

# On the Enumeration of Non-crossing Pairings of Well-balanced Binary Strings

Paul R. F. Schumacher\*

Texas A&M University at Qatar, Doha, Qatar

Catherine H. Yan†

Department of Mathematics, Texas A&M University,  
College Station, Texas

October 24, 2011

## Abstract

A non-crossing pairing on a binary string pairs ones and zeroes such that the arcs representing the pairings are non-crossing. A binary string is well-balanced if it is of the form  $1^{a_1}0^{a_1}1^{a_2}0^{a_2}\dots 1^{a_r}0^{a_r}$ . In this paper we establish connections between non-crossing pairings of well-balanced binary strings and various lattice paths in plane. We show that for well-balanced binary strings with  $a_1 \leq a_2 \leq \dots \leq a_r$ , the number of non-crossing pairings is equal to the number of lattice paths on the plane with certain right boundary, and hence can be enumerated by differential Goncarov polynomials. For the regular binary strings  $S = (1^k0^k)^n$ , the number of non-crossing pairings is given by the  $(k+1)$ -Catalan numbers. We present a simple bijective proof for this case.

## 1 Introduction

Motivated by problems that arose from the theory of random matrices and free probability, Kemp, Mahlborg, Rattan, and Smyth[4] studied non-crossing pairings of binary strings. Given a binary string  $S = 1^{a_1}0^{b_1}1^{a_2}0^{b_2}\dots$ , a non-crossing pairing on  $S$  matches ones and zeroes of  $S$  such that the arcs connecting the ones and zeroes are non-crossing if one puts the bits along a line and draws all the arcs above the bits. A binary string is *balanced* if the number of ones is the same as the number of zeroes. For a balanced binary string  $S$ , let  $NC_2(S)$  be the set of non-crossing pairings of  $S$ . It is easy to see that  $NC_2(S)$

---

\*Email: paul.schumacher@qatar.tamu.edu.

†Yan was supported in part by NSF grant DMS-0653846 and NSA grant H98230-11-1-0167. Email: cyan@math.tamu.edu.

is always non-empty. In [4], Kemp et.al. presented recurrence relations and functional equations satisfied by  $|NC_2(S)|$ , and proved general upper bounds in terms of the  $k$ -Catalan numbers. In this paper we examine a subset of this problem and further explore its properties.

**Definition 1.1.** *Let  $S = 1^{a_1}0^{b_1}1^{a_2}0^{b_2} \dots 1^{a_r}0^{a_r}$  be a binary string. We say that  $S$  is well-balanced if  $a_i = b_i$  for all  $i$ . We say that  $S$  is regular if  $a_i = b_j$  for all  $i, j$ .*

The main objective of this paper is to establish a connection between non-crossing pairings of well-balanced binary strings and lattice paths with a general right boundary. It is known in [4] that the non-crossing pairings for a regular binary string are counted by the  $k$ -Catalan numbers. In section 2 we give a bijective proof to complement the algebraic argument of [4]. We then list several lattice path representations for non-crossing pairings of regular binary strings. In section 3, we prove that for a well-balanced binary string  $S = 1^{a_1}0^{a_1} \dots 1^{a_{r+1}}0^{a_{i+1}}$  with  $a_1 \leq a_2 \leq \dots \leq a_{r+1}$ , the number of non-crossing pairings can be computed as the number of lattice paths from  $(0, 0)$  to  $(s_r, r)$  with the right boundary  $(s_1, s_2, \dots, s_r)$ , where  $s_i = a_1 + \dots + a_i$  for  $i = 1, 2, \dots, r$ . This result allows us to study  $NC_2(S)$  by the powerful machinery in symbolic computation and umbral calculus. As corollaries, we give a determinant formula, a linear recurrence relation, and a (shifted) generating function for  $|NC_2(S)|$ . Finally in Section 4, we describe a bijective proof for the well-balanced strings.

## 2 Regular Binary Strings

The number of non-crossing pairings for a regular binary string  $S = (1^k0^k)^n$  is equal to the  $n$ th  $(k + 1)$ -Catalan number. This result was stated by Kemp et al. [4, Prop 1.6], where they outlined an algebraic proof analyzing the recurrence relations of both structures. In this section we present a simple bijective proof, mapping non-crossing pairings on a regular binary string to a  $k$ -Catalan structure.

The Catalan structure we will be using is the generalization of representation (n) of Stanley's online list [13]. The basic structure for  $k = 2$  is  $n$  non-crossing chords pairing  $2n$  (labeled) points on a circle. The  $k$ -Catalan version is non-intersecting regions in a circle, each containing  $k$  of the  $nk$  points on the circle. An equivalent definition is given in [1, item (10)], where it is described as a family of  $n$  non-intersecting chord paths of length  $k - 1$  joining  $kn$  points on the circumference of a circle. For each region in the first description, the chord path can be obtained by joining the points in this region in clockwise order. The chord paths are non-intersecting since they are contained in non-intersecting regions. Alternatively, one can create chord groups where the first point in the region (in clockwise order) is connected to each other point in the region with a chord, as in [12]. Again these chord groups are non-intersecting with each other.

