



Runs and RSK Tableaux of Boolean Permutations

Emily Gunawan^{*}, Jianping Pan, Heather M. Russell
and Bridget Eileen Tenner[†]

Abstract. We define and construct the “canonical reduced word” of a boolean permutation, and show that the RSK tableaux for that permutation can be read off directly from this reduced word. We also describe those tableaux that can correspond to boolean permutations, and enumerate them. In addition, we generalize a result of Mazorchuk and Tenner, showing that the “run” statistic influences the shape of the RSK tableau of arbitrary permutations, not just of those that are boolean.

Mathematics Subject Classification. Primary 05A05; Secondary 20F55, 06A07, 05A19.

Keywords. Boolean permutation, Robinson–Schensted–Knuth correspondence, Permutation pattern, Reduced word, Run.

1. Introduction

In 1961, Schensted showed that the first part of an RSK partition $(\lambda_1(w), \lambda_2(w), \dots)$ is equal to the length of a longest increasing subsequence in a permutation [15]. In 1974, Greene generalized Schensted’s result, showing that the RSK partition of a permutation records the numbers of disjoint unions of increasing sequences of the permutation [8]. As he pointed out in his paper, this result was “somewhat surprising,” since there was no concrete interpretation of each individual part of the RSK partition below the first part.

The boolean elements in a Coxeter group are those elements whose principal order ideals in the Bruhat order are isomorphic to boolean algebras. This is

^{*}Research partially completed at the Isaac Newton Institute for Mathematical Sciences during the program Cluster algebras and representation theory (supported by EPSRC Grant Number EP/R014604/1). [†]Research partially supported by NSF Grant DMS-2054436 and Simons Foundation Collaboration Grant for Mathematicians 277603.

an important class of elements, established in [20], which has beautiful topological ([7, 9, 12, 13]) and representation theoretic ([10]) properties. Previous work established that the RSK shape of a boolean permutation has at most two rows. Recently, Mazorchuk and Tenner [10, Theorem 6.4] showed that the size of the RSK partition of a boolean permutation excluding the first row equals its run statistic, a statistic on the reduced words of a permutation. In Theorem 3.9 of the present paper, we generalize this result to all permutations, as follows.

Theorem

For any permutation $w \in S_n$, we have

$$\lambda_1(w) + \text{run}(w) = n.$$

Our result gives a concrete interpretation to the sum of all individual parts below the first row; in the case of fully commutative permutations, this gives meaning to the length of the second row and, by doing so, takes a step to address the missing meaning mentioned in [8].

Theorem 3.9 demonstrates a connection between the Coxeter-perspective and the pattern-perspective of permutations, via the RSK correspondence. From there, we focus exclusively on boolean permutations, defining a “canonical reduced word,” which we can construct in two (equivalent) ways. Not only does this word demonstrate the necessary run statistic, but we show, in fact, that it directly determines the entire RSK tableaux of the boolean permutation—without using the insertion algorithm.

The relationship between this canonical reduced word and the tableaux can be exploited further, allowing us to characterize precisely which tableaux are the RSK insertion and recording tableaux of boolean permutations. In particular, only those tableaux that we call “uncrowded” can correspond to these permutations under RSK. The uncrowded tableaux, themselves, have interesting combinatorics, as we demonstrate via their enumeration.

The paper is organized as follows. In Sect. 2, we introduce notation and terminology that we will use throughout the paper. We also review heap posets of boolean permutations, RSK insertion, and runs. In Sect. 3, we give a concrete interpretation of the run statistic in terms of the shape of arbitrary permutations in Theorem 3.9. Section 4 introduces the canonical reduced word of a boolean permutation, which can be defined (and constructed) from either an arbitrary reduced word or from its heap. The canonical word is used to construct the corresponding insertion and recording tableaux in Theorem 4.5, and it demonstrates the run statistic (Corollary 4.7). In Sect. 5, we characterize tableaux that are RSK tableaux of boolean permutations, which we call “uncrowded” tableaux. This is done in Corollaries 5.4 and 5.6. We conclude with Sect. 6, enumerating the uncrowded tableaux via a bijection with binary words in which each maximal block of 1s has odd length.

2. Background and Notation

Let S_n be the symmetric group on n elements. We represent permutations of S_n in *one-line notation* as $w = w(1)w(2) \cdots w(n)$. For each $i \in \{1, \dots, n-1\}$, let $s_i \in S_n$ denote the *simple reflection* (also called an *adjacent transposition*) swapping i and $i+1$, and fixing all other letters. The simple reflections generate S_n , meaning that every $w \in S_n$ can be decomposed as a product $w = s_{i_1} \cdots s_{i_\ell}$. The minimum ℓ among all such decompositions for w is the (*Coxeter*) *length* of w , denoted $\ell(w)$. In our proofs, we will make use of the fact that $\ell(w)$ is also the number of *inversions* in the one-line notation for w where an inversion is a pair of positions $i < j$, such that $w(i) > w(j)$. An expression $w = s_{i_1} \cdots s_{i_{\ell(w)}}$ is called a *reduced decomposition* of w . We ease this notation by writing such a decomposition as the *reduced word* $[i_1 \cdots i_{\ell(w)}]$. Let $R(w)$ denote the set of reduced words for w .

The *support* $\text{supp}(w)$ of a permutation w is the set of letters appearing in reduced words of w . Although simple reflections are subject to the Coxeter relations, this does not change the set of reflections appearing in any reduced decomposition, so $\text{supp}(w)$ is well defined. A permutation $w \in S_n$ has *full support* if $\text{supp}(w) = \{1, \dots, n-1\}$. For example, let $w = 51342 = s_4 s_2 s_3 s_2 s_4 s_1 \in S_5$. The permutation w has six inversions, so $\ell(w) = 6$ and $[423241] \in R(w)$. Since $\text{supp}(w) = \{1, 2, 3, 4\}$, we conclude that w has full support.

2.1. Fully Commutative Permutations and Boolean Permutations

Let $w \in S_n$ and $\sigma \in S_m$ with $m \leq n$. The permutation w is said to *contain the pattern* σ if w has a (not necessarily contiguous) subsequence whose elements are in the same relative order as σ . If w does not contain σ , we say that w *avoids* σ . For example, $w = 314592687$ contains the pattern 1423, because the subsequence 4968 (among others) is ordered in the same way as 1423. On the other hand, w avoids 3241, since it has no subsequence ordered in the same way as 3241. Note also that each inversion of a permutation is an instance of a 21-pattern.

Simple reflections satisfy *commutation relations* of the form $s_i s_j = s_j s_i$ when $|i - j| > 1$. An application of a commutation relation is called a *commutation move*. When referring to reduced words, we will say adjacent letters i and j in a reduced word *commute* when $|i - j| > 1$. Given a reduced word $[s]$ of a permutation, the equivalence class consisting of all words that can be obtained from $[s]$ by a sequence of commutation moves is the *commutation class* of $[s]$. A permutation whose set of reduced words forms a single commutation class is called *fully commutative*. The following proposition characterizes fully commutative permutations in terms of pattern avoidance.

Proposition 2.1 ([4]). *Let w be a permutation. The following are equivalent:*

- w is fully commutative,
- w avoids the pattern 321,
- no reduced word of w contains $i(i+1)i$ as a factor, for any i , and
- no reduced word of w contains $(i+1)i(i+1)$ as a factor, for any i .

This paper focuses on the subset of fully commutative permutations known as boolean permutations. While boolean permutations can be characterized using the language of the Bruhat order (see, for example, [3, Chapter 2]), the following result provides a description analogous to that of Proposition 2.1.

Proposition 2.2 ([20]). *Let w be a permutation. The following are equivalent:*

- w is boolean;
- w avoids the pattern 321 and 3412;
- there is a reduced word of w that consists of all distinct letters;
- every reduced word of w consists of all distinct letters.

2.2. Heaps of a Boolean Permutation

Heaps are posets used in [19] to study fully commutative elements of Coxeter groups. In this paper, heaps for boolean permutations provide a useful visualization of a key construction in Sect. 4.1. For a detailed list of attributions on the theory of heaps, see [18, solutions to Exercise 3.123(ab)].

By Proposition 2.2, a boolean permutation w has the property that each letter in $\text{supp}(w)$ appears exactly once in every reduced word of w . Thus, the relative positions of every pair of consecutive letters i and $i + 1$ in reduced words of w are fixed. This allows one to give the following simple description of the heap of a boolean permutation.

Definition 2.3. Given a boolean permutation $w \in S_n$, the heap H_w of w is the partial order on $\text{supp}(w)$ obtained via the transitive closure of the cover relations

$$i \prec i + 1 \text{ if } i \text{ appears to the left of } i + 1 \text{ in every reduced word of } w$$

and

$$i \succ i + 1 \text{ otherwise.}$$

We will refer to the Hasse diagram of a heap as a *heap diagram*.

