Algorithmic and algebraic aspects of unshuffling permutations

Samuele Giraudo, Stéphane Vialette

Algorithmic and algebraic aspects of unshuffling permutations

Samuele Giraudo, Stéphane Vialette

Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM, F-77454, Marne-la-Vallée, France

Abstract

A permutation is said to be a square if it can be obtained by shuffling two order-isomorphic patterns. The definition is intended to be the natural counterpart to the ordinary shuffle of words and languages. In this paper, we tackle the problem of recognizing square permutations from both the point of view of algebra and algorithms. On the one hand, we present some algebraic and combinatorial properties of the shuffle product of permutations. We follow an unusual line consisting in defining the shuffle of permutations by means of an unshuffling operator, known as a coproduct. This strategy allows to obtain easy proofs for algebraic and combinatorial properties of our shuffle product. We besides exhibit a bijection between square $(213, 231)$-avoiding permutations and square binary words. On the other hand, by using a pattern avoidance criterion on directed perfect matchings, we prove that recognizing square permutations is NP-complete.

Contents

☆This paper is an extended version of [GV16].

Email addresses: `samuele.giraudo@u-pem.fr` (Samuele Giraudo), `stephane.vialette@u-pem.fr` (Stéphane Vialette)

Introduction

The *shuffle product*, denoted by $\sqcup\!\sqcup$, is a well-known operation on words first defined by Eilenberg and Mac Lane [EML53]. Given three words u, v_1, and v_2, u is said to be a *shuffle* of v_1 and v_2 if it can be formed by interleaving the letters from v_1 and v_2 in a way that maintains the left-to-right ordering of the letters from each word. Besides purely combinatorial questions, the shuffle product of words naturally leads to the following computational problems:

(i) Given two words v_1 and v_2, compute the set $v_1 \sqcup\!\sqcup v_2$.

(ii) Given three words u, v_1, and v_2, decide if u is a shuffle of v_1 and v_2.

(iii) Given words u, v_1, ..., v_k, decide if u is in $v_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup v_k$.

(iv) Given a word u, decide if there is a word v such that u is in $v \sqcup\!\sqcup v$.

Even if these problems seem similar, they radically differ in terms of time complexity. Let us now review some facts about these. In what follows, n denotes the size of u and m_i denotes the size of each v_i. A solution to Problem *(i)* can be computed in

$$O\left((m_1 + m_2)\binom{m_1 + m_2}{m_1}\right) \tag{0.1}$$

time [Spe86]. An improvement and a generalization of Problem *(i)* has been proposed in [All00], where it is proven that given words v_1, ..., v_k, the iterated shuffle $v_1 \sqcup\!\sqcup \cdots \sqcup\!\sqcup v_k$ can be computed in

$$O\left(\binom{m_1 + \cdots + m_k}{m_1, \ldots, m_k}\right) \tag{0.2}$$

time. Problem *(ii)* is in P; it is indeed a classical textbook exercise to design an efficient dynamic programming algorithm solving it. It can be tested in $O\left(n^2/\log(n)\right)$ time [vLN82]. To the best of our knowledge, the first $O(n^2)$ time algorithm for this problem appeared in [Man83]. This algorithm can easily be extended to check in polynomial-time whether a word is in the shuffle of any fixed number of given words. Nevertheless, Problem *(iii)* is NP-complete [Man83, WH84]. This remains true even if the ground alphabet has size 3 [WH84]. Of particular interest, it is shown in [WH84] that Problem *(iii)* remains NP-complete even if all the words v_i, $i \in [k]$, are identical, thereby proving that, for two words u and v, it is NP-complete to decide whether or not u is in the iterated shuffle of v. Again, this remains true even if the ground alphabet has size 3. Let us now finally focus on Problem *(iv)*. It is shown in [BS14, RV13] that it is NP-complete to decide if a word u is a *square* (w.r.t. the shuffle), that is, a word u with the property that there exists a word v such that u is a shuffle of v with itself. Hence, Problem *(iv)* is NP-complete.

This paper is intended to study a natural generalization of $\sqcup\!\sqcup$, denoted by $\bullet$, as a shuffle of permutations. Roughly speaking, given three permutations π, σ_1, and σ_2, π is said to be a *shuffle* of σ_1 and σ_2 if π (viewed as a word) is a shuffle of two words that are order-isomorphic to σ_1 and σ_2. This shuffle product was first introduced by Vargas [Var14] under the name of *supershuffle*. Our intention in this paper is to study this shuffle product of permutations $\bullet$ both from a combinatorial and from a computational point of view by focusing on *square* permutations, that are permutations π being in the shuffle of a permutation σ with itself. Many other shuffle products on permutations appear in the literature. For instance, in [DHT02], the authors define the *convolution product* and the *shifted shuffle product*. For this last product, π is a shuffle of σ_1 and σ_2 if π is in the shuffle, as words, of σ_1 and the word obtained by incrementing all the letters of σ_2 by the size of σ_1. It is a simple exercise to prove that, given three permutations π, σ_1, and σ_2, deciding if π is in the shifted shuffle of σ_1 and σ_2 is in P.

This paper is organized as follows. In Section 2, we provide a precise definition of $\bullet$. We shall define $\bullet$ in terms of what we call the unshuffling operator Δ. The operator Δ is in fact a coproduct, endowing the linear span of all permutations with a coalgebra structure (see [JR79] or [GR14] for the definition of these algebraic structures). By duality, the unshuffling operator Δ leads to the definition of our shuffle operation on permutations. This approach has many advantages. First, some combinatorial properties of $\bullet$ depend on properties of Δ and those properties are easier to prove on the coproduct side. Second, this approach allows us to obtain a clear description of the multiplicities of the elements appearing in the shuffle of two permutations, which are of interest in their own right from a combinatorial point of view. Section 3 is devoted to showing that the problems related to the shuffle of words has links with the shuffle of permutations. In particular, we show that binary words that are square are in one-to-one correspondence with square permutations avoiding some patterns (Proposition 3.1). Next, Section 4 presents some algebraic and combinatorial properties of $\bullet$. We show that $\bullet$ is associative and commutative (Proposition 4.1), and that if a permutation is a square, its mirror, complement, and inverse are also squares (Proposition 4.3). Finally, Section 5 presents the most important result of this paper: the fact that deciding if a permutation is a square is NP-complete (Proposition 5.10). This result is obtained by exhibiting a reduction from the NP-complete pattern involvement problem [BBL98].

1. Notations and basic definitions

General notations

If S is a finite set, the cardinality of S is denoted by $|S|$, and if P and Q are two disjoint sets, $P \sqcup Q$ denotes the disjoint union of P and Q. For any nonnegative integer n, $[n]$ is the set $\{1, \dots, n\}$.

Words and permutations

We follow the usual terminology on words [CK97]. Let us recall here the most important ones. Let u be a word. The length of u (also called *size*) is denoted by $|u|$. The *empty word*, the only word of null length, is denoted by ϵ. We denote by $\widetilde{u}$ the *mirror image* of u, that is the word $u_{|u|}u_{|u|-1}\dots u_1$. If P is a subset of $[|u|]$, $u_{|P}$ is the subword of u consisting in the letters of u at the positions specified by the elements of P. If u is a word of integers and k is an integer, we denote by $u[k]$ the word obtained by incrementing by k all letters of u. The *shuffle* of two words u and v is the set recursively defined by

$$u \shuffle \epsilon = \{u\} = \epsilon \shuffle u \tag{1.1}$$

and

$$ua \shuffle vb = (u \shuffle vb)a \ \cup \ (ua \shuffle v)b, \tag{1.2}$$

were a and b are letters. For instance,

$$01 \shuffle 20 = \{0120, 0210, 0201, 2010, 2001\}. \tag{1.3}$$

A word u is a *square* if there exists a word v such that u belongs to $v \shuffle v$. For example, 202101 is a square since this word belongs to the set $201 \shuffle 201$.

We denote by S_n the set of permutations of size n and by S the set of all permutations. In this paper, permutations of a size n are specified by words of length n on the alphabet $[n]$ and without multiple occurrences of a letter, so that all above definitions about words remain valid on permutations. The only difference lies on the fact that we shall denote by $\pi(i)$ (instead of π_i) the i-th letter of any permutation π. For any nonnegative integer n, we write $\nearrow_n$ (resp. $\searrow_n$) for the permutation $12\dots n$ (resp. $n\,(n-1)\dots 1$). If π is a permutation of S_n, we denote by $\bar{\pi}$ the *complement* of π, that is the permutation satisfying $\bar{\pi}(i) = n - \pi(i) + 1$ for all $i \in [n]$. The *inverse* of π is denoted by π^{-1}.

If u is a word of integers where no letter occurs more than once, we define the *standarization* of u, $\mathrm{s}(u)$, to be the unique permutation of the same size as u such that for all $i, j \in [|u|]$, $u_i < u_j$ if and only if $\mathrm{s}(u)(i) < \mathrm{s}(u)(j)$. For instance,

$$\mathrm{s}(814637) = 613425. \tag{1.4}$$

In particular, the image of the map s is the set S of all permutations. Two words u and v having the same standarization are *order-isomorphic*. If σ is a permutation, we say that σ *occurs* in π if there is a set of indices P of $[|\pi|]$ such that σ and $\pi_{|P}$ are order isomorphic. When σ does not occur in π, π is said to *avoid* σ. The set of permutations of size n avoiding σ is denoted by $S_n(\sigma)$. The *pattern involvement problem* consists, given two permutations π and σ, in deciding if σ occurs in π. This problem is known to be NP-complete [BBL98].

Directed perfect matchings

A *directed graph* is an ordered pair $\mathcal{G} = (V, A)$ where V is a set whose elements are called *vertices* and A is a set of ordered pairs of vertices, called *arcs* (from a *source* vertex to a *sink* vertex). In this paper, we shall exclusively use $V \subset \mathbb{N}$. Notice that the aforementioned definition does not allow a directed graph to have multiple arcs with same source and target nodes. We shall not allow directed loops (that is, arcs that connect vertices with themselves). Two arcs are *independent* if they do not have a common vertex. An arc (i,i') *contains* an arc (j,j') if $\min(i,i') < \min(j,j') < \max(j,j') < \max(i,i')$. If no arc of $\mathcal{G}$ contains an other arc, we say that $\mathcal{G}$ is *containment-free*. Two arcs (i,i') and (j,j') are *crossing* if $\min(i,i') < \min(j,j') < \max(i,i') < \max(j,j')$. If no arcs of $\mathcal{G}$ are crossing, we say that $\mathcal{G}$ is *crossing-free*. A directed graph is a *directed matching* if all its arcs are independent. A directed matching is *perfect* if every vertex is either a source or a sink.

For any permutation π of an even size $2n$, a *directed perfect matching on* π is a pair $\mathcal{M} = (\mathcal{G}, \pi)$ where $\mathcal{G}$ is a directed perfect matching on the set $[2n]$ of vertices (see Figure 1). The *word of sources* (resp. *word of sinks*) of $\mathcal{M}$ is the subword $\pi(i_1)\pi(i_2)\ldots\pi(i_n)$ of π where

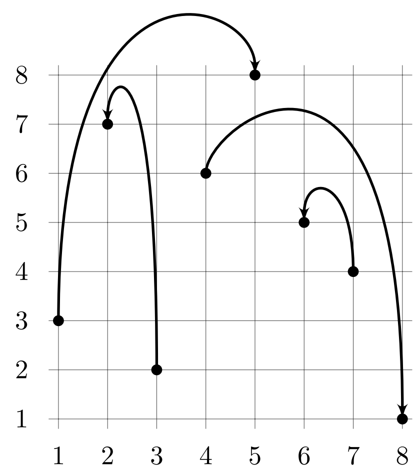


Figure 1: A directed perfect matching $\mathcal{M}$ on the permutation $\pi = 37268541$, represented on the permutation matrix of π. The set of vertices of $\mathcal{M}$ is $\{1,\ldots,8\}$ and the set of arcs of $\mathcal{M}$ is $\{(1,5),(3,2),(4,8),(7,6)\}$.

the indexes $i_1 < i_2 < \cdots < i_n$ are the sources (resp. sinks) of the arcs of $\mathcal{M}$. Figure 2 shows an example for these notions.

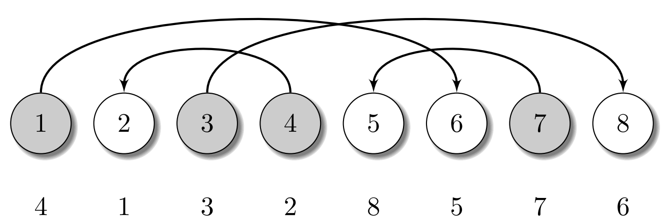


Figure 2: A directed perfect matching $\mathcal{M}$ on the permutation $\pi = 41328576$. The word of sources of $\mathcal{M}$ is **4327** and its word of sinks is 1856. Unlike in Figure 1, $\mathcal{M}$ is not drawn on the permutation matrix of π.

We describe here two notions of patterns for directed perfect matchings on permutations together with the notions of occurrences of patterns accompanying them. Let π be a permutation of size $2n$ and $\mathcal{M} = (\mathcal{G}, \pi)$ be a directed perfect matching on π.

1. An *unlabeled pattern* is a directed perfect matching $\mathcal{U} = ([2k], A)$, where $k \leqslant n$. We say that $\mathcal{M}$ contains an *unlabeled occurrence* of $\mathcal{U}$ if there is an increasing map $\phi : [2k] \to [2n]$ (*i.e.*, $i < j \in [2k]$ implies $\phi(i) < \phi(j)$) such that, if (i, i') is an arc of $\mathcal{U}$ then $(\phi(i), \phi(i'))$ is an arc of $\mathcal{G}$. Observe that this first notion of pattern occurrence does not depend on the permutation π. In other words, $\mathcal{M}$ contains an unlabeled occurrence of $\mathcal{U}$ if $\mathcal{G}$ contains a copy of $\mathcal{U}$ as a subgraph by changing some of its labels if necessary.

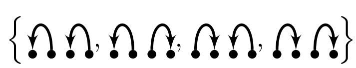
2. A *labeled pattern* is a directed perfect matching $\mathcal{P} = (\mathcal{U}, \sigma)$ on a permutation σ of size $2k$. We say that $\mathcal{M}$ contains a *labeled occurrence* of $\mathcal{P}$ if $\mathcal{M}$ contains an unlabeled occurrence of the directed perfect matching $\mathcal{U} = ([2k], A)$ such that $\mathrm{s}(\pi(\phi(1))\pi(\phi(2)) \ldots \pi(\phi(2k))) = \sigma$, where ϕ is a map defined as above. In other words, $\mathcal{M}$ contains a labeled occurrence of $\mathcal{P}$ if $\mathcal{G}$ contains a copy of $\mathcal{U}$ as a subgraph and the word consisting in the letters of π associated with each vertices of this copy in $\mathcal{G}$ is order-isomorphic to σ.