All three representations are equivalent. See Fig. 1 for an illustration with  $k = 3$  and  $n = 2$ , where the rightmost point on the circle is the initial point. In the following we shall use the first description in the bijection. The third one is convenient to extend to Catalan structures with a left weight, as demonstrated in [12].

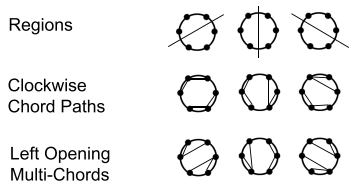


Figure 1: Three versions of representation (n) for the 2nd 3-Catalan number.

**Theorem 2.1.** *Let  $S = (1^k 0^k)^n$  be a regular binary string. Then  $|NC_2(S)| = C_n^{k+1} = \frac{1}{(k+1)n+1} \binom{(k+1)n+1}{n}$ , the  $n$ th  $(k+1)$ -Catalan number.*

*Proof.* We present a bijection between the set of non-crossing pairings for the binary string  $S = (1^k 0^k)^n$  and the set of ways to divide a circle inscribed with  $(k+1)n$  points into  $n$  non-intersecting regions each with  $k+1$  points. Given a non-crossing pairing of ones and zeroes of  $S$ , we first arrange the  $2kn$  01-bits clockwise on a circle, starting from the position  $(1, 0)$ . That is, the first one of  $S$  corresponds to the rightmost point on the circle. The pairing is then represented by  $kn$  non-crossing chords of the circle. Next we compress each consecutive run of  $k$  ones to a single point, and hence obtain  $n$  points  $Q_1, \dots, Q_n$ , where each  $Q_i$  is connected to  $k$  distinct zeroes on the circle with non-crossing chords. Now there are  $(k+1)n$  points in total. Taking the convex hull of  $Q_i$  and its  $k$  neighbors for each  $i$  gives  $n$  non-intersecting regions.

The inverse map is slightly more complex. Given a circle with  $n(k+1)$  points inscribed, divided into  $n$  distinct regions of  $k+1$  points, start with the fixed starting point and mark

the points in the positions  $1, (k+1)+1, 2(k+1)+1, \dots, (n-1)(k+1)+1$ . We claim that each region contains exactly one marked point. To see this, it is enough to show that no two marked points can belong to the same region. Assume otherwise and the points  $P_i$  and  $P_j$  at positions  $i(k+1)+1$  and  $j(k+1)+1$  (where  $i < j$ ) are in the same region  $A$ . Since the regions are non-intersecting, between any two consecutive points in the same region there are  $(k+1)t$  points for some integer  $t \geq 0$ . Assume that there are  $y$  points lying strictly between  $P_i$  and  $P_j$  that are also in the region  $A$ . Then  $0 \leq y \leq k-1$  and the total number of points between  $P_i$  and  $P_j$  is of the form  $y + (k+1)t$ . This leads to the equation  $y + (k+1)t = (j-i)(k+1) - 1$ , which implies that  $y \equiv -1 \pmod{k+1}$ . This contradicts the fact that  $0 \leq y \leq k-1$ .

Now we expand each marked point into  $k$  points labeled by 1, and label each unmarked point by 0. Within each region there are  $2k$  points,  $k$  with labeled 1 and  $k$  with 0. When reading clockwise, they are of the form  $0^t 1^k 0^{k-t}$ , which is equivalent to  $1^k 0^k$  by rotation and having exactly one non-crossing pairing. Laying all the points on a line and putting the non-crossing pairing for each region together gives us the non-crossing pairing of the appropriate regular binary string.  $\square$

As an example, in Fig. 2 we illustrate the correspondence between non-intersecting regions and non-crossing pairings of the regular binary string  $(1^2 0^2)^2$ .

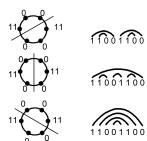


Figure 2: The corresponding regions and pairings of the binary string  $(1^2 0^2)^2$ .

The Catalan family contains many combinatorial structures, notably lattice paths. We list here some families of lattice paths that are also enumerated by the  $n$ th  $(k+1)$ -Catalan

number, and hence equi-numerated with the non-crossing pairings of the regular binary strings. For all the cases of Corollary 2.2, bijective proofs are known: cases 1 and 2 are proved in [12], case 3 is given in [1] and case 4 in [6]. In the next section we will prove that lattice paths with a given right boundary plays an important role in the enumeration of non-crossing pairings of well-balanced binary strings. This result can be viewed as a generalization of case 4.