The following proposition explains the connection between the heap of a boolean permutation and its complete set of reduced words.

Proposition 2.4 ([19, proof of Proposition 2.2] and [18, solutions to Exercise 3.123(ab)]). *If w is a boolean permutation, then the set of linear extensions of the heap H_w is the set of reduced words of w .*

We conclude this subsection with an example illustrating the heap of a boolean permutation.

Example 2.5. The permutation $w = 314569278 \in S_9$ has [21873456] as a reduced word, and the permutation is, therefore, boolean. The heap diagram for w is depicted in Fig. 1. By Proposition 2.4, the elements of $R(w)$ are exactly the linear extensions of this heap. For instance, [87213456] $\in R(w)$.

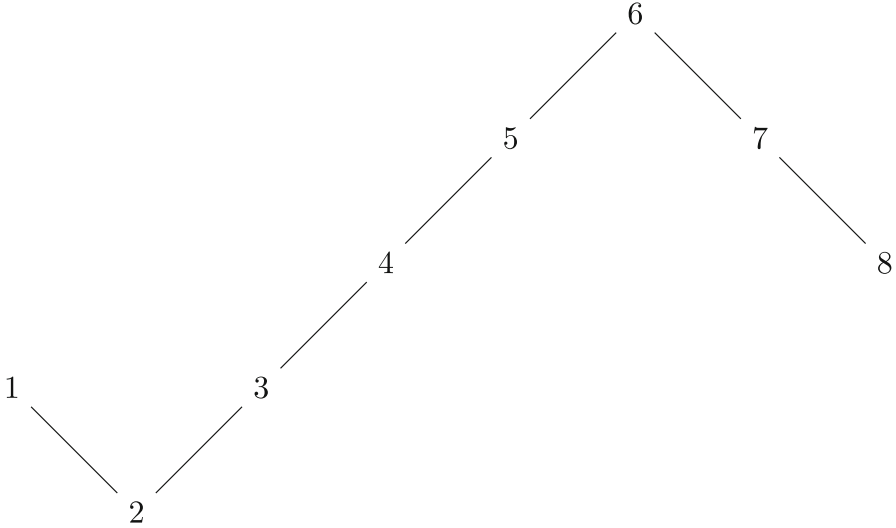


FIGURE 1. The heap diagram for the boolean permutation 314569278

2.3. Robinson–Schensted–Knuth Tableaux

Given a partition $\lambda = (\lambda_1, \lambda_2, \dots) \vdash n$, the *Young diagram of shape λ* is a top-and left-justified collection of n boxes, such that the i^{th} row has λ_i boxes. A *standard Young tableau of shape λ* is a filling of the Young diagram of shape λ by the values $1, \dots, n$, such that each value appears exactly once and values increase from left to right in rows and from top to bottom in columns.

The Robinson–Schensted–Knuth (RSK) correspondence as described in [15] is a bijection

$$w \mapsto (P(w), Q(w))$$

from S_n onto pairs of standard Young tableaux of size n having identical shape. The tableau $P(w)$ is called the *insertion tableau* of w , and the tableau $Q(w)$ is the *recording tableau* of w . The shape of these tableaux is called the *RSK partition of w* , denoted $\text{sh}(w) = (\lambda_1(w), \lambda_2(w), \dots)$. We will also write $P_i(w)$ to denote the partial insertion tableau constructed by the first i letters, $w(1) \cdots w(i)$, in the one-line notation for w . For more details, including a precise description of the RSK insertion algorithm, see, for example [17, Section 7.11].

The following symmetry result is a feature of the RSK insertion algorithm, and one that will simplify our own work.

Proposition 2.6 ([16]). *For any permutation w*

$$P(w^{-1}) = Q(w).$$

Schensted’s theorem [15, Theorem 1], stated below, articulates an important relationship between the RSK partition shape and the one-line notation for w .

Theorem 2.7. *Let w be a permutation with RSK partition shape $\text{sh}(w) = (\lambda_1(w), \lambda_2(w), \dots)$, and let $(\mu_1(w), \mu_2(w), \dots)$ denote the conjugate of $\text{sh}(w)$. The length of a longest increasing (resp., decreasing) subsequence in the one-line notation of w is $\lambda_1(w)$ (resp., $\mu_1(w)$).*

By Proposition 2.1, fully commutative permutations are 321-avoiding and therefore have decreasing subsequences of length at most two. It follows from Theorem 2.7, then, that the RSK partitions for fully commutative—and, in particular, for boolean—permutations have at most two rows. That observation is key to the arguments in this paper.

Theorem 2.7 is sufficient for the purposes of this paper, but we point out that Greene’s theorem [8, Theorem 3.1] is an important generalization of that result, and could perhaps be useful in extensions of our work. For more details about Greene’s theorem, see, for example, [14, Chapter 3].

2.4. Runs and Longest Increasing Subsequences

In this paper, we define a *run* as an increasing or decreasing sequence of consecutive integers; for example, 234 and 432 are runs, but 245 and 542 are not. Given a permutation w , let $\text{run}(w)$ denote the fewest number of runs needed to form a reduced word for w . A reduced word $[s] \in R(w)$ that can be written as the concatenation of $\text{run}(w)$ runs is called an *optimal run word* for w .

Example 2.8. Consider the permutation $w = 345619278 \in S_9$. Examining all reduced words for w shows that $\text{run}(w) = 3$, and thus the reduced words [21873456] and [87213456] given in Example 2.5 are both optimal run words for w . We can highlight their runs by writing them as $[21 \cdot 87 \cdot 3456]$ and $[87 \cdot 21 \cdot 3456]$, respectively. In contrast, $[82713456] \in R(w)$ is not optimal, because the string 82713456 cannot be written as the concatenation of three runs.

Recently, Mazorchuk and Tenner [10, Theorem 6.4] showed that, if w is a boolean permutation, then $\lambda_2(w) = \text{run}(w)$. Because the RSK partitions of boolean permutations have at most two rows, we can apply Theorem 2.7 to conclude that the length of a longest increasing subsequence of a boolean permutation $w \in S_n$ is equal to $n - \text{run}(w)$. In the present work, we will see that this result is not only true for boolean permutations, but for all permutations.

3. Runs and RSK Partitions

The goal of this section is to prove Theorem 3.9, which relates the first row of the RSK partition to the run statistic, for all permutations. This will generalize [10, Theorem 6.4], which was a result for boolean permutations, and the argument from that paper will guide this more general setting.

Proposition 3.1 ([10, Lemma 6.2 and Corollary 6.3]). *For any permutation $w \in S_n$, we have $n - \lambda_1(w) \leq \text{run}(w)$.*

Proof. Although [10] states these results in terms of boolean permutations, the same proofs work for all permutations. $\square$

It remains to show the other direction of the inequality, which we will do in Proposition 3.8. To that end, we will define a function ρ , which maps a permutation w to a shorter permutation (“shorter” in terms of Coxeter length). It multiplies w by a single run **on the left or right**, in such a way that a longest increasing subsequence in $\rho(w)$ is longer than that of w .

Definition 3.2. We define a map

$$\rho : (S_n \setminus \{12 \cdots n\}) \rightarrow S_n$$

as follows. Fix a permutation $w \in S_n$ that is not the identity permutation, and consider the lexicographically least longest increasing subsequence in the one-line notation for w . (In fact, any longest increasing subsequence would satisfy our needs.) Let $q \in [1, n]$ be the smallest value not appearing in this subsequence.

- If $q = 1$, then set $t := w^{-1}(1)$. Note that, by definition of q , we must have $t > 1$. Define r to be the run $(t - 1) \cdots 321$ and set $\rho(w) := w[r]$. In other words, $\rho(w)$ uses the run to slide 1 into the leftmost position of the permutation.
- Suppose that $q > 1$. If q appears to the right of $q - 1$ in the one-line notation for w , then set $t := w^{-1}(q)$ and $t' := w^{-1}(q - 1)$, so $t > t'$. In addition, $t > t' + 1$, because otherwise adding q to our lexicographically least longest increasing subsequence (which includes $q - 1$, by definition of q) would have made a longer increasing subsequence. Define r to be the (decreasing) run $(t - 1) \cdots (t' + 1)$ and set $\rho(w) := w[r]$. That is, $\rho(w)$ uses the run to slide q into position $t' + 1$, the position immediately to the right of $q - 1$.
- If $q > 1$ and q appears to the left of $q - 1$ in the one-line notation for w , then let $j \in [1, q - 1]$ be the smallest value appearing to the right of q in the one-line notation of w . Define r to be the (increasing) run $j(j + 1) \cdots (q - 1)$ and set $\rho(w) := [r]w$. In other words, there is a (possibly nonconsecutive) subsequence

$$q \qquad j \qquad (j + 1) \quad (j + 2) \quad \cdots \quad (q - 2) \quad (q - 1)$$

in w , and $\rho(w)$ transforms that subsequence into

$$j \qquad (j + 1) \quad (j + 2) \quad \cdots \quad (q - 2) \quad (q - 1) \qquad q$$

To make the process of Definition 3.2 more concrete, we consider a few examples. We will continue example (c) after a trio of lemmas, demonstrating those results as well.