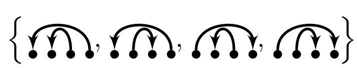
When $\mathcal{M}$ does not contain any unlabeled occurrence (resp. labeled occurrence) of an unlabeled pattern $\mathcal{U}$ (resp. labeled pattern $\mathcal{P}$), we say that $\mathcal{M}$ *avoids* $\mathcal{U}$ (resp. $\mathcal{P}$). This definition naturally extends to sets of patterns by setting that $\mathcal{M}$ *avoids* the set of unlabeled patterns (resp. labeled patterns) $U = \{\mathcal{U}_1, \ldots, \mathcal{U}_\ell\}$ (resp. $P = \{\mathcal{P}_1, \ldots, \mathcal{P}_\ell\}$) if $\mathcal{M}$ avoids every $\mathcal{U}_i$ of U (resp. $\mathcal{P}_i$ of P).

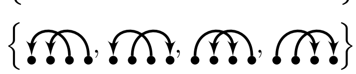
In this paper, we shall consider only patterns of size 4. The set of all unlabeled patterns of this size is

$$\mathcal{P} = \mathcal{P}_{\text{prec}} \cup \mathcal{P}_{\text{cont}} \cup \mathcal{P}_{\text{cros}}, \tag{1.5}$$

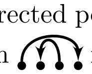
where

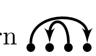
$\mathcal{P}_{\text{prec}} =$  (1.6)

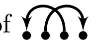
$\mathcal{P}_{\text{cont}} =$  (1.7)

$\mathcal{P}_{\text{cros}} =$  (1.8)

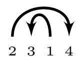
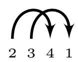
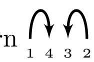
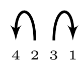
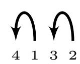
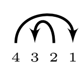
In these drawings, the vertices of each pattern are implicitly indexed from left to right by 1 to 4. Besides, any labeled pattern $\mathcal{P} = (\mathcal{U}, \sigma)$ is depicted by drawing $\mathcal{U}$ and by labeling all its vertices i by σ_i.

To give some examples of the previous notions, observe that a directed perfect matching $\mathcal{M}$ on a permutation contains an occurrence of the unlabeled pattern  if there are four vertices $i_1 < i_2 < i_3 < i_4$ of $\mathcal{M}$ such that (i_1, i_4) and (i_3, i_2) are arcs of $\mathcal{M}$. Moreover, $\mathcal{M}$ is *containment-free* (resp. *crossing-free*) if it avoids all patterns of $\mathcal{P}_{\text{cont}}$ (resp. $\mathcal{P}_{\text{cros}}$). For example, the directed perfect matching on the permutation of Figure 2

- contains exactly two unlabeled occurrences of the pattern  corresponding to the arcs $(1, 6)$ and $(4, 2)$, or $(3, 8)$ and $(7, 5)$;

- contains exactly one unlabeled occurrence of  corresponding to the arcs $(4, 2)$ and $(3, 8)$;
- avoids the unlabeled pattern .

The directed perfect matching on the permutation of Figure 1

- contains a labeled occurrence of the pattern  corresponding to the arcs $(1, 5)$ and $(3, 2)$;
- contains a labeled occurrence of the pattern  corresponding to the arcs $(1, 5)$ and $(4, 8)$;
- contains a labeled occurrence of the pattern  corresponding to the arcs arcs $(1, 5)$ and $(7, 6)$;
- contains a labeled occurrence of the pattern  corresponding to the arcs $(3, 2)$ and $(4, 8)$;
- contains a labeled occurrence of the pattern  corresponding to the arcs $(3, 2)$ and $(7, 6)$;
- contains a labeled occurrence of the pattern  corresponding to the arcs $(4, 8)$ and $(7, 6)$;
- avoids all other labeled patterns of size 4.

2. Shuffle product on permutations

The main purpose of this section is to give a formal definition of the shuffle product $\bullet$ on permutations. We shall define $\bullet$ by first defining a co-product called the unshuffling operator Δ on permutations. Then $\bullet$ is defined to be the dual of Δ. The reason that we define $\bullet$ in terms of Δ is due to the fact that many properties of $\bullet$ depend on properties of Δ and those properties are easier to prove on the co-product side. We invite the reader unfamiliar with the concepts of coproduct and duality to consult [JR79] or [GR14].

Let us denote by $\mathbb{Q}[S]$ the linear span of all permutations. We define a linear coproduct Δ on $\mathbb{Q}[S]$ in the following way. For any permutation π, we set

$$\Delta(\pi) = \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\pi_{|P_1}\right) \otimes \mathrm{s}\left(\pi_{|P_2}\right). \tag{2.1}$$

We call Δ the *unshuffling coproduct of permutations.* For instance,

$$\Delta(213) = \epsilon \otimes 213 + 2 \cdot 1 \otimes 12 + 1 \otimes 21 + 2 \cdot 12 \otimes 1 + 21 \otimes 1 + 213 \otimes \epsilon, \tag{2.2}$$

$$\Delta(1234) = \epsilon \otimes 1234 + 4 \cdot 1 \otimes 123 + 6 \cdot 12 \otimes 12 + 4 \cdot 123 \otimes 1 + 1234 \otimes \epsilon, \tag{2.3}$$

$$\begin{aligned} \Delta(1432) = \epsilon \otimes 1432 + 3 \cdot \mathbf{1} \otimes \mathbf{132} + 1 \otimes 321 + 3 \cdot 12 \otimes 21 \\ + 3 \cdot 21 \otimes 12 + 3 \cdot 132 \otimes 1 + 321 \otimes 1 + 1432 \otimes \epsilon. \end{aligned} \tag{2.4}$$

Observe that the coefficient of the tensor $1 \otimes 132$ is 3 in (2.4) because there are exactly three ways to extract from the permutation 1432 two disjoint subwords which are, respectively, order-isomorphic to the permutations 1 and 132.

We can now define our shuffle product $\bullet$ as the product that corresponds to the coproduct Δ under duality. From (2.1), for any permutation π, we have

$$\Delta(\pi) = \sum_{\sigma, \nu \in S} \lambda_{\sigma,\nu}^{\pi} \; \sigma \otimes \nu, \tag{2.5}$$

where the $\lambda_{\sigma,\nu}^{\pi}$ are nonnegative integers. By the definition (2.1) of Δ, the $\lambda_{\sigma,\nu}^{\pi}$ are equal to the number of different ways to extract from π two disjoint subwords respectively order-isomorphic to σ and ν. Now, by definition of duality, the dual product of Δ, denoted by $\bullet$, is a linear binary product on $\mathbb{Q}[S]$. It satisfies, for any permutations σ and ν,

$$\sigma \bullet \nu = \sum_{\pi \in S} \lambda_{\sigma,\nu}^{\pi} \; \pi, \tag{2.6}$$

where the coefficients $\lambda_{\sigma,\nu}^{\pi}$ are the ones of (2.5). We call $\bullet$ the *shuffle product of permutations*. For instance,

$$\begin{aligned} 12 \bullet 21 = 1243 + 1324 + 2 \cdot 1342 + 2 \cdot 1423 + 3 \cdot \mathbf{1432} + 2134 + 2 \cdot 2314 \\ + 3 \cdot 2341 + 2413 + 2 \cdot 2431 + 2 \cdot 3124 + 3142 + 3 \cdot 3214 + 2 \cdot 3241 \\ + 3421 + 3 \cdot 4123 + 2 \cdot 4132 + 2 \cdot 4213 + 4231 + 4312. \end{aligned} \tag{2.7}$$

Observe that the coefficient 3 of the permutation 1432 in (2.7) comes from the fact that the coefficient of the tensor $12 \otimes 21$ is 3 in (2.4).

Intuitively, the product $\bullet$ shuffles the values and the positions of the letters of the permutations. One can observe that the empty permutation ϵ is a unit for $\bullet$ and that this product is graded by the sizes of the permutations (*i.e.*, the product of a permutation of size n with a permutation of size m produces a sum of permutations of size $n + m$).

We say that a permutation π *appears* in the shuffle $\sigma \bullet \nu$ of two permutations σ and ν if the coefficient $\lambda_{\sigma,\nu}^{\pi}$ defined above is different from zero. In a more combinatorial way, this is equivalent to say that there are two sets P_1 and P_2 of disjoints indexes of letters of π satisfying $P_1 \sqcup P_2 = [|\pi|]$ such that the subword $\pi_{|P_1}$ is order-isomorphic to σ and the subword $\pi_{|P_2}$ is order-isomorphic to ν.

A permutation π is a *square* if there is a permutation σ such that π appears in $\sigma \bullet \sigma$. In this case, we say that σ is a *square root* of π. Equivalently, π is a square with σ as square root if and only if in the expansion of $\Delta(\pi)$, there is a tensor $\sigma \otimes \sigma$ with a nonzero coefficient. In a more combinatorial way, this is equivalent to saying that there are two sets P_1 and P_2

of disjoints indexes of letters of π satisfying $P_1 \sqcup P_2 = [|\pi|]$ such that the subwords $\pi_{|P_1}$ and $\pi_{|P_2}$ are order-isomorphic. Computer experiments give us the first numbers of square permutations with respects to their size, which are, from size 0 to 10,

$$1, 0, 2, 0, 20, 0, 504, 0, 21032, 0, 1293418. \tag{2.8}$$

This sequence (and its subsequence obtained by removing the 0's) is for the time being not listed in [Slo]. The first square permutations are listed in Table 1.

Size 0	Size 2	Size 4
ϵ	12, 21	1234, 1243, 1423, 1324, 1342, 4132, 3124, 3142, 3412, 4312, 2134, 2143, 2413, 4213, 2314, 2431, 4231, 3241, 3421, 4321

Table 1: The square permutations of sizes 0 to 4.

3. Binary square words and permutations

In this section, we shall show that the square binary words are in one-to-one correspondence with square permutations avoiding some patterns. This property establishes a link between the shuffle of binary words and our shuffle of permutations and allows us to obtain a new description of square binary words.

Let u be a binary word of length n with k occurrences of 0. We denote by btp (*Binary word To Permutation*) the map sending any such word u to the permutation obtained by replacing from left to right each occurrence of 0 in u by 1, 2, ..., k, and from right to left each occurrence of 1 in u by $k+1$, $k+2$, ..., n. For instance,

$$\mathrm{btp}(\mathbf{1}00\mathbf{1}0\mathbf{11}0\mathbf{1}0\mathbf{1}000) = \mathbf{C}12\mathbf{B}3\mathbf{A}\mathbf{9}4\mathbf{8}567, \tag{3.1}$$

where A, B, and C respectively stand for 10, 11, and 12. Observe that for any nonempty permutation π in the image of btp, there is exactly one binary word u such that $\mathrm{btp}(u0) = \mathrm{btp}(u1) = \pi$. In support of this observation, when π has an even size, we denote by $\mathrm{ptb}(\pi)$ (*Permutation To Binary word*) the word ua such that $|ua|_0$ and $|ua|_1$ are both even, where $a \in \{0, 1\}$. For instance,

$$\mathrm{ptb}(\mathbf{6}1\mathbf{54}23) = \mathbf{1}0\mathbf{11}00 \qquad \text{and} \qquad \mathrm{ptb}(1\mathbf{4}2\mathbf{3}) = 0\mathbf{1}0\mathbf{1}. \tag{3.2}$$

Proposition 3.1. *For any $n \geqslant 0$, the map* btp *restricted to the set of square binary words of length $2n$ is a bijection between this last set and the set of square permutations of size $2n$ avoiding the patterns* 213 *and* 231.

Proof of Proposition 3.1. The statement of the proposition is a consequence of the following claims implying that ptb is the inverse map of btp over the set of square binary words.

Claim 3.2. *The image of* btp *is the set of all permutations avoiding* 213 *and* 231.

Proof of Claim 3.2. Let us first show that the image of btp contains only permutations avoiding 213 and 231. Let u be a binary word, $\pi = \mathrm{btp}(u)$, and P_0 (resp. P_1) be the set of the positions of the occurrences of 0 (resp. 1) in u. By definition of btp, from left to right, the subword $v = \pi_{|P_0}$ is increasing and the subword $w = \pi_{|P_1}$ is decreasing, and all letters of w are greater than those of v. Now, assume that 123 occurs in π. Then, since v is increasing and w is decreasing, there is an occurrence of 3 (resp. 13, 23) in v and a relative occurrence of 21 (resp. 2, 1) in w. All these three cases contradict the fact that all letters of w are greater than those of v. A similar argument shows that π avoids 231 as well.

Finally, observe that any permutation π avoiding 213 and 231 necessarily starts by the smallest possible letter or the greatest possible letter. This property is then true for the suffix of π obtained by deleting its first letter, and so on for all of its suffixes. Thus, by replacing each letter a of π by 0 (resp. 1) if a has the role of a smallest (resp. greatest) letter, one obtains a binary word u such that $\mathrm{btp}(u) = \pi$. Hence, all permutations avoiding 213 and 231 are in the image of btp. □

Claim 3.3. *If u is a square binary word,* $\mathrm{btp}(u)$ *is a square permutation.*

Proof of Claim 3.3. Since u is a square binary word, there is a binary word v such that $u \in v \sqcup\!\sqcup v$. Then, there are two disjoint sets P and Q of positions of letters of u such that $u_{|P} = v = u_{|Q}$. Now, by definition of btp, the words $\mathrm{btp}(u)_{|P}$ and $\mathrm{btp}(u)_{|Q}$ have the same standarization σ. Hence, and by definition of the shuffle product of permutations, $\mathrm{btp}(u)$ appears in $\sigma \bullet \sigma$, showing that $\mathrm{btp}(u)$ is a square permutation. □

Claim 3.4. *If π is a square permutation avoiding* 213 *and* 231*,* $\mathrm{ptb}(\pi)$ *is a square binary word.*

Proof of Claim 3.4. Let π be a square permutation avoiding 213 and 231. By Claim 3.2, π is in the image of btp and hence, $u = \mathrm{ptb}(\pi)$ is a well-defined binary word. Since π is a square permutation, there are two disjoint sets P_1 and P_2 of indexes of letters of π such that $\pi_{|P_1}$ and $\pi_{|P_2}$ are order-isomorphic. This implies, by the definitions of btp and ptb, that $u_{|P_1} = u_{|P_2}$, showing that u is a square binary word. □

This ends the proof of Proposition 3.1 □

The number of square binary words is Sequence **A191755** of [Slo] beginning by

$$1, 0, 2, 0, 6, 0, 22, 0, 82, 0, 320, 0, 1268, 0, 5102, 0, 020632. \tag{3.3}$$

According to Proposition 3.1, this is also the sequence enumerating square permutations avoiding 213 and 231. Notice that it is conjectured in [HRS12] that the number of square binary words of length $2n$ is $\binom{2n}{n}\frac{2^n}{n+1} - \binom{2n-1}{n+1}2^{n-1} + O(2^{n-2})$.