**Corollary 2.2.** *Let  $S = (1^k 0^k)^n$  be a regular binary string. Then the set of non-crossing pairings on  $S$  corresponds to the following sets of lattice paths.*

1. *Lattice paths from  $(0, 0)$  to  $(kn, kn)$  with steps  $(0, k)$  and  $(1, 0)$ , never rising above the line  $y = x$ . Equivalently, this is the set of sequences from  $(0, 0)$  to  $(2kn, 0)$  with steps  $(k, k)$  and  $(1, -1)$ , never falling below the  $x$ -axis.*
2. *Lattice paths from  $(0, 0)$  to  $(2(kn + 1), 0)$ , starting with a  $(1, 1)$  step with the remaining steps in the path being  $n$  steps  $(k, k)$  and  $kn + 1$  steps  $(1, -1)$ , never falling below the  $x$ -axis, such that any maximal sequence of consecutive  $(1, -1)$  steps ending on the  $x$ -axis has odd length.*
3. *Lattice paths from  $(0, 0)$  to  $(kn, n)$  with steps  $(1, 0)$  or  $(1, 0)$ , never lying below the line  $x = ky$ .*
4. *Lattice paths from  $(0, 0)$  to  $(kn, n - 1)$  with steps  $(1, 0)$  or  $(1, 0)$ , never touching the line  $x = ky + k + 1$ .*

### 3 Well-balanced binary strings

For a well-balanced binary string  $S = 1^{a_1} 0^{a_1} 1^{a_2} 0^{a_2} \dots 1^{a_r} 0^{a_r} 1^{a_{r+1}} 0^{a_{r+1}}$ , we say that  $S$  consists of  $r + 1$  segments, where the  $i$ th segment is  $1^{a_i} 0^{a_i}$ . Again let  $NC_2(S)$  be the set of non-crossing pairings of ones and zeroes of  $S$ . Following [4], we denote by  $\phi^*(a_1, a_2, \dots, a_r, a_{r+1})$  the cardinality of  $NC_2(S)$ .

In this section we prove that for any sequence of non-decreasing integers  $a_1 \leq a_2 \dots \leq a_r \leq a_{r+1}$ ,  $\phi^*(a_1, \dots, a_{r+1})$  equals the number of lattice paths in the plane from  $(0, 0)$  to  $(x, r)$  with the right boundary  $(s_1, s_2, \dots, s_r)$ , where  $s_i = a_1 + \dots + a_i$ . This result leads to many algebraic properties and explicit formulas for  $\phi^*$ .

A lattice path  $P$  is a path in the plane with two kinds of steps: a unit north step  $N = (0, 1)$  or a unit east step  $E = (1, 0)$ . If  $x$  is a positive integer, a lattice path from the origin  $(0, 0)$  to the point  $(x, n)$  can be coded by a length  $n$  non-decreasing sequence  $(x_1, x_2, \dots, x_n)$ , where  $0 \leq x_i \leq x$  and  $x_i$  is the  $x$ -coordinate of the  $i$ -th north step. For example, the path  $EENENNEE$  is coded by  $(2, 3, 3)$ . In other words, the  $i$ -th north step of the path has coordinates  $(x_i, i - 1) \rightarrow (x_i, i)$ .

In general, let  $\mathbf{s}$  be a non-decreasing sequence with positive integer terms  $s_1, s_2, \dots, s_n$ . A lattice path from  $(0, 0)$  to  $(x, n)$  is one with the (weak) right boundary  $\mathbf{s}$  if  $x_i \leq s_i$  for  $1 \leq i \leq n$ . If  $x \geq s_n$ , then the number of lattice paths from  $(0, 0)$  to  $(x, n)$  with the

right boundary  $\mathbf{s}$  does not depend on  $x$ . Let  $Path_n(\mathbf{s})$  be the set of lattice paths from  $(0, 0)$  to  $(x, n)$  (where  $x \geq s_n$ ) with the right boundary  $\mathbf{s}$ , and  $LP_n(\mathbf{s})$  be the cardinality of  $Path_n(\mathbf{s})$ .

Given a sequence of positive integers  $a_1 \leq a_2 \leq \dots \leq a_r \leq a_{r+1}$ , we first notice a basic fact.

**Lemma 3.1.**  $\phi^*(a_1, a_2, \dots, a_r, a_{r+1})$  is independent of the value of  $a_{r+1}$ . That is,

$$\phi^*(a_1, a_2, \dots, a_r, a_{r+1}) = \phi^*(a_1, a_2, \dots, a_r, a_r) \quad \text{for all } a_{r+1} \geq a_r. \quad (1)$$

*Proof.* To see this, for any  $i > a_r$  let  $\alpha$  be the  $i$ th one in the last segment  $1^{a_{r+1}}0^{a_{r+1}}$ . For any zero before  $\alpha$ , it is easy to check that there are more ones than zeroes between this zero and  $\alpha$ . Hence in any non-crossing pairing  $\alpha$  must be paired with a zero in the last segment. There is only one way to do so, namely, we pair the last  $a_{r+1} - a_r$  ones in the  $(r+1)$ st segment with the  $a_{r+1} - a_r$  zeroes immediately after them. Hence (1) follows.  $\square$

Clearly,  $\phi^*(a) = 1$ . By convention, set  $\phi^*(\emptyset) = 1$ . Our main result is the following equation of  $\phi^*(a_1, a_2, \dots, a_r, a_{r+1})$ .