Example 3.3. (a) Let $u = 342516$. The lexicographically least longest increasing subsequence in u is 3456. The smallest value not appearing in 3456 is $q = 1$, so we use the first scenario of Definition 3.2 to compute $\rho(u)$. Set $t := u^{-1}(q) = 5$, producing the decreasing run $r = 4321$. We set $\rho(u) = u[r]$. Multiplying u by $[r]$ on the right is equivalent to

sliding $u(5) = 1$ into the leftmost position of the one-line notation, so $\rho(u) = 134256$.

- (b) Let $v = 142563$. The lexicographically least longest increasing subsequence in v is 1256. The smallest value not appearing in 1256 is $q = 3$. Since 3 appears to the right of $3 - 1$, we use the second scenario of Definition 3.2 to compute $\rho(v)$. We set $t := v^{-1}(q) = v^{-1}(3) = 6$ and $t' := v^{-1}(q - 1) = v^{-1}(2) = 3$, producing the decreasing run $r = 54$. We set $\rho(v) = v[r]$. Multiplying $v = 142563$ by $[r]$ on the right will slide $v(t) = v(6) = q = 3$ into position $t' + 1 = 4$, the position immediately to the right of $q - 1 = 2$, so $\rho(v) = 142356$.
- (c) Consider the permutation $w = 51642738$. The lexicographically least longest increasing subsequence in w is 1238. The smallest value not appearing in this longest subsequence is $q = 4$. Since 4 appears to the left of $4 - 1$ in w , we use the third scenario of Definition 3.2 to compute $\rho(w)$, producing $j = 2$ and the increasing run $r = 23$. Then multiplying w by $[r]$ on the left transforms the subsequence 423 of w into 234: $\rho(w) = [r]w = 51623748$.

There are important features of the image of a permutation under the map ρ , and we will want to take advantage of these later. To prepare for that, we now identify three of the map's key qualities.

Lemma 3.4. *For any nonidentity permutation w*

$$\ell(\rho(w)) + \ell([r]) = \ell(w),$$

where r is the run described in Definition 3.2. In particular, $\ell(\rho(w)) < \ell(w)$.

Proof. The minimality of q and the definition(s) of the run r in Definition 3.2 ensure that each simple reflection described by the letters in r undoes an inversion of the permutation. $\square$

While an application of the map ρ decreases the length statistic, it increases a different permutation statistic.

Lemma 3.5. *For any nonidentity permutation w , the length of a longest increasing subsequence in $\rho(w)$ is greater than the length of a longest increasing subsequence in w .*

Proof. The run r in Definition 3.2 was constructed so that the effect of multiplying w by $[r]$ is to insert q into the lexicographically least longest increasing subsequence in the permutation, without changing the relative order of any other letters in that subsequence. $\square$

After establishing Theorem 3.9, we shall see ρ increases the length of a longest increasing subsequence by 1.

Recall that Schensted's theorem relates those longest increasing subsequences to the size of the top row in the permutation's shape under the Robinson–Schensted–Knuth correspondence. This means that Lemma 3.5 can be written as

$$\lambda_1(\rho(w)) > \lambda_1(w).$$

Finally, because $\rho(w)$ and w differ by product with a run, we can say something about the relationship between $\text{run}(w)$ and $\text{run}(\rho(w))$.

Lemma 3.6. *For any nonidentity permutation w*

$$\text{run}(w) \leq \text{run}(\rho(w)) + 1.$$

Proof. Since $\rho(w)$ is equal to $w[r]$ or $[r]w$ by Definition 3.2, we can write w as the product of an optimal run word for $\rho(w)$ and the inverse permutation $([r])^{-1}$. If we write the run r as $r = r_1 \cdots r_h$, then $r_h \cdots r_1$ is also a run and it is a (in fact, the) reduced word for this $([r])^{-1}$. Thus, it is possible to write w as a product of $\text{run}(\rho(w)) + 1$ runs. This guarantees that an optimal run word for w has at most $\text{run}(\rho(w)) + 1$ runs, and so, $\text{run}(w) \leq \text{run}(\rho(w)) + 1$. $\square$

We demonstrate the preceding three lemmas using the third permutation from Example 3.3.

Example 3.7. Consider the permutation $w = 51642738$. As computed above, $\rho(w) = 51623748$.

- (a) The 21-patterns 42 and 43 appear in w but not in $\rho(w)$. All other inversions of w are also inversions of $\rho(w)$, so $\ell(\rho(w)) = \ell(w) - \ell([r]) = 10 - 2 = 8$, illustrating Lemma 3.4.
- (b) In the construction of $\rho(w)$, the value $q = 4$ was inserted into the increasing subsequence 1238 to form 12348. This is an increasing subsequence of $\rho(w)$ that is longer than any increasing subsequence found in w , illustrating Lemma 3.5.
- (c) We can compute $\text{run}(w) = 4$. For example, three of the optimal run words for w are

$$[32 \cdot 456 \cdot 321 \cdot 43], [32 \cdot 4321 \cdot 56 \cdot 43], \text{ and } [32 \cdot 4321 \cdot 543 \cdot 6].$$

We highlight these particular optimal run words, because their leftmost letters are “32,” which will be canceled after multiplication by $[r]$. That is, we find three corresponding reduced words of $\rho(w)$

$$[456 \cdot 321 \cdot 43], [4321 \cdot 56 \cdot 43], \text{ and } [4321 \cdot 543 \cdot 6].$$

Therefore, $\text{run}(\rho(w)) \leq 3 = \text{run}(w) - 1$, illustrating Lemma 3.6.

Lemmas 3.4, 3.5, and 3.6 give us tools for making inductive arguments involving the permutation statistics length, length of a longest increasing subsequence, and runs. In other words, these lemmas allow us to make inductive arguments toward achieving an upper bound for the function $\text{run}(w)$. Such a bound, in turn, could be combined with the lower bound proved in Proposition 3.1.

Proposition 3.8. *For any permutation $w \in S_n$, we have $\text{run}(w) \leq n - \lambda_1(w)$.*

Proof. If w is the identity, then $\lambda_1 = n$ and $\text{run}(w) = 0$, and the result is trivially true. Suppose that w is not the identity permutation. Assume, inductively, that the result holds for all permutations shorter than w . In particular, since Lemma 3.4 tells us that $\rho(w)$ is shorter than w , we have

$$\text{run}(\rho(w)) \leq n - \lambda_1(\rho(w)).$$

Next, note that, because λ_1 takes only integer values, we can rewrite Lemma 3.5 as

$$\lambda_1(w) \leq \lambda_1(\rho(w)) - 1.$$

Combining this with Lemma 3.6 gives

$$\begin{aligned} \text{run}(w) &\leq \text{run}(\rho(w)) + 1 \\ &\leq n - \lambda_1(\rho(w)) + 1 \\ &= n - (\lambda_1(\rho(w)) - 1) \\ &\leq n - \lambda_1(w). \end{aligned}$$

□

We can now prove the precise relationship between λ_1 and run , for any permutation.

Theorem 3.9. *For any permutation $w \in S_n$*

$$\lambda_1(w) + \text{run}(w) = n.$$

Proof. This follows from Propositions 3.1 and 3.8. □

Intuitively speaking, multiplying a permutation on the right by a run is the same as deleting an entry from the permutation and then inserting it somewhere else. For example, let $w = 253146 \in S_6$. If we multiply w on the right by the run $[321]$, we get $w[321] = 125346$. We can see that 1 is deleted from w and then inserted to the beginning. In fact, $\text{run}(w)$ is the minimum number of runs whose concatenation is a decomposition for w , where we drop the reduced condition.

Lemma 3.10. *The minimum number of runs whose concatenation is a decomposition for w is $\text{run}(w)$.*

Proof. Clearly, the minimum number of runs whose concatenation is a decomposition for w is at most $\text{run}(w)$. We prove the other inequality. Let $w \in S_n$ and let $w = \rho_1 \rho_2 \dots \rho_\ell$ where ρ_i is a run for $1 \leq i \leq \ell$. Since a run deletes one entry and inserts it somewhere else, it can change the length of a longest increasing subsequence by at most one. That is, the length of a longest increasing subsequence of the permutation $\rho_1 \rho_2 \dots \rho_\ell$ is at least $n - \ell$. Therefore, $\lambda_1(w)$ is at least $n - \ell$. It follows that $\ell \geq n - \lambda_1(w) = \text{run}(w)$. □

With the equivalent definition of $\text{run}(w)$ as being the minimum number of runs whose concatenation is a decomposition for w , our statistic $\text{run}(w)$ recovers the “Ulam distance” on permutations. Ulam’s metric was originally defined to study mutation of DNA sequences from the perspective of permutations [21] as well as to find the fastest way to sort a bridge hand of 13 cards [1]. An *Ulam move* deletes a value from the current permutation and places it at some other position. Correspondingly, the *Ulam distance* $U(\sigma, \tau)$ is the minimum number of Ulam moves needed to obtain τ from σ . See also [2]. In fact, our Theorem 3.9 is equivalent to a statement about Ulam’s distance

given in [6, Chapter 6B, Lemma 2]. Note that our proof is constructive (via Definition 3.2 of the map ρ), in contrast to the proof of the latter.