4. Algebraic issues

The aim of this section is to establish some of properties of the shuffle product of permutations $\bullet$. It is worth to note that, as we will see, algebraic properties of the unshuffling coproduct Δ of permutations defined in Section 2 lead to combinatorial properties of $\bullet$.

Proposition 4.1. *The shuffle product $\bullet$ of permutations is associative and commutative.*

Proof of Proposition 4.1. To prove the associativity of $\bullet$, it is convenient to show that its dual coproduct Δ is coassociative, that is

$$(\Delta \otimes I)\Delta = (I \otimes \Delta)\Delta, \tag{4.1}$$

where I denotes the identity map. This strategy relies on the fact that a product is associative if and only if its dual coproduct is coassociative. For any permutation π, we have

$$\begin{aligned}
(\Delta \otimes I)\Delta(\pi) &= (\Delta \otimes I) \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\pi_{|P_1}\right) \otimes \mathrm{s}\left(\pi_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \Delta\left(\mathrm{s}\left(\pi_{|P_1}\right)\right) \otimes I\left(\mathrm{s}\left(\pi_{|P_2}\right)\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \; \sum_{Q_1 \sqcup Q_2 = [|P_1|]} \mathrm{s}\left(\mathrm{s}\left(\pi_{|P_1}\right)_{|Q_1}\right) \otimes \mathrm{s}\left(\mathrm{s}\left(\pi_{|P_1}\right)_{|Q_2}\right) \otimes \mathrm{s}\left(\pi_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 \sqcup P_3 = [|\pi|]} \mathrm{s}\left(\pi_{|P_1}\right) \otimes \mathrm{s}\left(\pi_{|P_2}\right) \otimes \mathrm{s}\left(\pi_{|P_3}\right).
\end{aligned} \tag{4.2}$$

An analogous computation shows that $(I \otimes \Delta)\Delta(\pi)$ is equal to the last member of (4.2), whence the associativity of $\bullet$.

Finally, to prove the commutativity of $\bullet$, we shall show that Δ is cocommutative, that is for any permutation π, if in the expansion of $\Delta(\pi)$ there is a tensor $\sigma \otimes \nu$ with a coefficient λ, there is in the same expansion the tensor $\nu \otimes \sigma$ with the same coefficient λ. Clearly, a product is commutative if and only if its dual coproduct is cocommutative. Now, from the definition (2.1) of Δ, one observes that if the pair (P_1, P_2) of subsets of $[|\pi|]$ contributes to the coefficient of $\mathrm{s}\left(\pi_{|P_1}\right) \otimes \mathrm{s}\left(\pi_{|P_2}\right)$, the pair (P_2, P_1) contributes to the coefficient of $\mathrm{s}\left(\pi_{|P_2}\right) \otimes \mathrm{s}\left(\pi_{|P_1}\right)$. This shows that Δ is cocommutative and hence, that $\bullet$ is commutative. $\square$

Proposition 4.1 shows that $\mathbb{Q}[S]$ under the unshuffling coproduct Δ is a co-associative co-commutative coalgebra which implies, by duality, that $\mathbb{Q}[S]$ under $\bullet$ is an associative commutative algebra

Lemma 4.2. *The three linear maps*

$$\phi_1, \phi_2, \phi_3 : \mathbb{Q}[S] \to \mathbb{Q}[S] \tag{4.3}$$

linearly sending a permutation π to, respectively, $\widetilde{\pi}$, $\bar{\pi}$, and π^{-1} are endomorphisms of associative algebras.

Proof of Lemma 4.2. To prove, for $j = 1, 2, 3$, that ϕ_j is a morphism of associative algebras, we have to prove that for all permutations σ and ν,

$$\phi_j(\sigma \bullet \nu) = \phi_j(\sigma) \bullet \phi_j(\nu). \tag{4.4}$$

By duality, this is equivalent to showing that ϕ_j is a morphism of coalgebras, that is,

$$\Delta\phi_j = (\phi_j \otimes \phi_j)\Delta. \tag{4.5}$$

In the sequel, π is a permutation.

If P is a set of indexes of letters of π, we denote by $\widetilde{P}$ the set $\{|\pi| - i + 1 : i \in P\}$. Now, since the operation $\widetilde{\ }$ defines a bijection on the set of the subsets of $[|\pi|]$, and since the standardization operation commutes with the mirror operation on words without multiple occurrence of a letter, we have

$$\begin{aligned}
\Delta(\phi_1(\pi)) &= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\phi_1(\pi)_{|P_1}\right) \otimes \mathrm{s}\left(\phi_1(\pi)_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\widetilde{\pi}_{|P_1}\right) \otimes \mathrm{s}\left(\widetilde{\pi}_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\widetilde{\pi}_{|\widetilde{P_1}}\right) \otimes \mathrm{s}\left(\widetilde{\pi}_{|\widetilde{P_2}}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \widetilde{\mathrm{s}\left(\pi_{|P_1}\right)} \otimes \widetilde{\mathrm{s}\left(\pi_{|P_2}\right)} \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \phi_1\left(\mathrm{s}\left(\pi_{|P_1}\right)\right) \otimes \phi_1\left(\mathrm{s}\left(\pi_{|P_2}\right)\right) \\
&= (\phi_1 \otimes \phi_1)\Delta(\pi).
\end{aligned} \tag{4.6}$$

This shows that ϕ_1 is a morphism of coalgebras and hence, that ϕ_1 is a morphism of associative algebras.

Next, since by definition of the complementation operation on permutations, for any permutation τ and any indexes i and k, we have $\tau(i) < \tau(k)$ if and only if $\bar{\tau}(i) > \bar{\tau}(k)$, we have

$$\begin{aligned}
\Delta(\phi_2(\pi)) &= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\phi_2(\pi)_{|P_1}\right) \otimes \mathrm{s}\left(\phi_2(\pi)_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\bar{\pi}_{|P_1}\right) \otimes \mathrm{s}\left(\bar{\pi}_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \phi_2\left(\mathrm{s}\left(\pi_{|P_1}\right)\right) \otimes \phi_2\left(\mathrm{s}\left(\pi_{|P_2}\right)\right) \\
&= (\phi_2 \otimes \phi_2)\Delta(\pi).
\end{aligned} \tag{4.7}$$

This shows that ϕ_2 is a morphism of coalgebras and hence, that ϕ_2 is a morphism of associative algebras.

Finally, for any permutation τ, if P is a set of indexes of letters of τ, we denote by P_τ^{-1} the set $\{\tau(i) : i \in P\}$. Since the map sending a subset P of $[|\pi|]$ to P_π^{-1} is a bijection, and since $\mathrm{s}\left(\pi_{|P}\right)^{-1} = \mathrm{s}\left(\pi^{-1}_{|P_\pi^{-1}}\right)$, we have

$$\begin{aligned}
\Delta(\phi_3(\pi)) &= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\phi_3(\pi)_{|P_1}\right) \otimes \mathrm{s}\left(\phi_3(\pi)_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\pi^{-1}_{|P_1}\right) \otimes \mathrm{s}\left(\pi^{-1}_{|P_2}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\pi^{-1}_{|{P_1}_\pi^{-1}}\right) \otimes \mathrm{s}\left(\pi^{-1}_{|{P_2}_\pi^{-1}}\right) \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \mathrm{s}\left(\pi_{|P_1}\right)^{-1} \otimes \mathrm{s}\left(\pi_{|P_2}\right)^{-1} \\
&= \sum_{P_1 \sqcup P_2 = [|\pi|]} \phi_3\left(\mathrm{s}\left(\pi_{|P_1}\right)\right) \otimes \phi_3\left(\mathrm{s}\left(\pi_{|P_2}\right)\right) \\
&= (\phi_3 \otimes \phi_3)\Delta(\pi).
\end{aligned} \tag{4.8}$$

This shows that ϕ_3 is a morphism of coalgebras and hence, that ϕ_3 is a morphism of associative algebras. $\square$

We now use the algebraic properties of $\bullet$ exhibited by Lemma 4.2 to obtain combinatorial properties of square permutations.

Proposition 4.3. *Let π be a square permutation and σ be a square root of π. Then,*

(i) the permutation $\widetilde{\pi}$ is a square and $\widetilde{\sigma}$ is one of its square roots;

(ii) the permutation $\bar{\pi}$ is a square and $\bar{\sigma}$ is one of its square roots;

(iii) the permutation π^{-1} is a square and σ^{-1} is one of its square roots.

Proof of Proposition 4.3. All statements *(i)*, *(ii)*, and *(iii)* are consequences of Lemma 4.2. Indeed, since π is a square permutation and σ is a square root of π, by definition, π appears in the product $\sigma \bullet \sigma$. Now, by Lemma 4.2, for any $j = 1, 2, 3$, since ϕ_j is a morphism of associative algebras from $\mathbb{Q}[S]$ to $\mathbb{Q}[S]$, ϕ_j commutes with the shuffle product of permutations $\bullet$. Hence, in particular, one has

$$\phi_j(\sigma \bullet \sigma) = \phi_j(\sigma) \bullet \phi_j(\sigma). \tag{4.9}$$

Then, since π appears in $\sigma \bullet \sigma$, $\phi_j(\pi)$ appears in $\phi_j(\sigma \bullet \sigma)$ and appears also in $\phi_j(\sigma) \bullet \phi_j(\sigma)$. This shows that $\phi_j(\sigma)$ is a square root of $\phi_j(\pi)$ and implies *(i)*, *(ii)*, and *(iii)*. $\square$

Let us make an observation about Wilf-equivalence classes of permutations restrained on square permutations. Recall that two permutations σ and ν of the same size are *Wilf equivalent* if $|S_n(\sigma)| = |S_n(\nu)|$ for all $n \geqslant 0$. The well-known [SS85] fact that there is a single Wilf-equivalence class of permutations of size 3 together with Proposition 4.3 imply that 123 and 321 are in the same Wilf-equivalence class of square permutations, and that 132, 213, 231, and 312 are in the same Wilf-equivalence class of square permutations. Computer experiments show us that there are two Wilf-equivalence classes of square permutations of size 3. Indeed, the number of square permutations avoiding 123 begins by

$$1, 0, 2, 0, 12, 0, 118, 0, 1218, 0, 14272, \tag{4.10}$$

while the number of square permutations avoiding 132 begins by

$$1, 0, 2, 0, 11, 0, 84, 0, 743, 0, 7108. \tag{4.11}$$

Another consequence of Proposition 4.3 is that its makes sense to enumerate the sets of square permutations quotiented by the operations of mirror image, complement, and inverse. The sequence enumerating these sets begins by

$$1, 0, 1, 0, 6, 0, 81, 0, 2774, 0, 162945. \tag{4.12}$$

All Sequences (4.10), (4.11), and (4.12) (and their subsequences obtained by removing the 0s) are for the time being not listed in [Slo].

5. Algorithmic issues

This section is devoted to proving the NP-hardness of recognizing square permutations. As in the case of words, we shall use a linear graph framework where deciding whether a permutation is a square reduces to computing some specific matching in the associated linear graph [RV13, BS14]. We have, however, to deal with directed graphs/perfect matchings satisfying some precise properties. Let us first define two properties.

Definition 5.1 (Property $\mathbf{P_1}$)**.** *Let π be a permutation. A directed perfect matching $\mathcal{M}$ on π is said to have property $\mathbf{P_1}$ if it avoids the following set of unlabeled patterns:*

$\mathcal{P}_1 = \Big\{$  $\Big\}.$ (5.1)

Observe that the unlabeled patterns of $\mathcal{P}_1$ are the four of $\mathcal{P}_{\text{cont}}$ and the two of $\mathcal{P}_{\text{cros}}$ that have crossing edges in the opposite directions.

Definition 5.2 (Property $\mathbf{P_2}$)**.** *Let π be a permutation. A directed perfect matching $\mathcal{M}$ on π is said to have property $\mathbf{P_2}$ if, for any two distinct arcs (i, i') and (j, j') of $\mathcal{M}$, we have $\pi(i) < \pi(j)$ if and only if $\pi(i') < \pi(j')$.*

The rationale for introducing properties $\mathbf{P_1}$ and $\mathbf{P_2}$ stems from the following lemma.

Lemma 5.3. *Let π be a permutation. The following statements are equivalent:*

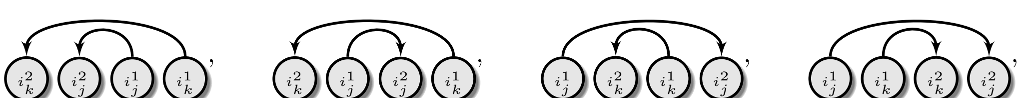
(i) The permutation π is a square.

(ii) There exists a directed perfect matching $\mathcal{M}$ on π satisfying properties $\mathbf{P_1}$ and $\mathbf{P_2}$.