**Theorem 3.2.** For a sequence of positive integers  $a_1 \leq a_2 \leq \dots \leq a_r \leq a_{r+1}$  with  $r \geq 1$ , let

$$\begin{aligned} s_1 &= a_1 \\ s_2 &= a_1 + a_2 \\ \dots &= \dots \\ s_r &= a_1 + a_2 + \dots + a_r. \end{aligned}$$

Then  $\phi^*(a_1, a_2, \dots, a_r, a_{r+1})$  equals the number of lattice paths with the right boundary  $(s_1, s_2, \dots, s_r)$ , that is

$$\phi^*(a_1, a_2, \dots, a_r, a_{r+1}) = LP_r(s_1, s_2, \dots, s_r). \quad (2)$$

We prove the theorem by a double induction on the length  $r$  and the first term  $a_1$ . First, when  $r = 1$ , the string  $S$  is of the form  $1^{a_1}0^{a_1}1^{a_2}0^{a_2}$  where  $a_1 \leq a_2$ . Given a non-crossing pairing, among the  $a_1$  1s in the first segment, assume that  $i$  of them are paired with zeroes in the first segment, and  $a_1 - i$  of them to zeroes in the second segment. There is only one way to arrange such a pairing. See Figure 3 for an illustration. As  $i$  ranges from 0 to  $a_1$ , we have  $\phi^*(a_1, a_2) = a_1 + 1 = LP_1(a_1)$ .

Let assume that equation (2) holds for all sequences of length less than or equal to  $r$ . We shall show that it is true for sequences of length  $r$ . To achieve this, we do induction on the first term  $a_1$ .

First we show equation (2) for  $a_1 = 1$ . This is achieved by showing that  $\phi^*(a_1, \dots, a_{r+1})$  and  $LP_r(s_1, \dots, s_r)$  satisfy the same recurrence relation.

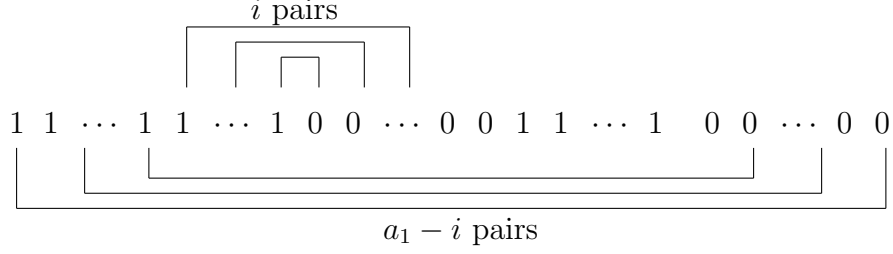


Figure 3: non-crossing pairings of  $1^{a_1}0^{a_1}1^{a_2}0^{a_2}$ .

**Lemma 3.3.** *The number of non-crossing pairings on well-balanced strings satisfies the recurrence*

$$\phi^*(1, a_2, \dots, a_r, a_{r+1}) = \sum_{i=1}^{r+1} \phi^*(a_2, \dots, a_i) \phi^*(a_{i+1}, \dots, a_{r+1}).$$

where we use the convention that  $\phi^*(\emptyset) = 1$ .

*Proof.* Each summand represents a choice of which segment to match the first zero in the string  $S = 101^{a_2}0^{a_2} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . If we match it to the one at the head of  $S$ , then we are left with the binary string  $1^{a_2}0^{a_2} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . There are  $\phi^*(a_2, \dots, a_{r+1})$  many non-crossing pairings in this case. Otherwise, we have to match the first zero to the very first one in the  $(i+1)$ st segment, for some  $i = 1, 2, \dots, r$ . Call this pair the *first pair*. In that case, the string  $S$  is broken into two pieces, one is nested under the first pair with bits  $1^{a_2}0^{a_2} \dots 1^{a_i}0^{a_i}$ , and the other is outside the first pair with bits  $1^{a_{i+1}}0^{a_{i+1}} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . They contribute  $\phi^*(a_2, \dots, a_i) \phi^*(a_{i+1}, \dots, a_{r+1})$  non-crossing pairings.  $\square$

Next we prove that the number of lattice paths with the right boundary  $(s_1, s_2, \dots, s_r)$ , where  $s_1 = 1$ , and  $s_i = 1 + a_2 + \dots + a_i$  for  $i = 2, 3, \dots, r$ , satisfies a similar recurrence.

**Lemma 3.4.**

$$LP_r(1, s_2, \dots, s_r) = LP_{r-1}(s_2 - 1, s_3 - 1, \dots, s_r - 1) + \sum_{i=1}^r LP_{i-1}(s_2 - 1, \dots, s_{i-1} - 1) \cdot LP_{r-i-1}(a_{i+1}, a_{i+1} + a_{i+2}, \dots, a_{i+1} + \dots + a_r).$$

*Proof.* Let  $x$  be any number larger than  $s_r$ . For lattice paths from  $(0, 0)$  to  $(x, r)$  with the right boundary  $(s_1, s_2, \dots, s_r)$ , we partition them according to the minimal  $i$  such that the point  $(s_i, i - 1)$  is on the lattice path.

Case 1. The path does not contain the lattice point  $(s_i, i - 1)$  for any  $i = 1, 2, \dots, r$ . Such a lattice path is strictly on the left of the boundary  $(s_1, s_2, \dots, s_r)$ . They are counted by  $LP_r(0, s_2 - 1, s_3 - 1, \dots, s_r - 1) = LP_{r-1}(s_2 - 1, s_3 - 1, \dots, s_r - 1)$ .