Note that, in general, applying the function ρ is not the same as applying an Ulam move, as in the last case of ρ . We give an algorithm for sorting a permutation w to the identity permutation by applying a shortest sequence of Ulam moves, as follows. First, we apply ρ repeatedly until we arrive at the identity permutation, giving us an optimal run word. Next, we read the runs of this optimal run word from right to left, one run at a time; each run corresponds to applying an Ulam move. In the following example, we demonstrate our algorithm for sorting a permutation.

Example 3.11. Continuing with Example 3.7, let $w = 51642738$. We apply ρ repeatedly: $\rho(w) = [23]w = 51623748$, $\rho^2(w) = [1234]\rho(w) = 12634758$, $\rho^3(w) = [345]\rho^2(w) = 12345768$, and $\rho^4(w) = [6]\rho^3(w) = 12345678$. This gives an optimal run word $[32 \cdot 4321 \cdot 543 \cdot 6]$ of w which tells us how to sort w using minimally many Ulam moves. Reading the four runs of this optimal run word from right to left, we apply the following Ulam moves to 51642738:

- (1) delete the sixth number, 7, from w and insert it after the number 3, producing 51642378;
- (2) delete the third number, 6, and insert it after the number 3, producing 51423678;
- (3) delete the first number, 5, and insert it after the number 3, producing 14235678;
- (4) delete the second number, 4, and insert it after the number 3, producing the identity permutation 12345678.

4. Canonical Reduced Words and the Second Row of RSK Tableaux

In this section, we construct and study a particular optimal run word for a boolean permutation w . We call it the *RSK canonical reduced word* (*canonical word* for short) of w and denote it by $\text{canon}(w)$. In Sect. 4.1, we will present two different algorithms to produce $\text{canon}(w)$: in Definition 4.1, we start from an arbitrary reduced word of w and apply commutation relations; in Definition 4.2, we construct $\text{canon}(w)$ from the heap H_w , which gives a convenient visualization of $\text{canon}(w)$. In Sect. 4.2, we establish an application of $\text{canon}(w)$. We show that $\text{canon}(w)$ directly produces $P(w)$ and $Q(w)$ without using the RSK insertion procedure.

4.1. Constructing Canonical Words

The first algorithm was inspired by a technique used in the proof of [10, Theorem 6.4]. Essentially, it uses commutation moves to push decreasing runs to the left within the word and increasing runs to the right, starting with runs on the smallest numbers.

Definition 4.1. Let w be a boolean permutation, and $[s] \in R(w)$ an arbitrary reduced word.

Step (1): Let a be the smallest value appearing in $[s]$. Apply commutation moves to push a run to the left or right of w according to the following instructions.

- (a) (*Push a singleton run to the left.*) If $a + 1$ does not appear in $[s]$, then define w' so that $w = [a]w'$.
- (b) (*Push a decreasing run to the left.*) If $a + 1$ appears to the left of a in $[s]$, then let $b \geq a + 1$ be the largest such that the run $b(b-1) \cdots a$ is a subsequence of s . Define w' so that $w = [b(b-1) \cdots a]w'$.
- (c) (*Push an increasing run to the right.*) If $a + 1$ appears to the right of a in $[s]$, then let $b \geq a + 1$ be the largest such that the run $a \cdots (b-1)b$ is a subsequence of s . Define w' so that $w = w'[a \cdots (b-1)b]$.

Step (2): If w' is not the identity permutation, repeat Step (4.1) on an arbitrary reduced word for w' . If w' is the identity, we are done.

The reduced word of w created by this algorithm is $\text{canon}(w)$.

Alternatively, we can construct the same canonical reduced word given in Definition 4.1 using the heap (defined in Sect. 2.2) of a boolean permutation.

Definition 4.2. Let H be the heap diagram of a boolean permutation w , drawn in increasing order from left to right. Start with two empty lists $\text{Dec} := \text{Dec}(H)$ and $\text{Inc} := \text{Inc}(H)$. We will scan the elements of H from left to right and fill these two lists with decreasing and increasing runs, respectively.

Step (1) Let a be the leftmost element (that is, smallest number) of H .

- (a) If $a + 1$ is not an element in H , then append the singleton run a to the list of Dec.
- (b) If $a \succ a + 1$ in H , then let b be the first extremal element of H (necessarily minimal) to the right of a . Append the decreasing run $b(b-1) \cdots a$ to the list Dec.
- (c) If $a \prec a + 1$ in H , then let b be the first extremal element of H (necessarily maximal) to the right of a . Prepend the increasing run $a(a+1) \cdots b$ to the list Inc.

Step (2) Let H' be the diagram obtained by removing the singleton a from H (Case (a)) or the elements $a, a + 1, \dots, b$ (Case (b) or (c)). If H' is not empty, redefine $H := H'$, and repeat Step (4.2). If H' is empty, we are done.

If Dec is nonempty, let $\text{canon}(\text{Dec})$ be the concatenation of the decreasing runs in Dec, with smaller indices appearing first; otherwise, let it be the empty word. If Inc is nonempty, let $\text{canon}(\text{Inc})$ be the concatenation of the increasing runs in Inc, with larger indices appearing first; otherwise, let it be the empty word. The reduced word $\text{canon}(w)$ is the concatenation of $\text{canon}(\text{Dec})$ and $\text{canon}(\text{Inc})$, in that order.

We can verify that Definitions 4.1 and 4.2 are equivalent, as follows. By Proposition 2.4, we know that a reduced word $[s]$ for a boolean permutation w corresponds to a linear extension of the heap H_w . In addition, we observe:

- (1) $x \in H_w$ if and only if $x \in [s]$;
- (2) $x \succ x + 1$ in H_w if and only if $x + 1$ is to the left of x in $[s]$;

(3) $x \prec x + 1$ in H_w if and only if $x + 1$ is to the right of x in $[s]$.

Therefore, we obtain the same decreasing run (including singleton) or increasing run at each iteration. Furthermore, appending a decreasing run is equivalent to pushing a decreasing run to the left, while prepending an increasing run is the same as pushing an increasing run to the right.

Next, we demonstrate this canonical word construction using two boolean permutations: one having full support and one not.

Example 4.3. Consider the boolean permutation $314627(10)589 \in S_{10}$. First, we construct $\text{canon}(w)$ following Definition 4.1.

- We start with an arbitrarily chosen reduced word: $[259136847] \in R(w)$.
- First, $a = 1$. Since $a + 1 = 2$ is to the left of a , and $b = 2$, we push the decreasing run $[21]$ to the left and write $w = [21 \cdot 5936847]$.
- Now, we look at $w' = [5936847]$. In this case, $a = 3$. Since $a + 1 = 4$ is to the right of a , and $b = 4$, we push the increasing run $[34]$ to the right: $w' = [59687 \cdot 34]$.
- Now, we look at $w'' = [59687]$. Here, $a = 5$, we have $a + 1 = 6$ to the right of a , and $b = 7$, so we push the increasing run $[567]$ to the right: $w'' = [98 \cdot 567]$.
- We are left with $[98]$, which is a run, so we are done.

Our steps produce the following:

$$[259136847] \rightsquigarrow [21 \cdot 5936847] \rightsquigarrow [21 \cdot 59687 \cdot 34] \rightsquigarrow [21 \cdot 98 \cdot 567 \cdot 34] = \text{canon}(w).$$

We could also have constructed $\text{canon}(w)$ using heaps, according to Definition 4.2. The heap H of w is shown in Fig. 2. As we go from left to right along the elements of H , we create $\text{Dec} = \{[21], [98]\}$ and $\text{Inc} = \{[567], [34]\}$. Then, $\text{canon}(\text{Dec}) = [21 \cdot 98]$ and $\text{canon}(\text{Inc}) = [567 \cdot 34]$. Their concatenation, in that order, produces the reduced word $\text{canon}(w)$ given above.

The boolean permutation in the next example does not have full support, so its heap diagram is disconnected.

Example 4.4. Consider the boolean permutation $w = 231548697(11)(10) \in S_{11}$. We can again construct $\text{canon}(w)$ following Definition 4.1.