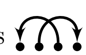
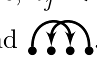
Proof of Lemma 5.3. Assume that *(i)* holds. Since π is a square, π has a square root, say σ. Let $2n = |\pi|$ (and hence $|\sigma| = n$). Then, by definition, there exist two sets

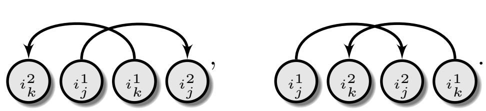
$$I^1 = \left\{i^1_1 < i^1_2 < \cdots < i^1_n\right\} \quad \text{and} \quad I^2 = \left\{i^2_1 < i^2_2 < \cdots < i^2_n\right\} \tag{5.2}$$

of disjoint indexes of letters of π such that $\pi_{|I^1}$ and $\pi_{|I^2}$ are both order-isomorphic to σ. Let $\mathcal{G} = (V, E)$ be the directed graph such that $V = [2n]$ and $E = \left\{\left(i^1_j, i^2_j\right) : j \in [n]\right\}$. It is easily seen that $\mathcal{M} = (\mathcal{G}, \pi)$ is a directed perfect matching since $I^1 \cap I^2 = \emptyset$ and $I^1 \cup I^2 = [2n]$. We first show that $\mathcal{M}$ avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$. Indeed, suppose, aiming at a contradiction, that such an occurrence appears for, say, arcs $\left(i^1_j, i^2_j\right)$ and $\left(i^1_k, i^2_k\right)$ of $\mathcal{M}$. Assuming without loss of generality $i^1_j < i^1_k$, we are left with the four configurations



(5.3)

where shadow nodes give the position in the permutation π. Then it follows that $i^2_j > i^2_k$. This is a contradiction since $i^1_j < i^1_k$ implies $j < k$, and hence, $i^2_j < i^2_k$. We now turn to proving that $\mathcal{M}$ also avoids the unlabeled patterns  and . Indeed, suppose, aiming at a contradiction, that such an occurrence appears for, say, arcs $\left(i^1_j, i^2_j\right)$ and $\left(i^1_k, i^2_k\right)$ of $\mathcal{M}$. Assuming without loss of generality $i^1_j < i^1_k$, we are left with the two configurations



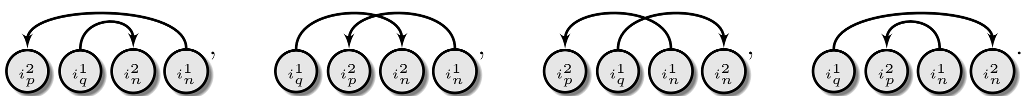
(5.4)

Then it follows that $i_j^2 > i_k^2$. Again, this is a contradiction since $i_j^1 < i_k^1$ implies $j < k$, and hence, $i_j^2 < i_k^2$. Finally, for any two distinct arcs $\left(i_j^1, i_j^2\right)$ and $\left(i_k^1, i_k^2\right)$ of $\mathcal{M}$, we have $\pi\left(i_j^1\right) < \pi\left(i_k^1\right)$ if and only if $\pi\left(i_j^2\right) < \pi\left(i_k^2\right)$ since we are comparing in both cases two elements (at positions j and k) in two patterns that are order-isomorphic to σ. Therefore, $\mathcal{M}$ satisfies properties $\mathbf{P_1}$ and $\mathbf{P_2}$, so that *(ii)* holds.

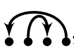
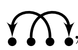
Assume now that *(ii)* holds. Let

$$I^1 = \left\{i_1^1 < i_2^1 < \cdots < i_n^1\right\} \quad \text{and} \quad I^2 = \left\{i_1^2 < i_2^2 < \cdots < i_n^2\right\} \tag{5.5}$$

such that I^1 is the set the sources of the arcs of $\mathcal{M}$ and I^2 is the set of the sinks of the arcs of $\mathcal{M}$. Let us first show that, for every $j \in [n]$, $\left(i_j^1, i_j^2\right)$ is an arc of $\mathcal{M}$. For that, we show that $\left(i_n^1, i_n^2\right)$ is an arc of $\mathcal{M}$. Suppose, aiming at a contradiction that this is false. Then, there exist two vertices i_p^2 and i_q^1 of $\mathcal{M}$ such that $\left(i_n^1, i_p^2\right)$ and $\left(i_q^1, i_n^2\right)$ are arcs of $\mathcal{M}$. Since $p < n$ and $q < n$, there is in $\mathcal{M}$ one of the four configurations



(5.6)

This is a contradiction since $\mathcal{M}$ satisfies property $\mathbf{P_1}$ and hence avoids the unlabeled patterns , , , and . Therefore, $\left(i_n^1, i_n^2\right)$ is an arc of $\mathcal{M}$. By iteratively applying the same reasoning, this also shows that all $\left(i_j^1, i_j^2\right)$, $j \in [n-1]$, are arcs of $\mathcal{M}$. Now, let p^1 be the word of sources and p^2 be the word of sinks of $\mathcal{M}$. Clearly p^1 and p^2 are disjoint in π (since $\mathcal{M}$ is a matching) and cover π (since $\mathcal{M}$ is perfect). Moreover, the fact that $\mathcal{M}$ satisfies $\mathbf{P_2}$ implies immediately that p^1 and p^2 are order-isomorphic. Hence, this shows that π is a square, so that *(i)* holds. $\square$

Observe that, given a square permutation $\pi \in S_{2n}$ and a directed perfect matching $\mathcal{M}$ on π satisfying properties $\mathbf{P_1}$ and $\mathbf{P_2}$, one can recover a square root of π by considering the standarization permutation of the word of sources (or, equivalently, the word of sinks) of $\mathcal{M}$. Figure 3 provides an illustration of Lemma 5.3 and of this observation.

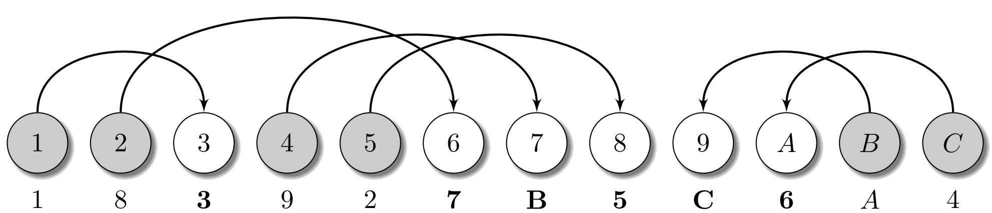
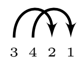


Figure 3: A directed perfect matching $\mathcal{M}$ on the permutation $\pi = 183927B5C6A4$ satisfying the properties $\mathbf{P_1}$ and $\mathbf{P_2}$. From $\mathcal{M}$, it follows that π is a square as it appears in the shuffle of $1892A4$ and $\mathbf{37B5C6}$, both being order-isomorphic to 145263, a square root of π.

Let π be a permutation. For the sake of clarity, we will say that a bunch of consecutive positions P of π is *above* (resp. *below*) another bunch of consecutive positions P' in π if $\pi(i) > \pi(j)$ (resp. $\pi(i) < \pi(j)$) for every $i \in P$ and every $j \in P'$. For example, σ_1 is above

σ_2 (in an equivalent manner, σ_2 is below σ_1) in Figure 6(a), whereas σ_1 is below σ_2 (in an equivalent manner, σ_2 is above σ_1) in Figure 6(b).

Moreover, if π is a permutation satisfying $\pi = \pi_1\sigma_1\pi_2\sigma_2\pi_3$ and $\mathcal{M}$ is a directed perfect matching on π, a (σ_1, σ_2)-*arc* (resp. (σ_2, σ_1)-*arc*) of $\mathcal{M}$ is any arc (i, j) (resp. (j, i)) of $\mathcal{M}$ such that the i-th letter of π belongs to σ_1 and the j-th letter of π belongs to σ_2.

Let us now state and prove some lemmas that will prove extremely useful for simplifying the proof of upcoming Proposition 5.10. First, whereas Lemma 5.3 states that a directed perfect matching on a permutation with Property $\mathbf{P_1}$ avoids some unlabeled patterns of length 4 (more specifically, it avoids the unlabeled patterns of $\mathcal{P}_1$), the following two lemmas state that a directed perfect matching on a permutation with Property $\mathbf{P_2}$ also avoids some additional labeled patterns. These two lemmas are easily proved by requiring Property $\mathbf{P_2}$. For example, an occurrence of the labeled pattern 

induces the existence of two arcs (i_1, i_3) and (i_2, i_4) with $i_1 < i_2 < i_3 < i_4$ and $\pi(i_4) < \pi(i_3) < \pi(i_2) < \pi(i_1)$.

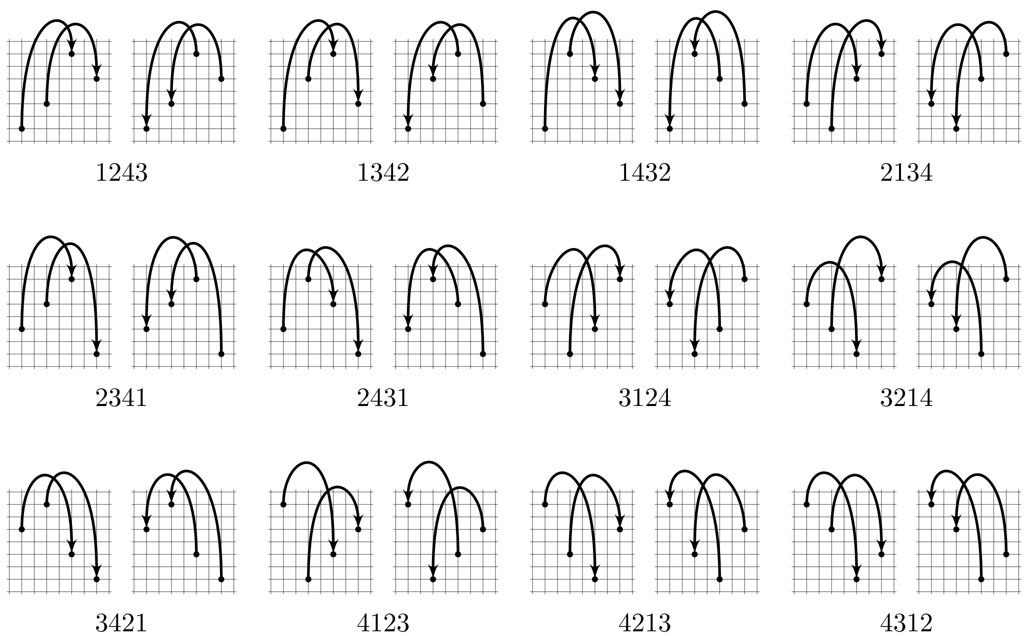
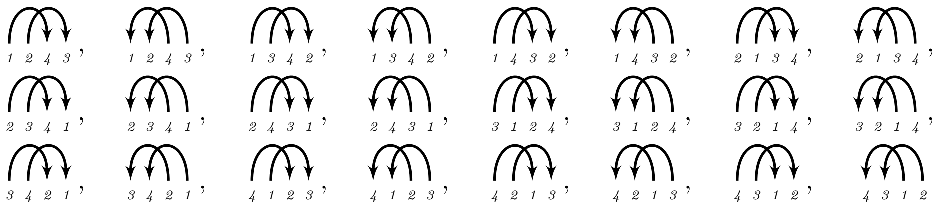


Figure 4: The labeled patterns with crossing edges avoided by any directed perfect matching on a permutation satisfying Property $\mathbf{P_2}$.

Lemma 5.4 (Forbidden crossing patterns)**.** *Let π be a permutation and $\mathcal{M}$ be a directed* *perfect matching on π satisfying Property* $\mathbf{P_2}$*. Then* $\mathcal{M}$ *avoids the following labeled patterns*



(5.7)

see Figure 4.

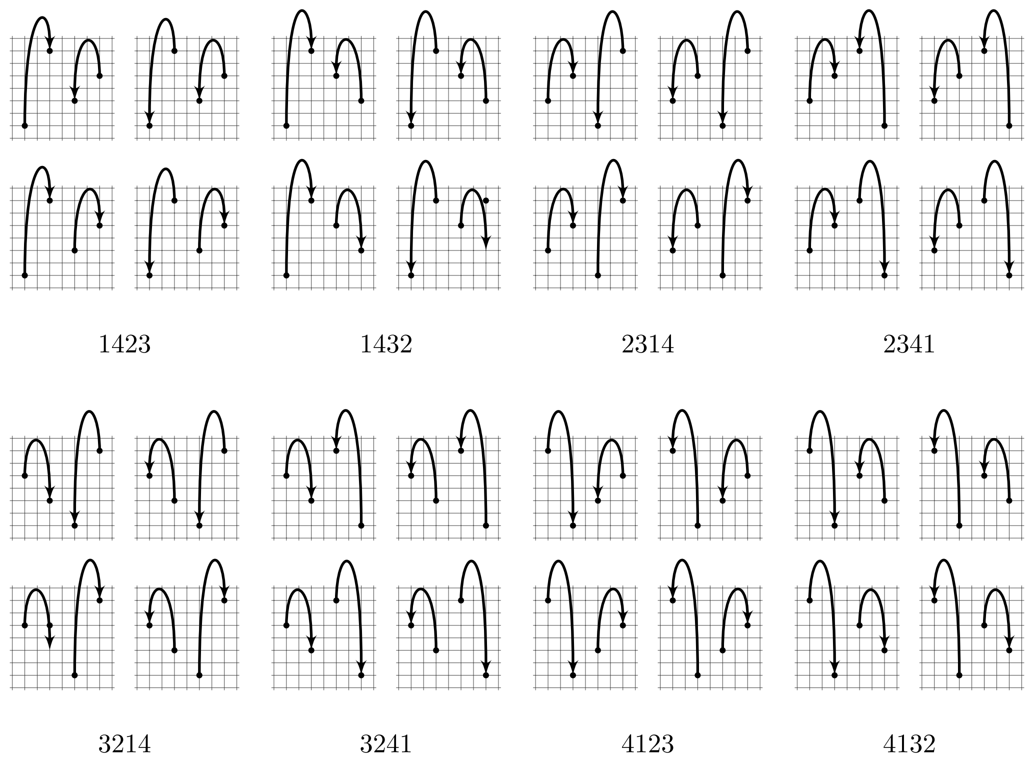
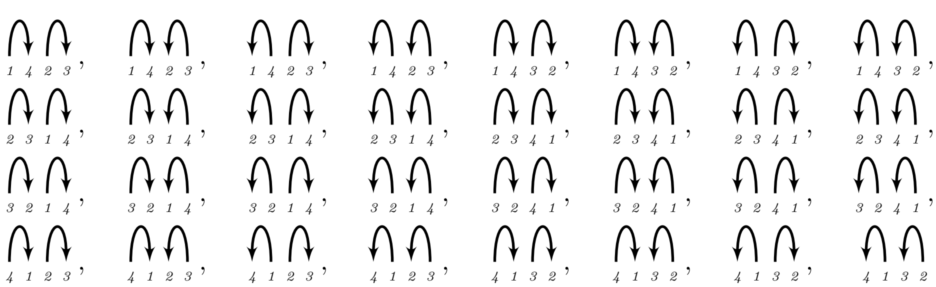


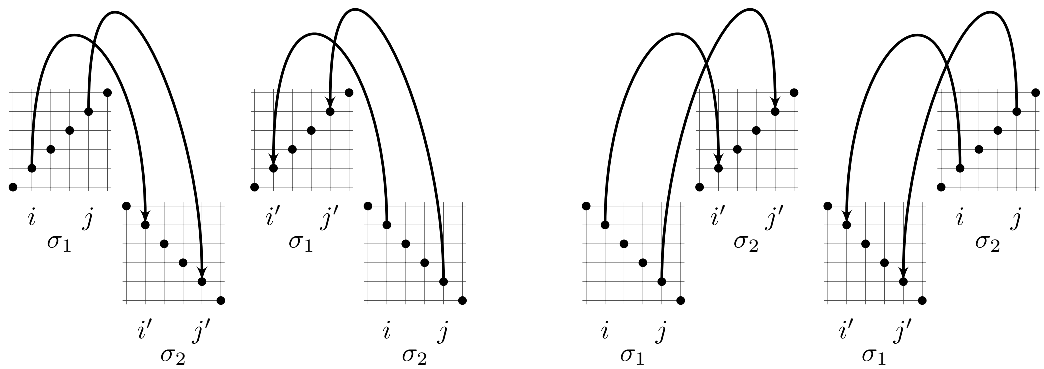
Figure 5: The labeled patterns with consecutive edges avoided by any directed perfect matching on a permutation satisfying Property $\mathbf{P_2}$.

