1)^n repmp_{u,n(j-1)+1}(z))}. \tag{47}$$

3. If u has the left end point embedding minimal overlapping property, then

$$\sum_{n\geq 0} t^n \sum_{w\in\mathbb{N}^n} x^{rep\text{-}u\text{-}\mathrm{mch}(w)} z^{\sum w} = \frac{1}{1-((\frac{1}{(1-z)})t+\sum_{n\geq 1} t^{n(j-1)+1}(x-1)^n lepmp_{u,n(j-1)+1}(z))}. \quad (48)$$

4 Future work

It is possible to extend our results on exact matchings to cover words in $[k]^*$ where $k \geq 3$ instead of $\{0,1\}^*$. The obvious thing to do is to code words in $\mathbb{N}^*$ as words in $[k]^*$. For example, if $k = 3$, then we could let $c(2n) = 0^n$ and $c(2n+1) = 1^{n+1}$ for all $n \geq 0$. Then code of a word $w = w_1 \dots w_n$ can be defined $c(w) = c(w_1)2c(w_2)2\dots c(w_{n-1})2c(w_n)$. For example, $c(502143) = 1^3 2202120^2 21^2$. The problem with this type of coding is that not all words in $\{0,1,2\}^*$ are of the form $c(w)$ for some $w \in \mathbb{N}^*$. For example, a word like 0101012010210101 is not of the form $c(w)$ for any word in $\mathbb{N}^*$. This makes it more difficult to transfer results about the distribution of end point embeddings to results about exact matchings of words in $\{0,1,2\}^*$. We have developed methods to do this, but they are considerably more complicated than the one presented in this paper. This will be the subject of a forthcoming paper of the authors and Thomas Langley.

References

[1] D. Beck, J. Remmel and T. Whitehead, *The combinatorics of transition matrices between the bases of the symmetric functions and the B_n analogues*, Discrete Math., 153 (1996) 3–27.

[2] F. Brenti, *Permutation enumeration, symmetric functions, and unimodality*, Pacific J. Math., 157 (1993) 1–28.

[3] F. Brenti, *A class of q-symmetric functions arising from plethysm*, J. Combin. Theory Ser. A, 91 (2000) 137–170.

[4] V. Dotsenko and A. Khoroshkin, *Anick-type resolutions and consecutive pattern avoidance*, arXiv:math/1002.2761.

[5] A. Duane and J. B. Remmel, *Minimal overlapping patterns in colored permutations*, Electron. J. Combin., 18 (2011) Article P25 (34 pp.).

[6] Ö. Eğecioğlu and J. Remmel, *Brick tabloids and connection matrices between bases of symmetric functions*, Discrete Appl. Math., 34 (1991) 107–120.

[7] I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, A Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons Inc, 1983.

[8] S. Heubach and T. Mansour, *Combinatorics and Compostions and Words*, CRC Press, 2010.

[9] S. Kitaev, A. Niedermaier, J. B. Remmel and A. Riehl, *Generalized pattern matching conditions for $C_k \wr S_n$*, ISRN Combinatorics, Volume 2013, Article ID 634823, 21 pages.

[10] T. M. Langley, *Alternative transition matrices for Brenti's q-symmetric functions and a class of q,t-symmetric functions on the hyperoctahedral group*, Proceedings of FPSAC 2002 (Formal Power Series and Algebraic Combinatorics), Melbourne, Australia.

[11] T. M. Langley and J. B. Remmel, *Enumeration of m-tuples of permuations and a new class of power bases for the space of symmetric functions*, Adv. in Appl. Math., 36 (2006) 30–66.

[12] A. Mendes and J. B. Remmel, *Permutations and words counted by consecutive patterns*, Adv. in Appl. Math., 37 (2006) 443–480.

[13] A. Mendes and J. Remmel, *Generating functions for statistics on $C_k \wr S_n$*, Sém. Lothar. Combin., 54A (2005/07) Art. B54At, (40 pp.).

[14] A. Mendes and J. B. Remmel, *Generating functions from symmetric functions*, preprint.

[15] A. Mendes, J. B. Remmel and A. Riehl, *Permutations with k-regular descent patterns*, in: S. Linton, N. Ruškuc, V. Vatter (Eds.), Permutation Patterns, London Math. Soc. Lecture Note Ser., 376, Cambridge Univ. Press, Cambridge, 2010, 259–286.

[16] D. Ram, J. B. Remmel and T. Whitehead, *Combinatorics of the q-basis of symmetric functions*, J. Comb. Theory Ser. A, 76 (1996) 231–271.

[17] J. D. Wagner, *The permutation enumeration of wreath products and cyclic and symmetric groups*, Adv. in Appl. Math., 30 (2003) 343–368.