- We start with an arbitrarily chosen reduced word: $[471(10)268] \in R(w)$.
- First, $a = 1$. Since $a + 1 = 2$ is to the right of a , and $b = 2$, we push the increasing run $[12]$ to the right and write $w = [47(10)68 \cdot 12]$.
- Now, we look at $w' = [47(10)68]$. In this case, $a = 4$. Since $a + 1 = 5$ does not appear in $[47(10)68]$, we write $[47(10)68] = [4 \cdot 7(10)68]$.
- Now, we look at $w'' = [7(10)68]$. Here $a = 6$. Since $a + 1 = 7$ is to the left of a , and $b = 7$, we push the decreasing run $[76]$ to the left and write $[7(10)68] = [76 \cdot (10)8]$.
- Now, we look at $w''' = [(10)8]$. Here, $a = 8$ and we push this singleton to the left and write $[(10)8] = [8 \cdot (10)]$.
- What remains is the singleton run $[(10)]$, so we are done.

$$\text{Row}_2(Q(w)) = \{i + 1 \mid i \text{ is the rightmost entry in a run of } \text{canon}(w)\}.$$

Proof. We first note that the two statements are equivalent, thanks to Proposition 2.6 and the fact that reduced words for the inverse permutation w^{-1} are exactly the reverse of the reduced words for w , from which it follows that $\text{canon}(w^{-1})$ is the reverse word of $\text{canon}(w)$ up to reordering the singleton runs. Therefore, it suffices to prove the first statement, about $\text{Row}_2(P(w))$.

It is straightforward to verify this result when w is the identity permutation and when the canonical reduced word for w consists of a single run. Assume now that the statement is true for all boolean permutations with canonical words having k runs for some $k \geq 1$. Consider a boolean permutation $w \in S_n$ having $k + 1$ runs in its canonical word.

Let a be the smallest value in the support of w . There are three cases:

1. $a + 1$ does not appear in any reduced word for w ,
2. $a + 1$ appears to the left of a in all reduced words for w , or
3. $a + 1$ appears to the right of a in all reduced words for w .

Case 1: If $a + 1$ does not appear in any reduced word for w , then a is its own maximal run in all reduced words of w . Thus, $w = [a]w' = w'[a]$ where w' is a boolean permutation, $\text{canon}(w) = [a]\text{canon}(w')$, $\text{supp}(w') \subseteq \{a + 2, \dots, n - 1\}$, and $\text{run}(w') = \text{run}(w) - 1$. Therefore, $\text{Row}_2(P(w')) \subseteq \{a + 3, \dots, n\}$. Furthermore, we have that $w'(i) = i$ for $i \leq a + 1$, that is, the one-line notation of w' is as follows:

$$w' = 1 \ 2 \ \cdots \ a \ (a + 1) \ w'(a + 2) \ \cdots \ w'(n).$$

Since $a + 1 \notin \text{Row}_2(P(w'))$ and the one-line notation for w is

$$w = 1 \ 2 \ \cdots \ (a + 1) \ a \ w'(a + 2) \ \cdots \ w'(n),$$

we have that $\text{Row}_2(P(w)) = \text{Row}_2(P(w')) \cup \{a + 1\}$, and thus, the inductive hypothesis on w' completes the argument.

Case 2: Suppose $a + 1$ is to the left of a in all reduced words for w . As in Definition 4.1, let $b \geq a + 1$ be the maximum value such that the decreasing run $[b(b - 1) \cdots a]$ is a subsequence of $\text{canon}(w)$. Then $w = [b \cdots (a + 1)a]w'$, where w' is a boolean permutation, $\text{canon}(w) = [b \cdots (a + 1)a]\text{canon}(w')$, $\text{supp}(w') \subseteq \{b + 1, \dots, n - 1\}$, and $\text{run}(w') = \text{run}(w) - 1$.

Because $1, \dots, b$ are not in the support of w' , these values are fixed by w' . The permutation w is the result of multiplying w' on the left by the run $[b(b - 1) \cdots a]$. Therefore

$$w = 12 \cdots (a - 1)(b + 1)a(a + 1) \cdots (b - 2)(b - 1)w(b + 1) \cdots w(n),$$

where, for $b + 1 \leq i \leq n$

$$w(i) = \begin{cases} b & \text{if } w'(i) = b + 1, \text{ and} \\ w'(i) & \text{otherwise.} \end{cases}$$

Since the first b values in the one-line notation of w' are fixed, the partial insertion tableau $P_b(w')$ is the 1-row insertion tableau of the identity permutation: $P(12 \cdots b)$. All entries to the right of b in the one-line notation of w' are larger than b , so none of the numbers $1, 2, \dots, b$ will get bumped to the

second row as we continue the insertion algorithm to produce $P(w')$. Thus all numbers in $\text{Row}_2(P(w'))$ are larger than b .

Meanwhile, the partial insertion tableau $P_b(w)$ has $1, \dots, (b-1)$ in the first row and $b+1$ in the second row

$$P_b(w) = \begin{array}{|c|c|c|c|c|c|c|c|} \hline 1 & 2 & & & \cdots & & & b-1 \\ \hline & & & & & & & \\ \hline \end{array}$$

All entries to the right of $b-1$ in the one-line notation of w are larger than $b-1$, so none of the numbers $1, 2, \dots, b-1$ will get bumped to the second row in the formation of $P(w)$.

If $b+1$ bumps some larger value $w'(j)$ [necessarily $w'(j) \neq b+1$] in the construction of $P(w')$, then the same value $w(j) = w'(j)$ would be bumped by b in the construction of $P(w)$. It follows that $\text{Row}_2(P(w)) = \text{Row}_2(P(w')) \cup \{b+1\}$, and the inductive hypothesis on w' completes the proof.

Case 3: Suppose $a+1$ is to the right of a in all reduced words for w . As in Definition 4.1, let $b \geq a+1$ be the maximum value in $\text{canon}(w)$ such that $[a \cdots (b-1)b]$ is a subsequence of $\text{canon}(w)$. Then $w = w'[a(a+1) \cdots b]$ where w' is a boolean permutation, $\text{canon}(w) = \text{canon}(w')[a(a+1) \cdots b]$, $\text{supp}(w') \subseteq \{b+1, \dots, n-1\}$, and $\text{run}(w') = \text{run}(w) - 1$.

The numbers $1, 2, \dots, b$ are again fixed points of w' , so the one-line notation of w' is of the form

$$w' = 12 \cdots (b-1)b w'(b+1) \cdots w'(n).$$

The fact that $w = w'[a(a+1) \cdots b]$ means that

$$w = 1 \cdots (a-1)(a+1)(a+2) \cdots b w'(b+1) a w'(b+2) \cdots w'(n).$$

Because $w'(b+1) > b$, the first $b+1$ values in the one-line notation of w' form an increasing sequence. Therefore the partial insertion tableau $P_{b+1}(w')$ is the 1-row tableau

$$P_{b+1}(w') = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 1 & 2 & & & \cdots & & & b & w'(b+1) \\ \hline \end{array}$$

Meanwhile, the partial insertion tableau $P_{b+1}(w)$ has $1, \dots, \widehat{a+1}, \dots, b, w'(b+1)$ in its first row and $a+1$ in the second row

$$P_{b+1}(w) = \begin{array}{|c|c|c|c|c|c|c|c|} \hline 1 & 2 & \cdots & a & a+2 & \cdots & b & w'(b+1) \\ \hline & & & & & & & \\ \hline \end{array}$$

The remaining steps of the RSK algorithm will bump exactly the same values for w and for w' . Hence, $\text{Row}_2(P(w)) = \text{Row}_2(P(w')) \cup \{a+1\}$, and the result follows from the inductive hypothesis on w' . $\square$

We demonstrate Theorem 4.5 by recalling a previous example.

Example 4.6. Let $w = 314627(10)589$, and recall that $\text{canon}(w) = [21 \cdot 98 \cdot 567 \cdot 34]$, as computed in Example 4.3. The RSK tableaux for w are

$$P(w) = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 2 & 5 & 7 & 8 & 9 \\ \hline 3 & 4 & 6 & 10 & & \\ \hline \end{array} \quad \text{and} \quad Q(w) = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 3 & 4 & 6 & 7 & 10 \\ \hline 2 & 5 & 8 & 9 & & \\ \hline \end{array},$$

confirming Theorem 4.5. That is, $\text{Row}_2(P(w)) = \{2+1, 9+1, 5+1, 3+1\}$ and $\text{Row}_2(Q(w)) = \{1+1, 8+1, 7+1, 4+1\}$.

Corollary 4.7. *If w is a boolean permutation, then $\text{canon}(w)$ is an optimal run word for w .*

Proof. This follows immediately from Theorem 3.9 and Theorem 4.5. $\square$

Another interesting consequence of this result is that certain values cannot appear together in the second row of $P(w)$ when w is boolean. We present an example of this here, and the result will be generalized in Sect. 5.