Lemma 5.5 (Forbidden precedence patterns)**.** *Let π be a permutation and* $\mathcal{M}$ *be a directed perfect matching on π satisfying Property* $\mathbf{P_2}$*. Then* $\mathcal{M}$ *avoids the following labeled patterns*



(5.8)

see Figure 5.



(a) An increasing pattern before and above a decreasing pattern.

(b) A decreasing pattern before and below an increasing pattern.

Figure 6: Illustration of Corollary 5.6.

A useful corollary of Lemma 5.4 reads as follows.

Corollary 5.6. *Let $\pi = \pi_1\, \sigma_1\, \pi_2\, \sigma_2\, \pi_3$ be a permutation and $\mathcal{M}$ be a directed perfect matching on π satisfying Properties $\mathbf{P_1}$ and $\mathbf{P_2}$. The following assertions hold.*

(i) If σ_1 is increasing, σ_2 is decreasing, and σ_1 is above σ_2 (see Figure 6(a)), then there is at most one arc between σ_1 and σ_2 in $\mathcal{M}$ (this arc can be a (σ_1, σ_2)-arc or a (σ_2, σ_1)-arc).

(ii) If σ_1 is decreasing, σ_2 is increasing, and σ_1 is below σ_2 (see Figure 6(b)), then there is at most one arc between σ_1 and σ_2 in $\mathcal{M}$ (this arc can be a (σ_1, σ_2)-arc or a (σ_2, σ_1)-arc).

Proof of Corollary 5.6. Suppose, aiming at a contradiction, that *(i)* does not hold. Since $\mathcal{M}$ has Property $\mathbf{P_1}$, it avoids the unlabelled patterns of $\mathcal{P}_1$. Then it follows that $\mathcal{M}$ contains (see Figure 6(a)) either two crossing (σ_1, σ_2)-arcs (a 3 4 2 1 labeled pattern) or two crossing (σ_2, σ_1)-arcs (a 3 4 2 1 labeled pattern). Hence, according to Lemma 5.4, $\mathcal{M}$ cannot have Property $\mathbf{P_2}$. This is the sought-after contradiction.

The proof for *(ii)* is similar (see Figure 6(b)) replacing the labeled patterns 3 4 2 1 and 3 4 2 1 by 2 1 3 4 and 2 1 3 4. □

Lemma 5.7. *Let $\pi = \pi_1\, \sigma_1\, \pi_2\, \sigma_2\, \pi_3$ be a permutation where σ_1 is an increasing pattern and σ_2 is (right) below σ_1, and $\mathcal{M}$ be a directed perfect matching on π that has Properties $\mathbf{P_1}$ and $\mathbf{P_2}$. If $\mathcal{M}$ contains a (σ_1, σ_2)-arc or a (σ_2, σ_1)-arc, then it does not contain a (σ_1, σ_1)-arc.*

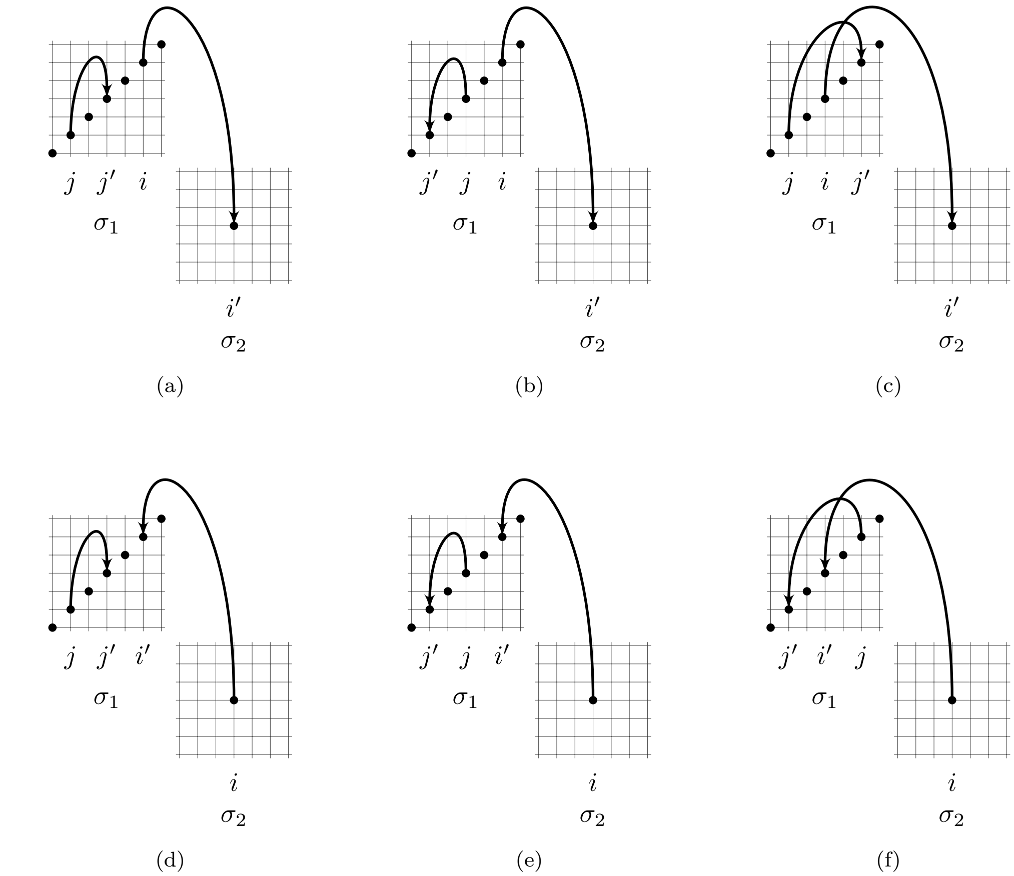
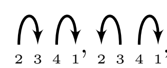
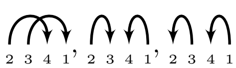
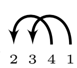


Figure 7: Illustration of Lemma 5.7.

Proof of Lemma 5.7. Suppose, aiming at a contradiction, that $\mathcal{M}$ contains a (σ_1, σ_2)-arc or a (σ_2, σ_1)-arc (i, i'), and a (σ_1, σ_1)-arc (j, j'). Since $\mathcal{M}$ has Property $\mathbf{P_1}$, it avoids the unlabeled patterns of $\mathcal{P}_1$. Therefore, $\mathcal{M}$ contains one of the following labeled patterns: ,  and  (see Figure 7). Hence, according to Lemmas 5.4 and 5.5, $\mathcal{M}$ cannot have Property $\mathbf{P_2}$. This is the sought-after contradiction. $\square$

Lemma 5.8. *Let π be a permutation and $\mathcal{M}$ be a directed perfect matching on π. If $\mathcal{M}$ has properties $\mathbf{P_1}$ and $\mathbf{P_2}$, then so does the directed perfect matching $\mathcal{M}^{\mathrm{r}}$ obtained from $\mathcal{M}$ by reversing each of its arcs.*

Proof of Lemma 5.8. It is immediate that $\mathcal{M}^{\mathrm{r}}$ satisfies Property $\mathbf{P_2}$, since, for any two arcs (i, i') and (j, j') of $\mathcal{M}$, we have $\pi(i) < \pi(j)$ if and only if $\pi(i') < \pi(j')$. As for Property $\mathbf{P_1}$, it is enough to observe that the set of unlabeled patterns $\mathcal{P}_1$ is closed by arc reversals. $\square$

A direct interpretation of Lemma 5.8 is that, if a permutation π is a square, one can exchange the roles of the two order-isomorphic patterns that cover π. This can also be seen as a consequence of Proposition 4.1 about the commutativity of $\bullet$. Besides, an immediate but useful consequence of Lemma 5.8 reads as follows.

Corollary 5.9. *Let π be a permutation and i and i' be two distinct indexes of π. There exists a directed perfect matching on π with Properties $\mathbf{P_1}$ and $\mathbf{P_2}$ that contains the arc (i, i') if and only of there exists a directed perfect matching on π with Properties $\mathbf{P_1}$ and $\mathbf{P_2}$ that contains the arc (i', i).*

Having disposed of these preliminary observations, we now turn to stating and proving the NP-hardness of the targeted problem.

Proposition 5.10. *Deciding whether a permutation is a square is NP-complete.*

Proof of Proposition 5.10. This decision problem is certainly in NP. To prove that it is NP-complete, we propose a reduction from the pattern involvement problem which is known to be NP-complete [BBL98].

Let $\pi \in S_n$ and $\sigma \in S_k$ be two permutations. Let us set

$$\begin{aligned}
N_4 &= 2(2n + k + 2) + 3,\\
N_3 &= 2(2N_4 + 2n + 2k + 4) + 3,\\
N_2 &= 2(2N_3 + 2N_4 + 2n + 2k + 4) + 3,\\
N_1 &= 2(2N_2 + 2N_3 + 2N_4 + 2n + 2k + 4) + 3.
\end{aligned} \tag{5.9}$$

Notice that N_1, N_2, N_3 and N_4 are polynomials in n. The crucial properties are that

(i) the integers N_1, N_2, N_3 and N_4 are odd;

(ii) the relation

$$N_i > \left(\sum_{i<j\leqslant 4} 2N_j\right) + 2n + 2k + 4 \tag{5.10}$$

holds for every $i \in [k]$.

To construct a new permutation μ from π and σ, we now turn to defining various gadgets (sequences of integers) that will act as building blocks. Recall that, for any permutation $p = p_1\, p_2 \cdots p_x$ of $[x]$ and any non-negative integer y, $p\ [y]$ stand for the sequence $(p_1 + y)\,(p + y)\ \cdots\ (p_x + y))$. Define

$$\begin{aligned}
\sigma' &= ((k+1)\ \sigma\ (k+2))\ [2N_2 + N_4 + 2n + k + 2],\\
\pi' &= ((n+1)\ \pi\ (n+2))\ [2N_2 + N_4 + n + k + 2],\\
\sigma'' &= \sigma\ [2N_2 + N_4],\\
\pi'' &= \pi\ [2N_2 + N_4 + k],\\
\nu_1 &= \nearrow_{N_1}\ [2N_2 + 2N_3 + 2N_4 + 2n + 2k + 4],\\
\nu_1' &= \nearrow_{N_1}\ [N_1 + 2N_2 + 2N_3 + 2N_4 + 2n + 2k + 4],\\
\nu_2 &= \searrow_{N_2}\ [N_2],\\
\nu_2' &= \searrow_{N_2},\\
\nu_3 &= \nearrow_{N_3}\ [2N_2 + 2N_4 + 2n + 2k + 4],\\
\nu_3' &= \nearrow_{N_3}\ [2N_2 + N_3 + 2N_4 + 2n + 2k + 4],\\
\nu_4 &= \searrow_{N_4}\ [2N_2 + N_4 + 2n + 2k + 4],\\
\nu_4' &= \searrow_{N_4}\ [2N_2].
\end{aligned} \tag{5.11}$$

We are now in position to define our target permutation μ (see Figure 8 for an illustration) as

$$\mu = \nu_1\ \nu_2\ \nu_1'\ \nu_3\ \sigma'\ \nu_4\ \nu_2'\ \nu_3'\ \pi'\ \nu_4'\ \pi''\ \sigma''. \tag{5.12}$$

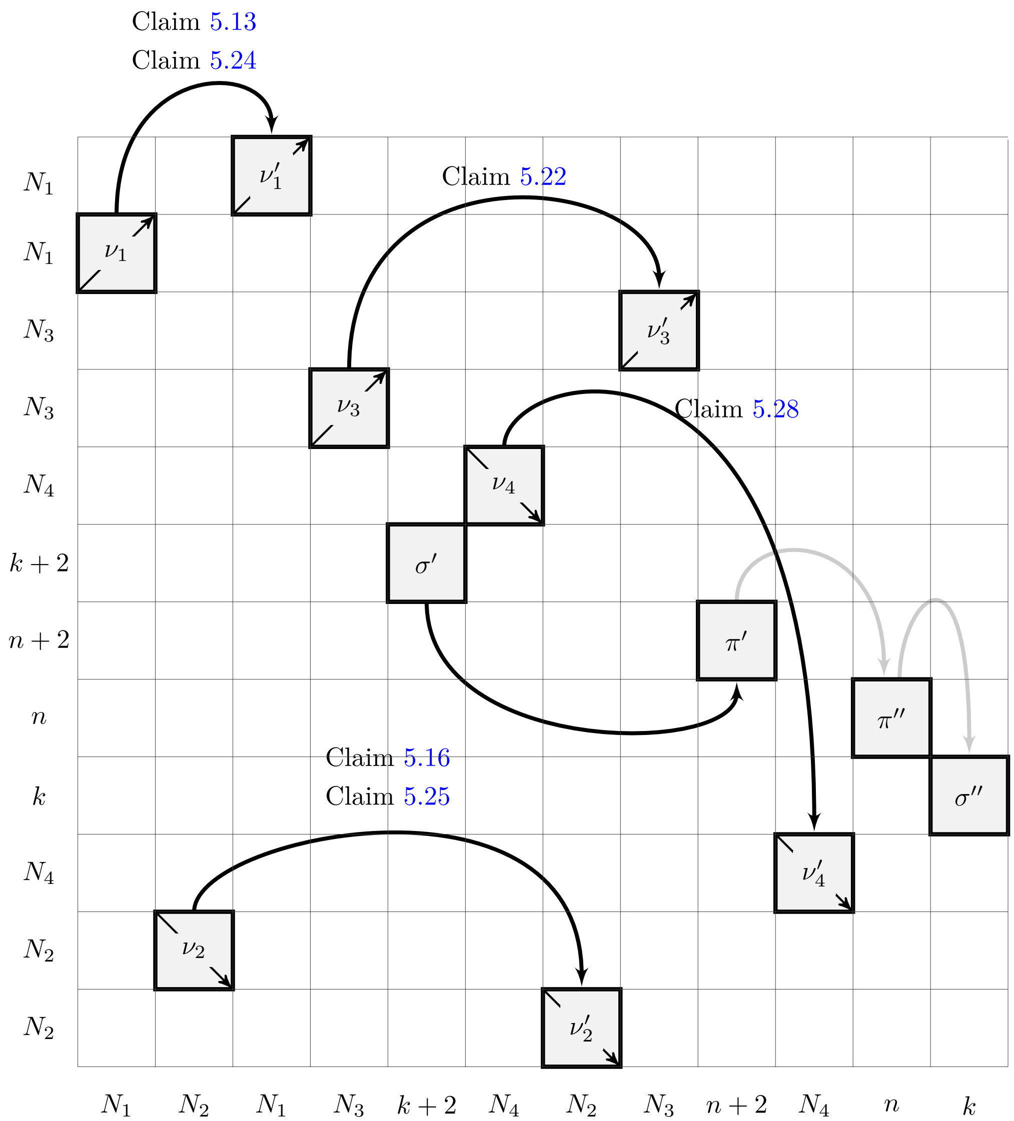


Figure 8: Schematic representation of the permutation μ used in the proof of Proposition 5.10. Black arcs denote the presence of at least one arc between two bunches of positions in μ. Grey arcs denote arcs that are only considered in the forward direction of the proof.