Case 2. The lattice path does not pass the points  $(s_1, 0), (s_2, 1), \dots, (s_{i-1}, i - 2)$ , but contains  $(s_i, i - 1)$ . Then the part from  $(0, 0)$  to  $(s_i, i - 1)$  is strictly on the left of  $(s_1 = 1, s_2, \dots, s_{i-1})$  and is counted by  $LP_{i-1}(0, s_2 - 1, s_3 - 1, \dots, s_{i-1} - 1) = LP_{i-2}(s_2 - 1, \dots, s_{i-1} - 1)$ . The next step on the path must be an  $N$ -step from  $(s_i, i - 1)$  to  $(s_i, i)$ . The

part from  $(s_i, i)$  to  $(x, r)$  is bounded by  $(s_{i+1}, s_{i+2}, \dots, s_r)$  and is counted by  $LP_{r-i}(s_{i+1} - s_i, s_{i+2} - s_i, \dots, s_r - s_i)$ . Summing over case 2 with  $i = 1, 2, \dots, r$  and combining with case one gives the desired recurrence formula.  $\square$

By the inductive hypothesis, we have that

$$LP_{r-1}(s_2 - 1, \dots, s_r - 1) = \phi^*(s_2 - 1, \dots, s_r - s_{r-1}, *) = \phi^*(a_2, \dots, a_r, a_{r+1}).$$

Similarly,

$$LP_{i-2}(s_2 - 1, \dots, s_{i-1} - 1) = \begin{cases} 1 & \text{if } i = 1 \\ \phi^*(a_2, \dots, a_i) & \text{for } i \geq 2 \end{cases}$$

$$LP_{r-i}(s_{i+1} - s_i, s_{i+2} - s_i, \dots, s_r - s_i) = \phi^*(a_{i+1}, \dots, a_r, a_{r+1}).$$

Hence Lemmas 3.3 and 3.4 imply equation (2) for  $a_1 = 1$ .

Now assume equation (2) is true for all non-decreasing integer sequences of length less than or equal to  $r$ , and true for all non-decreasing sequences of length  $r + 1$  with the first term less than  $a_1$ . We shall show it is true for the sequence  $a_1 \leq a_2 \leq \dots \leq a_r \leq a_{r+1}$ . This is again achieved by showing that  $\phi^*(a_1, \dots, a_{r+1})$  and  $LP_r(s_1, \dots, s_r)$  satisfy the same recurrence relation. We note that Lemma 3.5 has been given in [4, Thm 3.3]. Since the argument is short and parallel to that of Lemma 3.3, we include it here for completeness.

**Lemma 3.5.** *Let  $a_1 > 1$ . We have*

$$\phi^*(a_1, \dots, a_r, a_{r+1}) = \sum_{i=1}^{r+1} \phi^*(a_1 - 1, a_2, \dots, a_i) \phi^*(a_{i+1}, \dots, a_{r+1}). \quad (3)$$

*Proof.* Let  $\alpha$  be the first zero in the binary string  $S = 1^{a_1}0^{a_1} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . We partition the non-crossing pairings of the zeroes and ones of  $S$  according to the one that pairs with  $\alpha$ .

Case 1.  $\alpha$  is paired with the last 1 in the first segment. Such non-crossing pairings are counted by  $\phi^*(a_1 - 1, a_2, \dots, a_{r+1})$ . This is the last term in the summation of (3).

Case 2.  $\alpha$  is paired with the bit  $\beta$ , which is a 1 in the  $(i+1)$ th segment, where  $1 \leq i \leq r$ . Then  $\beta$  must be at the  $a_1$  position in the  $(i+1)$ th segment. Hence the string  $S$  is broken into two pieces, one is nested under the pair  $(\alpha, \beta)$  with bits  $0^{a_1-1}1^{a_2}0^{a_2} \dots 1^{a_i}0^{a_i}1^{a_1-1}$  which is rotated to the string  $1^{a_1-1}0^{a_1-1}1^{a_2}0^{a_2} \dots 1^{a_i}0^{a_i}$ , and the other is outside the pair  $(\alpha, \beta)$  with bits  $1^{a_{i+1}}0^{a_{i+1}} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . This gives the product

$$\phi^*(a_1 - 1, a_2, \dots, a_i) \phi^*(a_{i+1}, \dots, a_{r+1})$$

for  $1 \leq i \leq r$ .

Combining the above two cases, we obtain the recurrence (3).  $\square$

Let  $s_i = a_1 + \dots + a_i$  for  $i = 1, 2, \dots, r$ . Next we consider lattice paths with the right boundary  $(s_1, s_2, \dots, s_r)$ .

**Lemma 3.6.** For  $s_1 > 1$ , we have

$$LP_r(s_1, s_2, \dots, s_r) = \sum_{i=1}^{r+1} LP_{i-1}(s_1 - 1, \dots, s_{i-1} - 1) \cdot LP_{r-i}(a_{i+1}, a_{i+1} + a_{i+2}, \dots, a_{i+1} + \dots + a_r). \quad (4)$$

*Proof.* Again we partition the lattice paths from  $(0, 0)$  to  $(x, r)$  (where  $x \geq s_r$ ) with the right boundary  $(s_1, s_2, \dots, s_r)$  by the minimal  $i$  such that the path contains the point  $(s_i, i - 1)$ .