Corollary 4.8. *If w is a boolean permutation, then $\{i, i+1, i+2\} \not\subseteq \text{Row}_2(P(w))$ for all i .*

Proof. Let w be a boolean permutation with $i, i+1 \in \text{Row}_2(P(w))$. Thus, the canonical reduced word for w has one run, call it r_{i-1} , with leftmost element $i-1$ and another run, call it r_i , with leftmost element i . In particular, r_i is either a singleton or an increasing run.

Recall the algorithm used to construct $\text{canon}(w)$, given in Definition 4.1. If r_i is an increasing run, then $i+1$ must be part of r_i . On the other hand, if r_i is a singleton, then $i+1$ does not appear in any reduced word for w . In either case, it follows from Theorem 4.5 that $i+2 \notin \text{Row}_2(P(w))$. $\square$

5. Characterizing Boolean Insertion Tableaux

As we have observed, the insertion tableau of a boolean permutation has at most two rows. On the contrary, not every 2-row standard tableau is the insertion tableau of some boolean permutation. For example, as a consequence of Corollary 4.8, the following tableau is not the insertion tableau of any boolean permutation:

1	2	3	5
4	6	7	8

In this section, we will characterize the 2-row standard tableaux that are insertion tableaux for boolean permutations. We will rely heavily on the definition of the canonical reduced word from Sect. 4.1.

Definition 5.1. Let L be a set of integers. If, for all integers x and y , with $x > 0$, we have

$$|[y, y+2x] \cap L| \leq x+1,$$

then we will say that L is *uncrowded*. Otherwise, we say that L is *crowded*.

In other words, a set of integers L is crowded if L contains more than $x + 1$ of the integers in some interval of $2x + 1$ integers. We are interested in crowded and uncrowded sets as they pertain to standard tableaux. The following technical lemma is important for Theorem 5.3 and Definition 5.5.

Lemma 5.2. *Let T be a standard tableau with at most two rows, and let $R \subseteq \text{Row}_2(T)$. Then, $R \cup \{1\}$ is uncrowded if and only if R is uncrowded. In particular, $\text{Row}_2(T) \cup \{1\}$ is uncrowded if and only if $\text{Row}_2(T)$ is uncrowded.*

Proof. One direction of this statement is clear, since every subset of an uncrowded set is uncrowded. To prove the other direction, assume for the sake of contradiction that R is uncrowded, but $R \cup \{1\}$ is crowded. This means there is some minimal $x > 0$, such that $|[1, 2x + 1] \cap (R \cup \{1\})| > x + 1$.

Therefore, there are at least $x + 2$ elements in R from the set

$$\{1, 2, \dots, 2x - 1, 2x, 2x + 1\}.$$

Furthermore, since x is minimal, there are at most x elements in R from the set

$$\{1, 2, \dots, 2x - 1\}.$$

Hence, $\{2x, 2x + 1\} \subseteq R$ and there are exactly x elements in R from the set $\{1, 2, \dots, 2x - 1\}$. Since R is uncrowded, $\{1, 2, \dots, 2x - 1\} \cap R = \{2, 4, \dots, 2x - 2\}$. It follows that $\{2, 4, \dots, 2x, 2x + 1\} \subseteq R$.

This is a set of size $x + 1$, and there are only x positive integers smaller than $2x + 1$. However, $\text{Row}_1(T)$ requires at least $x + 1$ such numbers, so this is a contradiction. $\square$

With this lemma in hand, we can prove the main result in this section which will relate uncrowded sets to boolean permutations.

Theorem 5.3. *Let L be a subset of $\{1, \dots, n - 1\}$. Then, $L \cup \{0\}$ is uncrowded if and only if L is the set of leftmost letters in the runs of the canonical reduced word of a boolean permutation.*

Proof. First note that the result is easily checked when $|L|$ is small.

Let us now prove the direction that if $L \cup \{0\}$ is a crowded set then L cannot be the set of leftmost run letters in the canonical reduced word of a boolean permutation. Suppose that $L \cup \{0\}$ is crowded, and let us find a minimally wide set demonstrating this crowding: fix a pair of values $y \in L \cup \{0\}$ and $x > 0$, such that (1) $|[y, y + 2x] \cap (L \cup \{0\})| > x + 1$ and (2) x is minimal for this and all other possible values of y .

The second condition means no proper subset of $[y, y + 2x] \cap (L \cup \{0\})$ is crowded. This implies that one of the following cases holds:

- (i) $\{1, 3, 5, \dots, 2x - 1, 2x\} \subseteq L$,
- (ii) $\{y, y + 1, y + 2\} \subseteq L$, or
- (iii) $\{y, y + 1, y + 3, y + 5, \dots, y + 2x - 1, y + 2x\} \subseteq L$.

We will now prove in all cases that these cannot be leftmost run letters in the canonical word for a boolean permutation.

First, assume that Case (i) holds. Suppose by contradiction that $1, 3, 5, \dots, 2x - 1, 2x$ are leftmost elements in the runs of the canonical reduced word $\text{canon}(b)$ of some boolean permutation b . Then, Theorem 4.5 tells us that $\text{Row}_2(P(b)) \supseteq \{2, 4, 6, \dots, 2x, 2x + 1\}$. However, $\{2, 4, 6, \dots, 2x, 2x + 1\}$ is uncrowded and $\{1, 2, 4, 6, \dots, 2x, 2x + 1\}$ is crowded. By Lemma 5.2, we conclude that this case is impossible.

Case (ii) is Corollary 4.8.

Next, assume that Case (iii) holds. Suppose the elements of the set $\{y, y + 1, y + 3, \dots, y + 2x - 1\}$ are leftmost letters in the canonical reduced word $\text{canon}(b)$ of a boolean permutation b . Then, $\text{canon}(b)$ must include the run

$$y(y - 1) \cdots,$$

which forces those remaining leftmost letters to come from the runs

$$\begin{aligned} &(y + 1)(y + 2) \\ &(y + 3)(y + 4) \\ &\vdots \\ &(y + 2x - 1)(y + 2x), \end{aligned}$$

where the rightmost values are either in those runs or are not in the support of b . In either case, it would be impossible for $y + 2x$ to be a leftmost run letter in $\text{canon}(b)$.

Thus, if $L \cup \{0\}$ is crowded, then L cannot be the set of leftmost run letters of the canonical reduced word of a boolean permutation.

We now show that, whenever $L \cup \{0\}$ is uncrowded, there is a canonical word for some boolean permutation whose leftmost run letters are exactly the elements of L . For $L = \emptyset$, the empty word for the identity permutation satisfies these conditions. Now, say $L \cup \{0\}$ is uncrowded, and assume, inductively, that our result holds for all L' with $|L'| < |L|$. Write $L = \{m_0 < \dots < m_k\}$, and note that, for all i , we have $m_i \geq 2i + 1$. Indeed, if there is an $i > 0$, such that $m_i \leq 2i$, then $L \cup \{0\}$ is crowded, since

$$|[0, 2i] \cap (L \cup \{0\})| = |\{0, m_0, \dots, m_i\}| = i + 2 > i + 1.$$

First, consider the case that $m_i = 2i + 1$ for all i . Then, $L = \{1, 3, \dots, 2k + 1\}$, and the following word—comprised of two-element increasing runs—is the canonical word for a boolean permutation with leftmost run letters exactly the elements of L : $[(2k + 1)(2k + 2) \cdots 34 \cdot 12]$.

Otherwise, choose the smallest j such that $m_j > 2j + 1$, and define

$$L' := \{z - m_j : z \in L \text{ and } z > m_j\}.$$

Note that $L' \cup \{0\} \subseteq \{z - m_j : z \in L \cup \{0\}\}$. The latter set is uncrowded, since it is a shift of the uncrowded set $L \cup \{0\}$. This means $L' \cup \{0\}$ is a subset of an uncrowded set and is therefore also uncrowded. By the inductive hypothesis, there exists a canonical word $[s]$ for a boolean permutation with leftmost run letters exactly the elements of L' .

Define $[t]$ to be the word obtained by adding m_j to each letter in $[s]$. Since $[t]$ is simply a shift of $[s]$, we see that $[t]$ is a canonical word for a boolean

permutation with leftmost run letters $L \setminus \{m_0, \dots, m_j\} = L \setminus \{1, 3, \dots, 2j - 1, m_j\}$. By construction, all letters of $[t]$ are larger than m_j . Since $m_j > 2j + 1$, we also have $m_j - 1 > 2j$. Thus, we conclude that the following word is a canonical word for a boolean permutation with leftmost run letters given by L :

$$[m_j(m_j - 1) \cdot t \cdot (2j - 1)(2j) \cdot \dots \cdot 34 \cdot 12].$$

□

Combining Theorem 5.3 and Theorem 4.5, we can characterize the sets $\text{Row}_2(P(w))$ when w is boolean.