It is immediate that μ can be constructed in polynomial-time in n and k. We claim that σ occurs in π if and only if there exists a directed perfect matching $\mathcal{M}$ on μ that has Properties $\mathbf{P_1}$ and $\mathbf{P_2}$ (that is, by Lemma 5.3, μ is a square).

Suppose first that σ occurs in π and fix any occurrence. Construct a directed matching $\mathcal{M}$ on μ as follows (all arcs are oriented to the right):

(1) $\mathcal{M}$ contains N_1 pairwise crossing (ν_1, ν_1')-arcs.

(2) $\mathcal{M}$ contains N_2 pairwise crossing (ν_2, ν_2')-arcs.

(3) $\mathcal{M}$ contains N_3 pairwise crossing (ν_3, ν_3')-arcs.

(4) $\mathcal{M}$ contains N_4 pairwise crossing (ν_4, ν_4')-arcs.

(5) $\mathcal{M}$ contains $k+2$ pairwise crossing (σ', π')-arcs as depicted in Figure 9. More precisely,

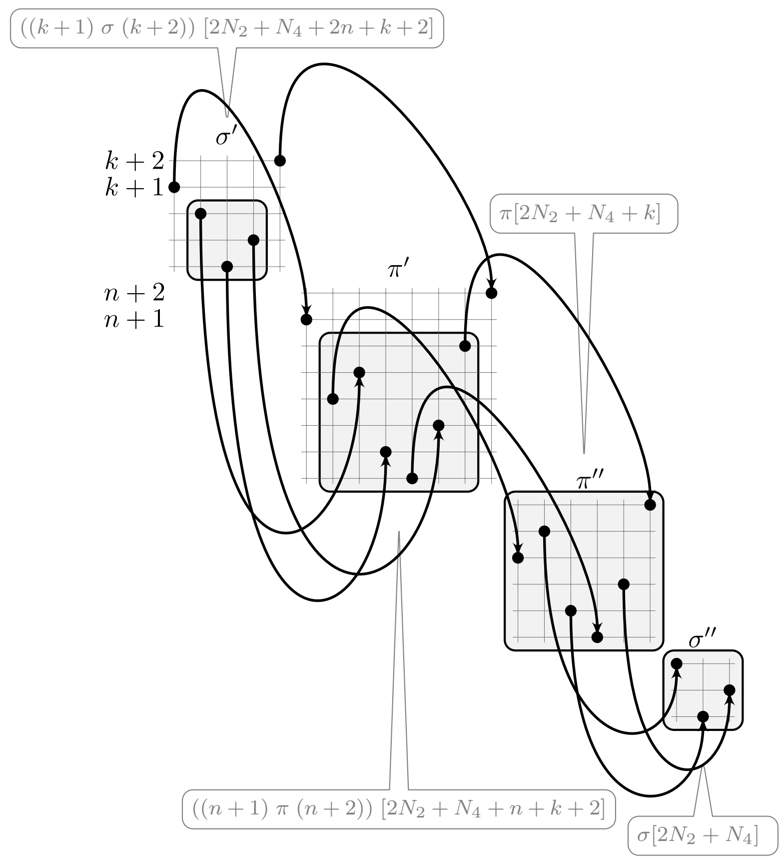


Figure 9: Illustration of the directed perfect matching $\mathcal{M}$ between gadgets σ', π', π'' and σ'' assuming two input permutation $\sigma = 312$ and $\pi = 4\mathbf{52}1\mathbf{3}6$ (where a specific occurrence of σ in π is depicted in bold).

(i) the first position of σ' (*i.e.*, $(2N_1 + N_2 + N_3) + 1$) is linked to the first position of π' (*i.e.*, $(2N_1 + 2N_2 + 2N_3 + N_4 + k + 2) + 1$),

(ii) the last position of σ' (*i.e.*, $(2N_1 + N_2 + N_3) + k + 2$) is linked to the last position of π' (*i.e.*, $(2N_1 + 2N_2 + 2N_3 + N_4 + k + 2) + n + 2$), and all other positions in σ' are linked by means of k pairwise crossing arcs to the positions in π' that correspond to the fixed occurrence of σ in π. (Notice that we use here the fact that σ occurs in π).

(6) $\mathcal{M}$ contains $n-k$ pairwise crossing (π', π'')-arcs as depicted in Figure 9. More precisely, all positions in π' that do not correspond to the fixed occurrence of σ in π are linked by means of $n-k$ pairwise crossing arcs to the positions in π'' that do not correspond to the fixed occurrence of σ in π.

(7) $\mathcal{M}$ contains k pairwise crossing (π'', σ'')-arcs as depicted in Figure 9. More precisely, the positions in π'' that correspond to the fixed occurrence of σ in π are linked by means of k pairwise crossing arcs to all positions in σ''. (Notice that, again, we use here the fact that σ occurs in π).

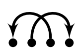
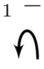
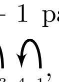
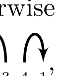
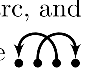
It can be easily checked (probably referring to Figure 8) that $\mathcal{M}$ is perfect and has Properties $\mathbf{P_1}$ and $\mathbf{P_2}$.

Conversely, suppose that there exists an directed perfect matching $\mathcal{M}$ on μ that has Properties $\mathbf{P_1}$ and $\mathbf{P_2}$. We show that σ occurs as a pattern in π. Whereas the directed perfect matching $\mathcal{M}$ may not be as regular as in the forward direction, the main idea is to prove that $\mathcal{M}$ contains enough structure (more precisely, $k+2$ (σ', π')-arcs) so that we can conclude that σ occurs in π. We have divided the reverse direction into a set of basic claims that progressively defines and refines the overall structure of $\mathcal{M}$.

Claim 5.11. *We may assume that there is no (ν_1', ν_1)-arc in $\mathcal{M}$.*

Proof of Claim 5.11. We first observe that, according to Property $\mathbf{P_1}$, since $\mathcal{M}$ avoids the unlabeled patterns of $\mathcal{P}_1$, $\mathcal{M}$ cannot contain both a (ν_1, ν_1')-arc and a (ν_1', ν_1)-arc. Now, if $\mathcal{M}$ does not contain a (ν_1', ν_1)-arc we are done. Otherwise, $\mathcal{M}$ does contain some (ν_1', ν_1)-arcs and no (ν_1, ν_1')-arc, and the result follows from Lemma 5.8. $\square$

Claim 5.12. *There is neither a (ν_1, ν_2)-arc nor a (ν_2, ν_1)-arc in $\mathcal{M}$.*

Proof of Claim 5.12. First, according to Corollary 5.6, there exists at most one arc between ν_1 and ν_2 in $\mathcal{M}$ (this arc can be a (ν_1, ν_2)-arc or a (ν_2, ν_1)-arc). Suppose now, aiming at a contradiction, that there exists either one (ν_1, ν_2)-arc or one (ν_2, ν_1)-arc, say $a = (i, i')$, in $\mathcal{M}$. In this case, according to Lemma 5.7, $\mathcal{M}$ does not contain any (ν_1, ν_1)-arc. We now claim that $\mathcal{M}$ contains $N_1 - 1$ pairwise crossing (ν_1, ν_1')-arcs (and $i = 1$) if a is a (ν_1, ν_2)-arc, or $N_1 - 1$ pairwise crossing (ν_1', ν_1)-arcs (and $i' = 1$) if a is a (ν_2, ν_1)-arc (recall here that  and  are forbidden patterns in $\mathcal{M}$). Indeed, observe first that $N_1 - 1 > |\nu_3| + |\sigma'| + |\nu_4| + |\nu_2'| + |\nu_3'| + |\pi'| + |\nu_4'| + |\pi''| + |\sigma_1|$. Therefore, there exists at least one (ν_1, ν_1')-arc if a is a (ν_1, ν_2)-arc or at least one (ν_1', ν_1)-arc if a is a (ν_2, ν_1)-arc. Hence, if $\mathcal{M}$ does not contain $N_1 - 1$ pairwise crossing (ν_1, ν_1')-arcs or $N_1 - 1$ pairwise crossing (ν_1', ν_1)-arcs, then it contains one of the following labeled patterns: , , , ,  and . Applying Lemma 5.4 and Lemma 5.5 yields a contradiction. Then it follows that $\mathcal{M}$ contains $N_1 - 1$ pairwise crossing (ν_1, ν_1')-arcs (and $i = 1$) if a is a (ν_1, ν_2)-arc, or $N_1 - 1$ pairwise crossing (ν_1', ν_1)-arcs (and $i' = 1$) if a is a (ν_2, ν_1)-arc. But it follows from Claim 5.11 that $\mathcal{M}$ does not contain any (ν_1', ν_1)-arc, and hence $\mathcal{M}$ contains $N_1 - 1$ pairwise crossing (ν_1, ν_1')-arcs and a is a (ν_1, ν_2)-arc (since  is forbidden).

We now observe that $|\nu_1| = |\nu_1'| = N_1$. Hence, since $\mathcal{M}$ is perfect, there exists a position in ν_1' that is not involved in a (ν_1, ν_1')-arc in $\mathcal{M}$. We rule out this configuration by considering two cases:

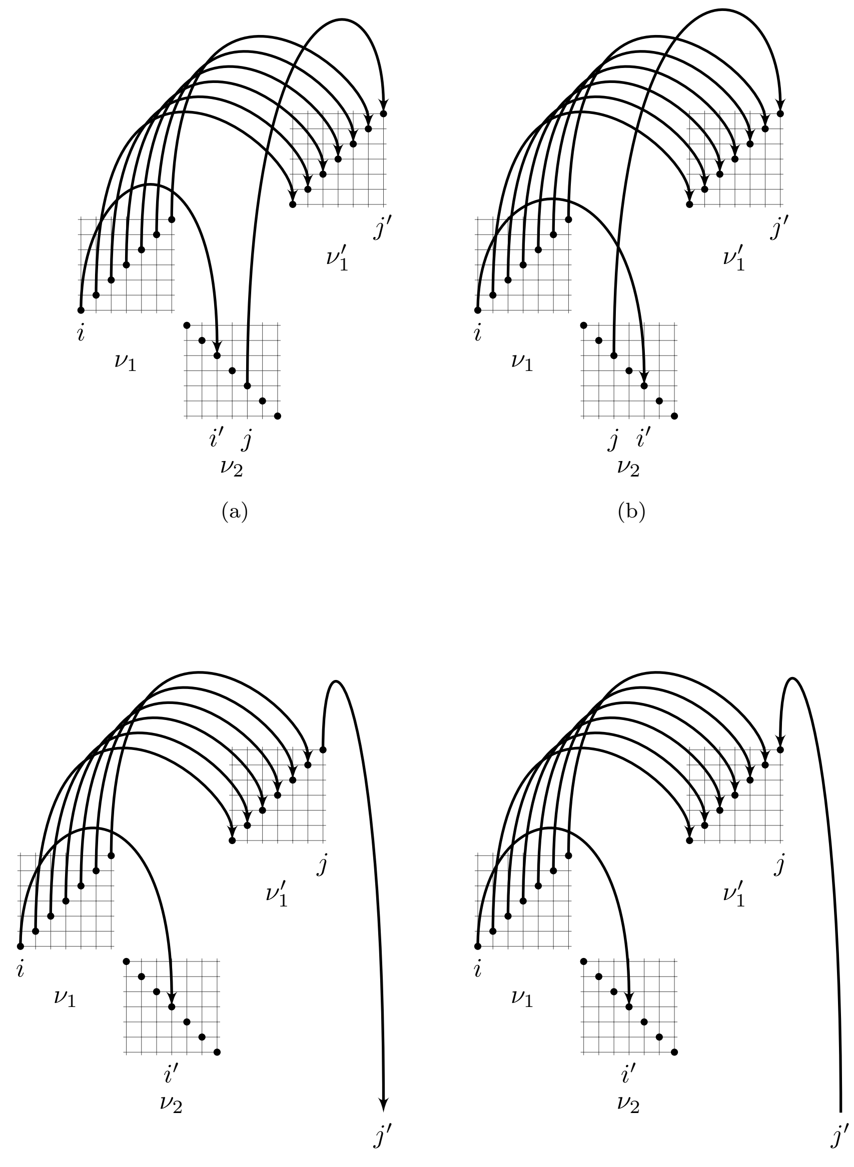
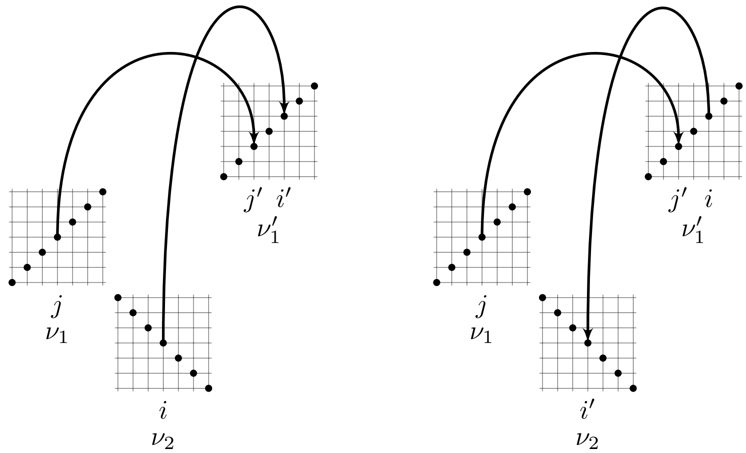


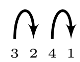
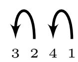
Figure 10: Illustration of Claim 5.12.