Case 1. The path does not contain any of the points  $\{(s_i, i - 1) : i = 1, 2, \dots, r\}$ . There are  $LP_r((s_1 - 1, s_2 - 1, \dots, s_r - 1))$  many such paths. This gives the last term in the summation of (4).

Case 2. The path does not contain  $(s_1, 0), \dots, (s_{i-1}, i - 2)$  but contains  $(s_i, i - 1)$ . Then the part from  $(0, 0)$  to  $(s_i, i - 1)$  is strictly on the left of  $(s_1, s_2, \dots, s_{i-1})$  and is counted by  $LP_{i-1}(s_1 - 1, s_2 - 1, \dots, s_{i-1} - 1)$ . The next step on the path must be an N-step from  $(s_i, i - 1)$  to  $(s_i, i)$ . The part from  $(s_i, i)$  to  $(x, r)$  is bounded by  $(s_{i+1}, \dots, s_r)$ , and is counted by  $LP_{r-i}(s_{i+1} - s_i, s_{i+2} - s_i, \dots, s_r - s_i)$ . Note that  $s_j - s_{i+1} = a_{i+1} + \dots + a_j$ , we obtain the recurrence (4).  $\square$

Theorem 3.2 follows from the recurrences (3) and (4) and the inductive hypothesis.  $\square$

One advantage of equating a new combinatorial structure to lattice paths is that lattice paths are a classical subject in combinatorial theory, whose enumeration can be systematically approached by symbolic computation and umbral calculus, see, for example, see the classical book by Mohanty [6] and the extensive work done by H. Niederhausen and his collaborators [7, 8, 9, 10, 11, 2, 3]. In particular, it is known that lattice paths within general boundaries can be computed by a determinant formula. See Theorem 1 of [6, p.32]. Applying Theorem 3.2, we obtain the following determinant formula for the value of  $\phi^*(a_1, \dots, a_r, a_{r+1})$ .

**Corollary 3.7.** Let  $1 \leq a_1 \leq a_2 \leq \dots \leq a_r \leq a_{r+1}$  be a non-decreasing integer sequence of length  $r + 1$ . Then

$$\phi^*(a_1, a_2, \dots, a_{r+1}) = \det M,$$

where  $M$  is a  $r \times r$  matrix whose entries are given by  $M_{i,j} = \binom{s_i+1}{j-i+1}$ .

We check some easy cases.

1. For  $r = 1$ ,  $s_1 = a_1$ , and any  $a_2 \geq a_1$ , we have  $\phi^*(a_1, a_2) = \det_{1 \times 1}(a_1 + 1) = a_1 + 1$ .
2. For  $r = 2$ ,  $s_1 = a_1$  and  $s_2 = a_1 + a_2$ , for any  $a_1 \leq a_2 \leq a_3$  we have

$$\begin{aligned} \phi^*(a_1, a_2, a_3) &= \det \begin{pmatrix} s_1 + 1 & \binom{s_1+1}{2} \\ 1 & s_2 + 1 \end{pmatrix} \\ &= (s_1 + 1)(s_2 + 1 - \frac{s_1}{2}) \\ &= (a_1 + 1)(a_2 + 1 + \frac{a_1}{2}). \end{aligned}$$

3. For the regular binary strings  $S = (1^k 0^k)^n$ ,

$$\phi^*(k, k, \dots, k) = LP_{n-1}(k, 2k, \dots, (n-1)k)$$

is the number of lattice paths from  $(0, 0)$  to  $(x, n-1)$  (where  $x > (n-1)k$ ) that never touches the line  $x = ky + k + 1$ . Using the formula (1.11) of [6, p.9], we have

$$\begin{aligned} |NC_2((1^k 0^k)^n)| &= \frac{k+1}{k+1+(k+1)(n-1)} \binom{k+1+(k+1)(n-1)}{n-1} \\ &= \frac{1}{(k+1)n+1} \binom{(k+1)n+1}{n}, \end{aligned}$$

which is the  $n$ -th  $(k+1)$ -Catalan number.

Another approach in lattice path counting is to use the analytical theory of biorthogonal polynomials. In [5] the second author, together with Kung and Sun, proved that  $LP_r(s_1, \dots, s_r)$  can be represented by differential Goncarov polynomials, which is a sequence of polynomials that are bi-orthogonal to a given sequence of linear operators, each of which is a formal power series of the (backward) difference operator. Explicitly, let  $\tilde{g}_r(x; b_1, \dots, b_r)$  be the difference Goncarov polynomials with the variable  $x$  and parameters  $b_1, \dots, b_r$ . Then Theorem 4.1 of [5] states that

$$LP_r(s_1, \dots, s_r) = \frac{1}{r!} \tilde{g}_r(x; x - s_1 - 1, x - s_2 - 1, \dots, x - s_r - 1), \forall x \in \mathbb{R}.$$

In particular, letting  $x = 1$ , we derive the following equation.