Corollary 5.4. *Let X be a subset of $\{2, \dots, n\}$, and set $L := \{x - 1 : x \in X\}$. The set X is equal to $\text{Row}_2(P(w))$ for some boolean permutation $w \in S_n$ if and only if $L \cup \{0\}$ is uncrowded.*

This result together with Lemma 5.2 motivates us to define a special class of tableaux.

Definition 5.5. Let T be a standard tableau having at most two rows. We say that T is *uncrowded* if the set $\text{Row}_2(T)$ is uncrowded.

Note that a 1-row tableau, for which $\text{Row}_2(T) = \emptyset$, is always uncrowded. We can rephrase Definition 5.5 to say a tableau is uncrowded if and only if it is the insertion tableau of a boolean permutation. Note that this does not mean that an uncrowded tableau can *only* occur as an insertion tableau of a boolean permutation. Because the RSK insertion algorithm is a bijection between permutations and *pairs* of standard tableaux, multiple permutations can have the same insertion tableau. For example, both 3142 (boolean) and 3412 (not Boolean) have the following insertion tableau which is uncrowded, because $\{3, 4\}$ is uncrowded:

1	2
3	4.

The inverse of a boolean permutation is also boolean, because reversing a reduced word will not introduce any repetition among its letters. Therefore, due to Proposition 2.6, there is a statement about recording tableaux that is analogous to the insertion tableau result given in Corollary 5.4.

Corollary 5.6. *Let Q be a standard tableau with at most two rows. Then, Q is the RSK recording tableau of some boolean permutation if and only if $\text{Row}_2(Q)$ is uncrowded.*

In other words, both the insertion and recording tableaux of boolean permutations follow the same characterization: their second rows are uncrowded. However the converse is not true. For example, 3412 is not a boolean permutation, but has uncrowded insertion and recording tableaux.

6. Enumerating Uncrowded Tableaux

In this section, we enumerate uncrowded tableaux via a bijection to a certain set of binary words. A *maximal run* in a word is a factor of maximally many identical symbols, and it is common to use *run-length* to describe the length of a maximal run. For example, the binary word 0111001 has run-lengths 1, 3, 2, and 1, when read from left to right.

Let U_n be the set of standard uncrowded tableaux of size n (including the 1-row tableau), and let X_n be the set of 01-words of length $n - 1$ in which all run-lengths of 1s are odd. We define a map $f : X_n \rightarrow U_n$ as follows. (See Example 6.3.) If x is the word whose letters are all 0s, then let

$$f(x) = \begin{array}{|c|c|c|c|} \hline 1 & 2 & \cdots & n \\ \hline \end{array}.$$

Otherwise, given $x = x_1x_2 \cdots x_{n-1} \in X_n$, we will construct a 2-row tableau $f(x)$. Define $\alpha(x) \subset [1, n]$ and $\beta(x) \subset [2, n]$ as follows:

- $1 \in \alpha(x)$.
- If $x_i = 0$, then $n + 1 - i \in \alpha(x)$.
- If we have a maximal run of 1s starting at index i and ending at index $i + 2k$, then
 - ★ $n + 1 - i \in \beta(x)$,
 - ★ if $2k > 0$, then $n + 1 - j \in \beta(x)$ for all $j \in \{i + 1, i + 3, i + 5, \dots, i + 2k - 1\}$, and
 - ★ if $2k > 0$, then $n + 1 - j \in \alpha(x)$ for all $j \in \{i + 2, i + 4, i + 6, \dots, i + 2k\}$.

Let $f(x)$ be the tableau whose first (resp., second) row is the increasing sequence of entries in $\alpha(x)$ (resp., $\beta(x)$).

We first verify f is a well-defined map from X_n to U_n .

Lemma 6.1. *For each $x \in X_n$, we have $f(x) \in U_n$.*

Proof. We must establish two facts: $\beta(x)$ is uncrowded, and $f(x)$ is a standard tableau.

The requirement that each run-length of 1s is odd means that, for each even length interval of integers $[i, i + 2z]$, we can have at most

$$n + 1 - i \text{ and } \{n + 1 - j \mid j \in \{i + 1, i + 3, \dots, i + 2z - 1\}\}, \text{ or } \\ \{n + 1 - j \mid j \in \{i, i + 2, \dots, i + 2z\}\}$$

in $\beta(x)$. Therefore, we can have at most $z + 1$ integers in $\beta(x)$ from $[i, i + 2z]$. Therefore, $\beta(x)$ is uncrowded.

To see that $\alpha(x)$ and $\beta(x)$ form the first and second rows, respectively, of a standard tableau, we observe that the following is an injective map sending an entry in the second row to a smaller entry in the first row: For a maximally long factor $x_i \cdots x_{i+2k}$ of 1s in x , consider the entries of the interval $[n + 1 - (i + 2k - 1), n + 1 - i]$ in $\beta(x)$.

- If $2k = 0$, then the corresponding singleton set in β is $\{n + 1 - i\}$. Either $i = n - 1$ or $x_{i+1} = 0$. If $i = n - 1$, then $n + 1 - i = n + 1 - (n - 1) = 2$ and we

map $2 \in \beta(x)$ to $1 \in \alpha(x)$. If $x_{i+1} = 0$, then $n+1-(i+1) = n-i \in \alpha(x)$, and we map $n+1-i \in \beta(x)$ to $n-i \in \alpha(x)$.

- If $2k \geq 2$, then we have

$$\{n+1-i\} \cup \{n+1-j \mid j = i+1, i+3, \dots, i+2k-1\} \subseteq \beta(x)$$

and

$$\{n+1-j \mid j = i+2, i+4, \dots, i+2k\} \subseteq \alpha(x).$$

Either $i+2k = n-1$ or $x_{i+2k+1} = 0$. If $i+2k = n-1$, then we map $n+1-(2k-1) \in \beta(x)$ to $1 \in \alpha(x)$. Otherwise, the fact that $x_{i+2k+1} = 0$ gives us an entry in $\alpha(x)$ that is smaller than everything in $\{n+1-i\} \cup \{n+1-j \mid j = i+1, i+3, \dots, i+2k-1\} \subseteq \beta(x)$. This allows us to send each element in $\beta(x)$ to a unique smaller entry in $\alpha(x)$.

Therefore, each $f(x)$ is indeed in U_n . $\square$

In fact, not only does f send elements from X_n to the set U_n , but we can invert this process. That is, the map f is a bijection.

Proposition 6.2. *The map $f : X_n \rightarrow U_n$ is a bijection.*

Proof. We describe the inverse map $g : U_n \rightarrow X_n$ of f .

If $T \in U_n$ is the 1-row tableau, then let $g(T)$ be the word whose letters are all 0s.

Otherwise, let $g(T)$ be the 01-word $x_1x_2 \cdots x_{n-1}$ constructed using the following algorithm. Set $x_i := 0$ for $1 \leq i \leq n-1$. Let β denote the second row of T . Since T is standard, we have $\beta \subset [2, n]$.

- Let z be the largest number in β . Note that $z > 1$, and set $x_{n+1-z} := 1$.
 - ★ If $z-1 \notin \beta$, then let $\beta' := \beta \setminus \{z\}$.
 - ★ Otherwise the uncrowded condition guarantees the existence of a maximal integer $k \geq 1$, such that

$$[z-2k, z] \cap \beta = \{z, z-1, z-3, \dots, z-(2k-1)\}.$$

Set

$$x_j := 1 \text{ for } n+2-z \leq j \leq n+(2k+1)-z,$$

and let $\beta' := \beta \setminus \{z, z-1, z-3, \dots, z-(2k-1)\}$.

- If β' is empty, then we are done. Otherwise, redefine $\beta := \beta'$ and iterate the process.

We first check that the algorithm is well defined. More precisely, we need to prove that $n+(2k+1)-z$ is less than or equal to $n-1$, which is equivalent to proving that z is greater than or equal to $2k+2$. Suppose, for the sake of contradiction, that $z \leq 2k+1$. Because $z-(2k-1) \in \beta$, we have $z-(2k-1) \geq 2$, and thus, $z = 2k+1$. The $k+1$ elements of $S := \{z, z-1, z-3, \dots, z-(2k-1)\}$ are in $\text{Row}_2(T)$. When $z = 2k+1$, we have $S = \{2, 4, \dots, 2k-1, 2k\}$. This means that $S \subseteq \text{Row}_2(T)$ is uncrowded and $S \cup \{1\}$ is crowded. Lemma 5.2 shows this is impossible, and we have reached a contradiction.

Next we show that $x_{n+(2k+2)-z}$, if it exists, will stay equal to 0 at the conclusion of each iteration. To see this, note that $z-(2k+1)$ is not in the

second row of T , so the next largest element in the iteration will be less than $z - (2k + 1)$.