- There exists a (ν_2, ν'_1)-arc (j, j') in $\mathcal{M}$ (we cannot have a (ν'_1, ν_2)-arc since the unlabeled pattern is forbidden), see Figure 10(a) and Figure 10(b). Then it follows that $\mathcal{M}$ contains the labeled pattern 2 1 3 4 (with arc (j, j') and any (ν, ν'_1)-arc). Applying Lemma 5.4 yields the sought-after contradiction.



(a) (b)

Figure 11: Illustration of Claim 5.15.

- There exists an arc (j, j') $j = 2N_1 + N_2$ and $j' > 2N_1 + N_2$, or $j' = 2N_1 + N_2$ and $j > 2N_1 + N_2$, see Figure 10(c) and Figure 10(d). Then it follows that $\mathcal{M}$ contains one of the two following labeled patterns:  and  (with arc (i, i') and arc (j, j')). Applying Lemma 5.5 yields the sought-after contradiction.

$\square$

Claim 5.13. *There is at least one (ν_1, ν_1')-arc in $\mathcal{M}$.*

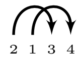
Proof of Claim 5.13. Suppose, aiming at a contradiction, that there is no (ν_1, ν_1')-arc in $\mathcal{M}$. Then it follows that there exists an arc (i, i') in $\mathcal{M}$ that is neither a (ν_1, ν_1)-arc (since N_1 is odd) nor a (ν_1, ν_2)-arc (Claim 5.12) nor a (ν_1, ν_1') (by our contradiction hypothesis). (In other words, $i \leqslant N_1$ and $i' > 2N_1 + N_2$, or $i' \leqslant N_1$ and $i > 2N_1 + N_2$.) Therefore, since $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$), there is neither a (ν_2, ν_2)-arc nor a (ν_1', ν_1')-arc in $\mathcal{M}$. Then it follows that $\mathcal{M}$ contains either N_1 arcs (j, j') with $N_1 + N_2 < j \leqslant 2N_1 + N_2$ and $j' > 2N_1 + N_2$ (if $i \leqslant N_1$ and $i' > 2N_1 + N_2$), or N_1 arcs (i, i') with $N_1 + N_2 < j' \leqslant 2N_1 + N_2$ and $j > 2N_1 + N_2$ (if $i' \leqslant N_1$ and $i > 2N_1 + N_2$), otherwise $\mathcal{M}$ would not be containment-free. But $|\nu_1'| = N_1 > |\nu_3| + |\sigma'| + |\nu_4| + |\nu_2'| + |\nu_3'| + |\pi'| + |\nu_4'| + |\pi''| + |\sigma''|$. This is a contradiction. $\square$

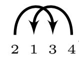
The above claim will be complemented in upcoming Claim 5.24.

Claim 5.14. *There is no (ν_2, ν_2)-arc in $\mathcal{M}$.*

Proof of Claim 5.14. Combine Claim 5.13 together with the fact that $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$). $\square$

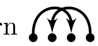
Claim 5.15. *There is neither a (ν_2, ν_1')-arc nor a (ν_1', ν_2)-arc in $\mathcal{M}$.*

Proof of Claim 5.15. First, according to Corollary 5.6–*(ii)*, there exists either at most one (ν_2, ν_1')-arc and no (ν_1', ν_2)-arc, or at most one (ν_1', ν_2)-arc and no (ν_2, ν_1')-arc (i, i') in $\mathcal{M}$ (see Figure 11). Now from Claim 5.13, there exists at least one (ν_1, ν_1')-arc, say (j, j'), in $\mathcal{M}$. Hence, since $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$), $\mathcal{M}$ contains one of the following labeled patterns: 

and 

. Applying Lemma 5.4 yields the sought-after contradiction. $\square$

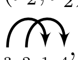
Claim 5.16. *There is at least one (ν_2, ν_2')-arc in $\mathcal{M}$.*

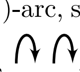
Proof of Claim 5.16. First, according to Claim 5.13, there exists at least one (ν_1, ν_1')-arc in $\mathcal{M}$ and hence, since $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$) there is no (ν_2', ν_2)-arc in $\mathcal{M}$. Now, suppose, aiming at a contradiction, that there is no (ν_2, ν_2')-arc in $\mathcal{M}$. Notice that there is neither a (ν_1, ν_2)-arc (Claim 5.12) nor a (ν_2, ν_1)-arc (Claim 5.12) nor a (ν_2, ν_2)-arc (Claim 5.14) nor a (ν_2, ν_1')-arc (Claim 5.15) nor a (ν_1', ν_2)-arc (Claim 5.15) in $\mathcal{M}$. But $|\nu_2| = N_2 > |\nu_3| + |\sigma'| + |\nu_4| + |\nu_3'| + |\pi'| + |\nu_4'| + |\pi''| + |\sigma''|$. Hence $\mathcal{M}$ cannot be a directed perfect matching, thereby contradicting our hypothesis about $\mathcal{M}$. $\square$

Claim 5.17. *There is neither a (ν_1', ν_1')-arc, nor a (ν_1', ν_3)-arc, nor a (ν_3, ν_1')-arc, nor a (ν_1', σ')-arc, nor a (σ', ν_1')-arc, nor a (ν_1', ν_4)-arc nor a (ν_4, ν_1')-arc nor a (ν_3, ν_3)-arc, nor a (ν_3, σ')-arc, nor a (σ', ν_3)-arc, nor a (ν_3, ν_4)-arc, nor a (ν_4, ν_3)-arc, nor a (σ', σ')-arc, nor a (σ', ν_4)-arc, nor a (ν_4, σ')-arc, nor a (ν_4, ν_4)-arc in $\mathcal{M}$.*

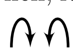
Proof of Claim 5.17. Combine Claim 5.16 with the fact that $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$). $\square$

Claim 5.18. *There is neither a (ν_2', ν_3')-arc, nor a (ν_3', ν_2')-arc, nor a (ν_2', π')-arc, nor a (π', ν_2')-arc, nor a (ν_2', ν_4')-arc, nor a (ν_4', ν_2')-arc, nor a (ν_2', π'')-arc, nor a (π'', ν_2')-arc nor a (ν_2', σ'')-arc, nor a (σ'', ν_2')-arc in $\mathcal{M}$.*

Proof of Claim 5.18. Suppose aiming at a contradiction that $\mathcal{M}$ contains a (ν_2', ν_3')-arc, a (ν_3', ν_2')-arc, a (ν_2', π')-arc, a (π', ν_2')-arc, a (ν_2', ν_4')-arc, a (ν_4', ν_2')-arc, a (ν_2', π'')-arc, a (π'', ν_2')-arc a (ν_2', σ'')-arc or a (σ'', ν_2')-arc, say (i, i'). We now observe that ν_3', π', ν_4', π'' and σ'' are all right above of both ν_2 and ν_2'. Furthermore, according to Claim 5.16, there exists a (ν_2, ν_2')-arc, say (j, j'). Then, it follow that $\mathcal{M}$ contains one of the following labeled patterns: 

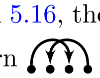
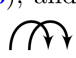
, 

, 

and 

(see Figure 12). Applying Lemma 5.5 and Lemma 5.4 yields the sought-after contradiction. $\square$

Claim 5.19. *There is neither a (ν_3, ν_2')-arc nor a (ν_2', ν_3)-arc in $\mathcal{M}$.*

Proof of Claim 5.19. Suppose, aiming at a contradiction, that there exists (ν_3, ν_2')-arc or a (ν_2', ν_3)-arc (j, j') in $\mathcal{M}$. According to Claim 5.16, there exists at least one (ν_2, ν_2')-arc (i, i') in $\mathcal{M}$. Since $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$), there is no (ν_2', ν_3)-arc in $\mathcal{M}$ (see Figure 13), and hence (j, j') is a (ν_3, ν_2')-arc. Then it follows that $\mathcal{M}$ contains the labeled pattern 

(see Figure 13). Applying Lemma 5.4 yields the sought-after contradiction. $\square$

Claim 5.20. *There is at most one (ν_2, ν_3)-arc or at most one (ν_3, ν_2)-arc in $\mathcal{M}$.*

Proof of Claim 5.20. Apply Corollary 5.6. $\square$

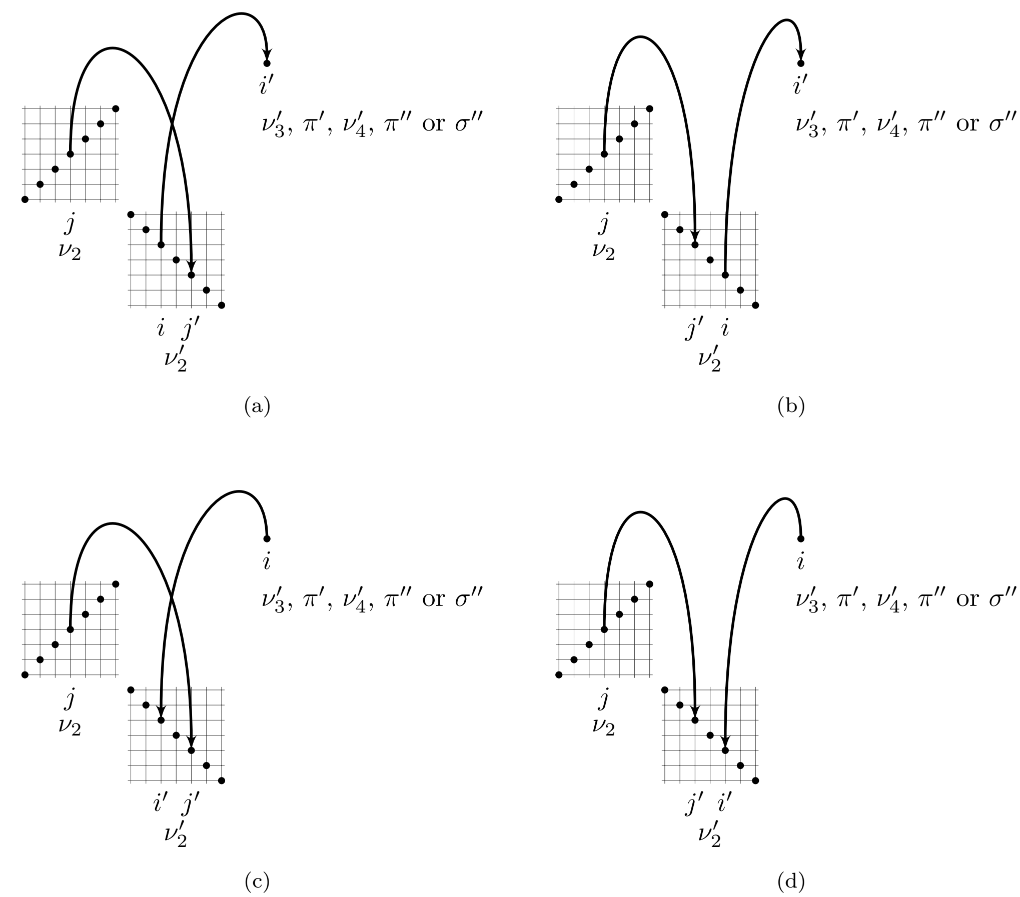
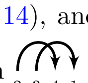


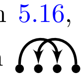
Figure 12: Illustration of Claim 5.18.

We will see soon (upcoming Claim 5.25) that there exists actually no (ν_2, ν_3)-arc in $\mathcal{M}$.

Claim 5.21. *There is neither a (ν_1, ν_3)-arc, nor a (ν_3, ν_1)-arc in $\mathcal{M}$.*

Proof of Claim 5.21. Suppose, aiming at a contradiction, that there exists a (ν_1, ν_3)-arc or a (ν_3, ν_1)-arc, say (j, j'), in $\mathcal{M}$. According to Claim 5.13, there exists at least one (ν_1, ν_1')-arc, say (i, i'), in $\mathcal{M}$. Since $\mathcal{M}$ avoids the unlabeled pattern *CrossingRL* (Property $\mathbf{P_1}$), there is no (ν_3, ν_1)-arc in $\mathcal{M}$ (see Figure 14), and hence (j, j') is a (ν_1, ν_3)-arc. Then it follows that $\mathcal{M}$ contains the labeled pattern  (see Figure 14). Applying Lemma 5.4 yields the sought-after contradiction. $\square$

Claim 5.22. *There exists a (ν_3, ν_3')-arc in $\mathcal{M}$.*

Proof of Claim 5.22. First, according to Claim 5.16, there exists at least one (ν_2, ν_2')-arc in $\mathcal{M}$. Since $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$) there is no (ν_3', ν_3)-arc in $\mathcal{M}$. Now, suppose, aiming at a contradiction, that there is no (ν_3, ν_3')-arc in $\mathcal{M}$. Combining Claim 5.17, Claim 5.19, Claim 5.20 Claim 5.21 together with our hypothesis,

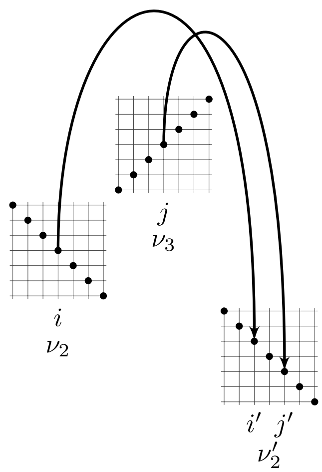


Figure 13: Illustration of Claim 5.19.