**Corollary 3.8.** *Let  $a_1 \leq a_2 \leq \dots \leq a_{r+1}$  and  $s_i = a_1 + \dots + a_i$ . Then*

$$\phi^*(a_1, \dots, a_r, a_{r+1}) = \frac{1}{r!} \tilde{g}_r(1; -s_1, -s_2, \dots, -s_r). \quad (5)$$

Goncarov polynomials bear many nice analytical and combinatorial properties, including the expansion formula, a linear recurrence, the Appell relations, difference relations, summation formula, shift invariant formula, and binomial expansion. See [5] for an expository article on the theory of Goncarov polynomials and its application in combinatorics. Our result in the present paper allows one to apply these powerful mechanisms to the enumeration of non-crossing pairings of well-balanced binary strings. As examples, we list the linear recurrence and the Appell relation of differential Goncarov polynomials, which give a linear recurrence and a shifted generating function for  $\phi^*$ . In fact, many of the combinatorial properties of  $\phi^*$  discussed in [4] can be deduced from the known results of lattice paths and Goncarov polynomials.

- Linear Recurrence for Differential Goncarov Polynomials.

$$x^{(n)} = \sum_{i=0}^n \binom{n}{i} b_i^{(n-i)} \tilde{g}_i(x; b_0, b_1, \dots, b_{i-1}) \quad (6)$$

where  $x^{(n)}$  is the rising factorial  $x^{(n)} = x(x+1) \cdots (x+n-1)$ .

- Appell relation for Differential Goncarov Polynomials.

$$(1-t)^{-x} = \sum_{n=0}^{\infty} \tilde{g}_n(x; b_0, \dots, b_{n-1}) \frac{t^n}{n!(1-t)^{b_n}}. \quad (7)$$

Letting  $x = 1$  we have the following results for  $\phi^*(a_1, \dots, a_{r+1})$  with  $a_1 \leq a_2 \leq \dots \leq a_{r+1}$ .

**Corollary 3.9.** *Linear recurrence for  $\phi^*$*

$$1 = \sum_{i=0}^n (-1)^{n-i} \frac{(s_i)_n}{(n-i)!} \phi^*(a_1, a_2, \dots, a_{i+1}),$$

where  $(x)_n$  is the falling factorial  $(x)_n = x(x-1)\cdots(x-n+1)$ .

**Corollary 3.10.** *Appell relation for  $\phi^*$ .*

$$\sum_{i=0}^{\infty} t^i (1-t)^{s_{i+1}} \phi^*(a_1, \dots, a_{i+1}) = \frac{1}{1-t},$$

where  $s_i = a_1 + \dots + a_i$ .

We remark that Theorem 3.2 is only proved for non-decreasing sequences  $a_1 \leq a_2 \leq \dots \leq a_{r+1}$ . It is easy to check that in general,

$$\phi^*(a_1, \dots, a_i, a_{i+1}, \dots, a_r) \neq \phi^*(a_1, \dots, a_{i+1}, a_i, \dots, a_r).$$

Finding explicit formulas for  $\phi^*(a_1, \dots, a_r)$  for general integer sequences is still an open problem.

## 4 A bijective proof for Theorem 3.2

Inspired by the recurrence proof in Section 3, we construct a bijection between the set of non-crossing pairings of a well-balanced binary string to the set of lattice paths with appropriate right boundary. The bijection is established via an intermediate object—the ordered left-weighted ballot sequences, or the ordered left-weighted parenthesization, which are special cases of the left-weighted Catalan extensions studied and cataloged by the first author in [12].

**Definition 4.1.** *Given a sequence of positive integers  $\mathbf{b} = (b_1, b_2, \dots, b_k)$  such that  $\sum_{i=1}^k b_i = n$ , an ordered ballot sequence of left-weight  $\mathbf{b}$  is an arrangement of  $b_1, b_2, \dots, b_k$  and  $n$  copies of  $-1$  such that (i) the terms  $b_1, \dots, b_k$  appear in order, and (ii) the partial sums are always non-negative.*

For example, the ordered ballot sequences of left-weight  $(2, 1)$  are listed below.

$$21 - - - \quad 2 - 1 - - \quad 2 - - 1 -$$

An equivalent description is the ordered left-weighted parenthesization, where each integer  $b_i$  represents a left parenthesis of weight  $b_i$ , (or an inseparable block of  $b_i$  many left-parentheses), and each  $-1$  represents one right-parenthesis. For each  $b_i$ , we draw a line from  $b_i$  to the last right-parenthesis that matches with it, and refer to it as the *region* of the left-parenthesis  $b_i$ . See the following figure for an illustration, where we also indicate which right parentheses are matched with  $b_i$ .

$$\begin{array}{ccc} 2 & 1 & ) & ) & ) \\ \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} \end{array} \quad \begin{array}{ccc} 2 & ) & 1 & ) & ) \\ \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} \end{array} \quad \begin{array}{ccc} 2 & ) & ) & 1 & ) \\ \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} & \underbrace{\quad} \end{array}$$