Now, it is apparent that the above algorithm indeed gives a 01-word. Moreover, all run-lengths of 1s are odd; therefore, $g(T) \in X_n$. It is straightforward to check that $f \circ g = \mathbf{1}_{U_n}$ and $g \circ f = \mathbf{1}_{X_n}$. Therefore, we obtain a bijection between U_n and X_n . $\square$

We demonstrate this bijection with an example, beginning with the map f .

Example 6.3. Suppose we have the 01-word

$$x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ x_7 \ x_8 \ x_9 \ x_{10} \ x_{11} \ x_{12} \ x_{13} \ x_{14} \ x_{15} \ x_{16} \ x_{17} \\ x = 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \in X_{18}.$$

Then, $\alpha(x) = \{1, 2, 3, 6, 7, 9, 12, 14, 16, 17\}$, $\beta(x) = \{4, 5, 8, 10, 11, 13, 15, 18\}$ and

$$T = f(x) = \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 2 & 3 & 6 & 7 & 9 & 12 & 14 & 16 & 17 \\ \hline 4 & 5 & 8 & 10 & 11 & 13 & 15 & 18 & & \\ \hline \end{array} \in U_{18}. \quad (6.1)$$

For example, the letter $x_{13} = 0$ tells us that $n + 1 - 13 = 6 \in \alpha(x)$, while the maximal block $x_8 \cdots x_{12}$ of 1s tells us that

- ★ $n + 1 - j \in \beta(x)$ for all $j \in \{8, 8 + 1, 8 + 3\}$, so $\{11, 10, 8\} \subseteq \beta(x)$ and
- ★ $n + 1 - j \in \alpha(x)$ for all $j \in \{8 + 2, 8 + 4\}$, so $\{9, 7\} \subseteq \alpha(x)$.

Next, we illustrate the inverse map g .

Example 6.4. Suppose we have $T \in U_{18}$ as given in (6.1), and let $\beta = \{18, 15, 13, 11, 10, 8, 5, 4\}$ denote the second row of T .

- ★ $z = 18 \in \beta$ and $z - 1 \notin \beta$. Thus, $x_1 = 1$.
- ★ $z = 15 \in \beta$ and $z - 1 \notin \beta$. Thus, $x_4 = 1$.
- ★ $z = 13 \in \beta$ and $z - 1 \notin \beta$. Thus, $x_6 = 1$.
- ★ $z = 11 \in \beta$ and $z - 1, z - 3 \in \beta$, and thus, $k = 2$. Therefore, $x_8 = x_9 = x_{10} = x_{11} = x_{12} = 1$.
- ★ $z = 5 \in \beta$ and $z - 1 \in \beta$, and thus, $k = 1$. Therefore, $x_{14} = x_{15} = x_{16} = 1$.

Therefore, we get $g(T) = 1001010111101110$, and this is exactly the word x from Example 6.3.

The bijection developed above, between uncrowded tableaux U_n and the set X_n of 01-words having odd run-lengths of 1s, allows us to enumerate the set U_n , and related subsets. We present those results in the following corollary. Note that we must adjust for indexing in parts (b) and (c) of the result, because both of the sequences referenced there are for binary words of length n , not $n - 1$.

Corollary 6.5. (a) *The number of uncrowded tableaux in U_n are counted by the sequence [11, A028495], which is known to count the 01-words in X_n .*

Below is a table for the first few entries of $|U_n|$

n	1	2	3	4	5	6	7	8	9	10
$ U_n $	1	2	3	6	10	19	33	61	108	197

- (b) The number of 2-row uncrowded tableaux in U_n , which is $U_n - 1$, are counted by the sequence [11, A077865], which is known to count the 01-words in X_n , not including the all-0s word.
- (c) The uncrowded tableaux in U_n having n in the second row are counted by the sequence [11, A006053], which is known to count the 01-words in X_n that start with the letter 1. Below is a table for the first few entries for the sequence

n	1	2	3	4	5	6	7	8	9	10
$a(n)$	0	1	1	3	4	9	14	28	47	89

Acknowledgements

The authors would like to thank the 2021–2022 Research Community in Algebraic Combinatorics program at ICERM, through which this research took place. The authors thank the organizers and staff for putting together this invigorating and inspiring workshop series. The authors are also grateful to Carolina Benedetti for helpful discussions. This work also benefited from computation using SAGEMATH [5]. Finally, we thank the anonymous reviewers whose suggestions helped improve and clarify this paper.

Data Availability Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of Interest The authors declare that they have no conflicts of interest.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

References

- [1] D. J. Aldous and P. Diaconis. Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem. *Bulletin of the*

- American Mathematical Society*, 36:413–432, 1999. <https://doi.org/10.1090/s0273-0979-99-00796-x>
- [2] M. Bóna and M. Bruner. *Log-concavity, the Ulam distance and involutions*. Preprint [arXiv:1502.05438](https://arxiv.org/abs/1502.05438).
- [3] A. Björner and F. Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.
- [4] S. C. Billey, W. Jockusch, and R. P. Stanley. Some combinatorial properties of Schubert polynomials. *J. Algebraic Combin.*, 2:345–374, 1993. <https://doi.org/10.1023/A:1022419800503>
- [5] The Sage Developers. *Sage Mathematics Software (Version 9.3)*. The Sage Development Team, 2021.
- [6] P. Diaconis. Group representations in probability and statistics. Institute of Mathematical Statistics Lecture Notes-Monograph Series, 11. *Institute of Mathematical Statistics, Hayward, CA*, 1988. vi+198 pp.
- [7] M. Donten-Bury and L. Escobar and I. Portakal. (2023) *Complexity of the usual torus action on Kazhdan-Lusztig varieties*. *Algebr. Comb.* 6 3: 835–861. <https://doi.org/10.5802/alco.279>
- [8] C. Greene. An extension of Schensted’s theorem. *Advances in Math.*, 14:254–265, 1974. [https://doi.org/10.1016/0001-8708\(74\)90031-0](https://doi.org/10.1016/0001-8708(74)90031-0)
- [9] P. Karuppuchamy. On Schubert Varieties. *Communications in Algebra*, 41:1365–1368, 2013. <https://doi.org/10.1080/00927872.2011.635620>
- [10] V. Mazorchuk and B. E. Tenner. Intersecting principal Bruhat ideals and grades of simple modules. *Comb. Theory*, 2(1):14–31, 2022. <https://doi.org/10.5070/C62156886>
- [11] OEIS Foundation Inc. (2022). The On-Line Encyclopedia of Integer Sequences. <http://oeis.org>.
- [12] K. Ragnarsson and B. E. Tenner. Homotopy type of the boolean complex of a Coxeter group. *Adv. Math.*, 22:409–430, 2009. <https://doi.org/10.1016/j.aim.2009.05.007>
- [13] K. Ragnarsson and B. E. Tenner. Homology of the boolean complex. *J. Algebraic Combin.*, 34:617–639, 2011. <https://doi.org/10.1007/s10801-011-0285-5>
- [14] B. E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.
- [15] C. Schensted. Longest increasing and decreasing subsequences. *Canadian J. Math.*, 13:179–191, 1961. <https://doi.org/10.4153/CJM-1961-015-3>
- [16] M. P. Schützenberger. Quelques remarques sur une construction de Schensted. *Math. Scand.*, 12:117–128, 1963. <https://doi.org/10.7146/math.scand.a-10676>

- [17] R. Stanley. *Enumerative Combinatorics, Volume 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, first edition, 1999.
- [18] R. P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
- [19] J. R. Stembridge. On the fully commutative elements of Coxeter groups. *J. Algebraic Combin.*, 5(4):353–385, 1996. <https://doi.org/10.1023/A:1022452717148>
- [20] B. E. Tenner. Pattern avoidance and the Bruhat order. *J. Comb. Theory Ser. A*, 114:888–905, 2007. <https://doi.org/10.1016/j.jcta.2006.10.003>
- [21] S. E. Ulam. Some ideas and prospects in biomathematics. *Annual review of biophysics and bioengineering*, 1:277–292, 1972. <https://doi.org/10.1146/annurev.bb.01.060172.001425>

Emily Gunawan*
Department of Mathematics and Statistics
University of Massachusetts Lowell
Lowell MA
USA
e-mail: emily_gunawan@uml.edu

Jianping Pan
Department of Mathematics
North Carolina State University
Raleigh NC
USA
e-mail: jpan9@ncsu.edu

Heather M. Russell
Department of Mathematics and Statistics
University of Richmond
Richmond VA
USA
e-mail: hrussell@richmond.edu

Bridget Eileen Tenner[†]
Department of Mathematical Sciences
DePaul University
Chicago IL
USA
e-mail: bridget@math.depaul.edu

Communicated by Vasu Tewari
Received: 9 July 2023.
Accepted: 4 February 2024.