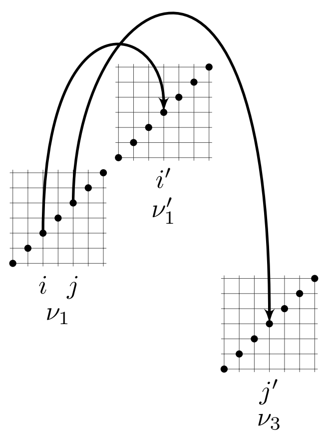
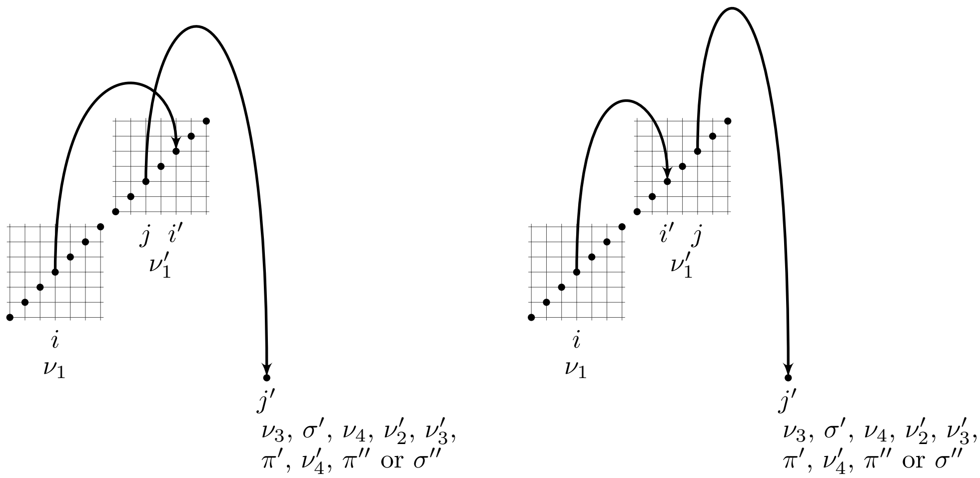


Figure 14: Illustration of Claim 5.21

we conclude that $N_3 - 1$ positions in ν_3 are involved in arcs of $\mathcal{M}$ that are neither (ν_1, ν_3)-arcs, nor (ν_3, ν_1)-arcs, nor (ν_2, ν_3)-arcs, nor (ν_3, ν_2)-arcs, nor (ν_1', ν_3)-arcs, nor (ν_3, ν_1')-arcs, nor (ν_3, ν_3)-arcs, nor (ν_3, σ')-arcs, nor (σ', ν_3)-arcs, nor (ν_3, ν_4)-arcs, nor (ν_4, ν_3)-arcs, nor (ν_3, ν_3')-arcs, nor (ν_3', ν_3)-arcs. But $N_3 - 1 > |\pi'| + |\nu_4'| + |\pi''| + |\sigma''| = N_4 + 2n + 2k + 4$, and hence $\mathcal{M}$ is not a perfect matching. This is the sought-after contradiction. $\square$

Claim 5.23. *There is neither a (σ', σ')-arc, nor a (σ', ν_4)-arc, nor a (ν_4, σ')-arc, nor a*



(a) (b)

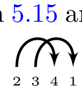
Figure 15: Illustration of Claim 5.24.

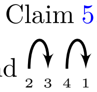
(σ', ν_2')-arc, nor a (ν_2', σ')-arc, nor a (ν_4, ν_4)-arc nor a (ν_4, ν_2')-arc, nor a (ν_2', ν_4)-arc, nor a (ν_2', ν_2')-arc in $\mathcal{M}$.

Proof of Claim 5.23. Combine Claim 5.22 together with the fact that $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$). □

The next two claims state that $\mathcal{M}$ actually contains all (ν_1, ν_1')-arcs and all (ν_2, ν_2')-arcs.

Claim 5.24. *$\mathcal{M}$ contains N_1 pairwise crossing (ν_1, ν_1')-arcs.*

Proof of Claim 5.24. First, according to Claim 5.16, $\mathcal{M}$ contains at least one (ν_1, ν_1')-arc. Now, suppose, aiming at a contradiction, that $\mathcal{M}$ does not contain N_1 (ν_1, ν_1')-arcs. Combining Claim 5.12, Claim 5.15, Claim 5.15 and Claim 5.17, we conclude that $\mathcal{M}$ contains one of the two following labeled patterns: 

and 

(see Figure 15). Applying Lemma 5.4 or Lemma 5.5 yields the sought-after contradiction. □

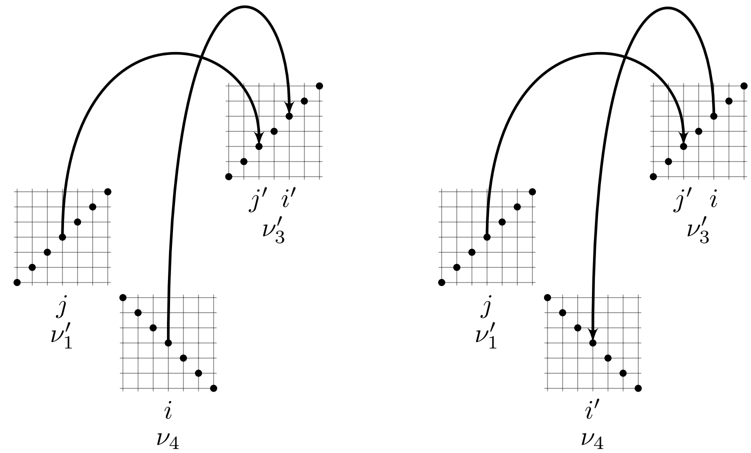
Claim 5.25. *$\mathcal{M}$ contains N_2 pairwise crossing (ν_2, ν_2')-arcs.*

Proof of Claim 5.25. The key idea is to focus on ν_2' and combine Claim 5.16, Claim 5.19, Claim 5.18 and Claim 5.23. □

Claim 5.26. *There is neither a (ν_2, ν_4)-arc nor a (ν_4, ν_2)-arc in $\mathcal{M}$.*

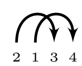
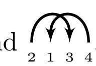
Proof of Claim 5.26. According to Claim 5.25, all positions in ν_2 and ν_2' are involved in (ν_2, ν_2')-arcs in $\mathcal{M}$. □

Claim 5.27. *There is neither a (ν_4, ν_3')-arc nor a (ν_3', ν_4)-arc in $\mathcal{M}$.*

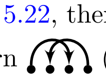


(a) (b)

Figure 16: Illustration of Claim 5.27.

Proof of Claim 5.27. First, according to Corollary 5.6–*(ii)*, there exists either at most one (ν_4, ν'_3)-arc and no (ν'_3, ν_4)-arc, or at most one (ν'_3, ν_4)-arc and no (ν_4, ν'_3)-arc (i, i') in $\mathcal{M}$ (see Figure 16). Now from Claim 5.22, there exists at least one (ν_3, ν'_3)-arc, say (j, j'), in $\mathcal{M}$. Hence, since $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$), $\mathcal{M}$ contains one of the two following labeled patterns:  and . Applying Lemma 5.4 yields the sought-after contradiction. $\square$

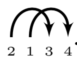
Claim 5.28. *There is at least one (ν_4, ν'_4)-arc in $\mathcal{M}$.*

Proof of Claim 5.28. First, according to Claim 5.22, there is at least one (ν_3, ν'_3)-arc in $\mathcal{M}$. Therefore, since $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$), there is no (ν'_4, ν_4)-arc in $\mathcal{M}$. Now, suppose, aiming at a contradiction, that there is no (ν_4, ν'_4)-arc in $\mathcal{M}$. First, according to Claim 5.24 and Claim 5.25, there is neither a (ν_1, ν_4)-arc nor a (ν_4, ν_1)-arc nor a (ν'_1, ν_4)-arc nor a (ν_2, ν_4)-arc nor a (ν_4, ν_2)-arc nor a (ν'_1, ν_4)-arc nor a (ν_4, ν'_1)-arc nor a (ν_3, ν_4)-arc not a (ν_4, ν_3)-arc nor a (σ', ν_4)-arc nor a (ν_4, σ')-arc nor a (ν_4, ν_4)-arc nor a (ν_4, ν'_2)-abstract nor a (ν'_2, ν_4)-arc in $\mathcal{M}$. Furthermore, according to Claim 5.27, there is neither a (ν_4, ν'_3)-arc nor a $\nu'_3, \nu_4)$-arc in $\mathcal{M}$. But $N_4 > |\pi'| + |\pi''| + |\sigma''|$, and hence $\mathcal{M}$ is not a direct perfect matching, which contradicts our hypothesis about $\mathcal{M}$. $\square$

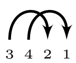
Claim 5.29. *There is neither a (π', π')-arc nor a (σ', π'')-arc nor a (π'', σ')-arc nor a (σ', σ'')-arc nor a (σ'', σ')-arc in $\mathcal{M}$.*

Proof of Claim 5.29. Combine Claim 5.28 together with the fact that $\mathcal{M}$ has Property $\mathbf{P_1}$ and hence is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$). $\square$

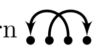
Claim 5.30. *There is neither a (σ', ν'_3)-arc nor a (ν'_3, σ')-arc in $\mathcal{M}$.*

Proof of Claim 5.30. First, according to Claim 5.22, there is at least one (ν_3, ν'_3)-arc in $\mathcal{M}$. Therefore, since $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$), there is no (ν'_3, σ')-arc in $\mathcal{M}$. Now, suppose, aiming at a contradiction, that there is a (σ', ν'_3)-arc in $\mathcal{M}$. Hence, since $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$), $\mathcal{M}$ contains the labeled pattern . Applying Lemma 5.4 yields the sought-after contradiction. □

Claim 5.31. *There is neither a (σ', ν'_4)-arc nor a (ν'_4, σ')-arc in $\mathcal{M}$.*

Proof of Claim 5.31. First, according to Claim 5.28, there is at least one (ν_4, ν'_4)-arc in $\mathcal{M}$. Therefore, since $\mathcal{M}$ is containment-free (*i.e.*, it avoids the unlabeled patterns of $\mathcal{P}_{\text{cont}}$), and avoids  (Property $\mathbf{P_1}$), there is no (ν'_4, σ')-arc in $\mathcal{M}$. Now, suppose, aiming at a contradiction, that there is a (σ', ν'_4)-arc in $\mathcal{M}$. Hence, $\mathcal{M}$ contains the labeled pattern  Applying Lemma 5.4 yields the sought-after contradiction. □

Claim 5.32. *There is no (π', σ')-arc in $\mathcal{M}$.*

Proof of Claim 5.32. Combine Claim 5.28 together with the fact that $\mathcal{M}$ avoids the unlabeled pattern  (Property $\mathbf{P_1}$). □

Combining the above claims, we conclude that there are $k+2$ (σ', π')-arcs in $\mathcal{M}$. Recall that

$$\sigma' = ((k+1)\ \sigma\ (k+2))\ [2N_2 + N_4 + 2n + k + 2] \tag{5.13}$$

and that

$$\pi' = ((n+1)\ \pi\ (n+2))\ [2N_2 + N_4 + n + k + 2]. \tag{5.14}$$

Then it follows we have at least k (possibly $k+1$ or $k+2$) independent (σ', π')-arcs (a, a') in $\mathcal{M}$ with

$$2N_1 + N_2 + N_3 + 1 < a < 2N_1 + N_2 + N_3 + k + 2 \tag{5.15}$$

and

$$2N_1 + N_2 + N_3 + (k+2) + 1 < a' < 2N_1 + N_2 + N_3 + (k+2) + n + 2. \tag{5.16}$$

Therefore, by our hypothesis about $\mathcal{M}$, σ occurs as a pattern in π. □

6. Conclusion and perspectives

There are a number of further directions of investigation in this general subject. They cover several areas: algorithmic, combinatorics, and algebra. Let us mention several — not necessarily new— open problems that are, in our opinion, the most interesting. How many permutations of S_{2n} are squares? How many $(213, 231)$-avoiding permutations of S_{2n} are squares? (Equivalently, by Proposition 3.1, how many binary strings of length $2n$ are squares; see also Problem 4 in [HRS12])? How hard is the problem of deciding whether a $(213, 231)$-avoiding permutation is a square (Problem 4 in [HRS12], see also [BS14, RV13])? Given two permutations π and σ, how hard is the problem of deciding whether σ is a square root of π? As for algebra, one can ask for a complete algebraic study of $\mathbb{Q}[S]$ as a graded associative algebra for the shuffle product $\bullet$. Describing a generating family for $\mathbb{Q}[S]$, defining multiplicative bases of $\mathbb{Q}[S]$, and determining whether $\mathbb{Q}[S]$ is free as an associative algebra are worthwhile questions.

References

[All00] C. Allauzen. Calcul efficace du shuffle de k mots. Technical report, Institut Gaspard Monge, Université Marne-la-Vallée, 2000.

[BBL98] P. Bose, J. F. Buss, and A. Lubiw. Pattern Matching for Permutations. *Inform. Process. Lett.*, 65(5):277–283, 1998.

[BS14] S. Buss and M. Soltys. Unshuffling a square is NP-hard. *J. Comput. Syst. Sci.*, 80(4):766–776, 2014.

[CK97] C. Choffrut and J. Karhumäki. *Combinatorics of Words, in G. Rozenberg and A. Salomaa (eds), Handbook of Formal Languages.* Springer-Verlag, 1997.

[DHT02] G. Duchamp, F. Hivert, and J.-Y. Thibon. Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras. *Int. J. Algebr. Comput.*, 12(5):671–717, 2002.

[EML53] S. Eilenberg and S. Mac Lane. On the groups of $H(\Pi, n)$. I. *Ann. of Math. (2)*, 58:55–106, 1953.

[GR14] D. Grinberg and V. Reiner. Hopf Algebras in Combinatorics. `arXiv:1409.8356[math.CO]`, 2014.

[GV16] S. Giraudo and S. Vialette. Unshuffling Permutations. *Latin American Theoretical Informatics Symposium*, 9644:509–521, 2016.

[HRS12] D. Henshall, N. Rampersad, and J. Shallit. Shuffling and Unshuffling. *Bulletin of the EATCS*, 107:131–142, 2012.

[JR79] S. A. Joni and G.-C. Rota. Coalgebras and bialgebras in combinatorics. *Stud. Appl. Math.*, 61(2):93–139, 1979.

[Man83] A. Mansfield. On the computational complexity of a merge recognition problem. *Discrete Appl. Math.*, 5:119–122, 1983.

[RV13] R. Rizzi and S. Vialette. On Recognizing Words That Are Squares for the Shuffle Product. In A.A Bulatov and A.M. Shur, editors, *8th International Computer Science Symposium in Russia, CSR 2013, Ekaterinburg, Russia*, pages 235–245, 2013.

[Slo] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. `https://oeis.org/`.

[Spe86] J.-C. Spehner. Le calcul rapide des melanges de deux mots. *Theor. Comput. Sci.*, 47:181–203, 1986.

[SS85] R. Simion and F. W. Schmidt. Restricted permutations. *Eur. J. Combin.*, 6(4):383–406, 1985.

[Var14] Y. Vargas. Hopf algebra of permutation pattern functions. *26th International Conference on Formal Power Series and Algebraic Combinatorics*, pages 839–850, 2014.

[vLN82] J. van Leeuwen and M. Nivat. Efficient Recognition of Rational Relations. *Inform. Process. Lett.*, 14(1):34–38, 1982.

[WH84] M. K. Warmuth and D. Haussler. On the complexity of iterated shuffle. *J. Comput. Syst. Sci.*, 28(3):345–358, 1984.