Figure 4: Ordered left-weighted parentheses and the corresponding regions

Let  $a_1 \leq a_2 \leq \dots \leq a_r, a_{r+1}$  be a sequence of positive integers, and let  $s_i = a_1 + \dots + a_i$  for  $i = 1, 2, \dots, r$ . We will use  $LwP(a_1, \dots, a_r)$  to denote the set of ordered left-weighted parentheses with the weight  $(a_1, \dots, a_r)$ . Let  $S = 1^{a_1}0^{a_1} \dots 1^{a_{r+1}}0^{a_{r+1}}$ . We shall construct two bijections, bijection  $f$  from  $Path_r(s_1, \dots, s_r)$  to  $LwP(a_1, \dots, a_r, a_{r+1})$  and bijection  $g$  from  $LwP(a_1, \dots, a_{r+1})$  to  $NC_2(S)$ . We will then combine  $f$  and  $g$  to create a bijection from  $Path_r(s_1, \dots, s_r)$  to  $NC_2(S)$ .

**Bijection  $f$  from  $Path_r(s_1, \dots, s_r)$  to  $LwP(a_1, \dots, a_r, a_{r+1})$ .**

For any path  $P$  from  $(0, 0)$  to  $(s_r, r)$  with the right boundary  $s_1, \dots, s_r$ , write the path  $P$  as a sequence of  $E$  and  $N$ s. There are  $r$   $N$ -steps and  $s_r$  many  $E$  steps.

1. Add an  $N$  at the beginning of  $P$ , and  $a_{r+1}$  many  $E$  at the end of  $P$ .
2. Replace the  $i$ th  $N$  with a left-parenthesis of weight  $a_i$ .
3. Replace each  $E$  with a right parenthesis.

For example, let the positive integers  $a_i$  be  $2, 2, 3, 4$ , then the right boundary is  $\mathbf{s} = (2, 4, 7)$ . For a path  $P$  coded by  $(1, 1, 2)$ , the steps in  $P$  are  $ENNENEEEE$ , then  $f(P) = 2)234))))))$ .

It is easy to see that  $f(P) \in LwP(a_1, \dots, a_r, a_{r+1})$  whenever  $P \in Path_r(s_1, \dots, s_r)$ . The steps are easy to reverse, so  $f$  is a bijection.

**Bijection  $g$  from  $LwP(a_1, \dots, a_{r+1})$  to  $NC_2(S)$ , where  $S = 1^{a_1}0^{a_1} \dots 1^{a_{r+1}}0^{a_{r+1}}$  and  $a_1 \leq a_2 \leq \dots \leq a_{r+1}$ .**

The map  $g$  is defined inductively. First set  $g(\emptyset) = \emptyset$ . For  $r = 0$ , the unique element of  $LwP(a)$  is mapped to the unique non-crossing pairing of  $1^a0^a$ .

Assume that  $g$  is a well-defined bijection from the set  $LwP(a_1, \dots, a_l)$  to the set  $NC_2(1^{a_1}0^{a_1} \dots 1^{a_l}0^{a_l})$  for all  $l \leq r$  and  $a_1 \leq a_2 \leq \dots \leq a_l$ . We shall define  $g$  on  $LwP(a_1, \dots, a_{r+1})$ .



- ference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 2000), vol. 143, 2000, pp. 141–160.
- [3] K. Humphreys and H. Niederhausen, *Counting lattice paths taking steps in infinitely many directions under special access restrictions*, Theoret. Comput. Sci. **319** (2004), no. 1-3, 385–409.
  - [4] T. Kemp, K. Mahlborg, A. Rattan, and C. Smyth, *Enumeration of non-crossing pairings on bit strings*. J. Combin. Theory Ser. A **118** (2011), no. 1, 129–151.
  - [5] J. Kung, X. Sun, and C. H. Yan, *Goncarov-type polynomials and applications in combinatorics*, Preprint (2005).
  - [6] S. G. Mohanty, *Lattice path counting and applications*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1979, Probability and Mathematical Statistics.
  - [7] H. Niederhausen, *Sheffer polynomials for computing exact Kolmogorov-Smirnov and Rényi type distributions*, Ann. Statist. **9** (1981), no. 5, 923–944.
  - [8] H. Niederhausen, *The enumeration of restricted random walks by Sheffer polynomials with applications to statistics*, J. Statist. Plann. Inference **14** (1986), no. 1, 95–114.
  - [9] H. Niederhausen, *Symmetric Sheffer sequences and their applications to lattice path counting*, J. Statist. Plann. Inference **54** (1996), no. 1, 87–100.
  - [10] H. Niederhausen, *Lattice paths between diagonal boundaries*, Electron. J. Combin. **5** (1998), Research Paper 30, 37 pp. (electronic).
  - [11] H. Niederhausen, *Recursive initial value problems for Sheffer sequences*, Discrete Math. **204** (1999), no. 1-3, 319–327.
  - [12] P. R. F. Schumacher, *A left weighted Catalan extension*, Integers **10** (2010), 771–792.
  - [13] R. P. Stanley, *Exercises on Catalan and related numbers*, <http://math.mit.edu/~rstan/ec/catalan.pdf>.