# Ranking Binary Unlabelled Necklaces in Polynomial Time

Duncan Adamson(✉)

Reykjavik University, Reykjavik, Iceland
duncana@ru.is

**Abstract.** Unlabelled Necklaces are an equivalence class of cyclic words under both the rotation (cyclic shift) and the relabelling operations. The relabelling of a word is a bijective mapping from the alphabet to itself. The main result of the paper is the first polynomial-time algorithm for ranking unlabelled necklaces of a binary alphabet. The time-complexity of the algorithm is $O(n^6 \log^2 n)$, where $n$ is the length of the considered necklaces. The key part of the algorithm is to compute the rank of any word with respect to the set of unlabelled necklaces by finding three other ranks: the rank over all necklaces, the rank over symmetric unlabelled necklaces, and the rank over necklaces with an enclosing labelling. The last two concepts are introduced in this paper.

## 1 Introduction

For classes of words under lexicographic (or dictionary) order, a unique integer can be assigned to every word corresponding to the number of words smaller than it. Such an integer is called the *rank* of a word. The *ranking* problem asks to compute the rank of a given word. Ranking has been studied for various objects including partitions [13], permutations [9,10], combinations [12], etc.

The ranking problem is straightforward for the set of all words over a finite alphabet (assuming the standard lexicographic order), however this ceases to be the case once additional symmetry is introduced. One such example is combinatorial necklaces [6]. A *necklace*, also known as a *cyclic word*, is an equivalence class of all words under the cyclic rotation operation, also known as a cyclic shift. Necklaces are classical combinatorial objects and they remain an object of study in other contexts such as total search problems [4] or circular splicing systems [3]. The first class of cyclic words to be ranked were *Lyndon words* - fixed length aperiodic cyclic words - by Kociumaka et al. [7] who provided an $O(n^3)$ time algorithm, where $n$ is the length of the word. An algorithm for ranking necklaces - fixed length cyclic words - was given by Kopparty et al. [8], without tight bounds on the complexity. A quadratic algorithm for ranking necklaces was provided by Sawada et al. [11]. More recently algorithms have been presented for ranking multidimensional necklaces [1] and bracelets [2].

*Our Results.* This paper presents the first polynomial time algorithm for ranking *binary unlabelled necklaces*. Informally, binary unlabelled necklaces can be


Published by Springer Nature Switzerland AG 2022
Y.-S. Han and G. Vaszil (Eds.): DCFS 2022, LNCS 13439, pp. 15–29, 2022.
https://doi.org/10.1007/978-3-031-13257-5_2

though of as necklaces over a binary alphabet with the additional symmetry over the *relabelling* operation, a bijection from the set of symbols to itself. Considered in terms of binary values, the words 0001 and 1110 are equivalent under the relabelling operation, however 1010 and 1100 are not. We provide an $O(n^6 \log^2 n)$ time algorithm for ranking an unlabelled binary necklace within the set of unlabelled binary necklaces of length $n$.

## 2 Preliminaries

Let $\Sigma$ be a finite alphabet. For the remainder of this work we assume $\Sigma$ to be $\{0,1\}$ where $0 < 1$. We denote by $\Sigma^*$ the set of all words over $\Sigma$ and by $\Sigma^n$ the set of all words of length $n$. The notation $\bar{w}$ is used to clearly denote that the variable $\bar{w}$ is a word. The length of a word $\bar{w} \in \Sigma^*$ is denoted by $|\bar{w}|$. We use $\bar{w}_i$, for any $i \in \{1, \dots, |\bar{w}|\}$ to denote the $i^{th}$ symbol of $\bar{w}$. Given two words $\bar{w}, \bar{u} \in \Sigma^*$, the *concatenation operation* is denoted by $\bar{w} : \bar{u}$, returning the word of length $|\bar{w}| + |\bar{u}|$ where $(\bar{w} : \bar{u})_i$ equals either $\bar{w}_i$, if $i \leq |\bar{w}|$ or $\bar{u}_{i-|\bar{w}|}$ if $i > |\bar{w}|$. The $t^{th}$ power of a word $\bar{w}$, denoted by $\bar{w}^t$, equals $\bar{w}$ repeated $t$ times.

Let $[n]$ be the ordered sequence of integers from 1 to $n$ inclusive and let $[i,j]$ be the ordered sequence of integers from $i$ to $j$ inclusive. Given two words $\bar{u}, \bar{v} \in \Sigma^*$, $\bar{u} = \bar{v}$ if and only if $|\bar{u}| = |\bar{v}|$ and $\bar{u}_i = \bar{v}_i$ for every $i \in [|\bar{u}|]$. A word $\bar{u}$ is *lexicographically smaller* than $\bar{v}$ if there exists an $i \in [|\bar{u}|]$ such that $\bar{u}_1\bar{u}_2 \dots \bar{u}_{i-1} = \bar{v}_1\bar{v}_2 \dots \bar{v}_{i-1}$ and $\bar{u}_i < \bar{v}_i$. Given two words $\bar{v}, \bar{w} \in \Sigma^*$ where $|\bar{v}| \neq |\bar{w}|$, $\bar{v}$ is smaller than $\bar{w}$ if $\bar{v}^{|\bar{w}|} < \bar{w}^{|\bar{v}|}$ or $\bar{v}^{|\bar{w}|} = \bar{w}^{|\bar{v}|}$ and $|\bar{v}| < |\bar{w}|$. For a given set of words $\mathbf{S}$, the *rank* of $\bar{v}$ with respect to $\mathbf{S}$ is the number of words in $\mathbf{S}$ that are smaller than $\bar{v}$.

The *subword* of a cyclic word $\bar{w} \in \Sigma^n$ denoted $\bar{w}_{[i,j]}$ is the word $\bar{u}$ of length $n + j - i + 1 \bmod n$ such that $\bar{u}_a = \bar{w}_{i+a \bmod n}$, i.e. the word such that the $a^{th}$ symbol of $\bar{u}$ corresponds to the symbol at position $i + a \bmod n$ of $\bar{w}$. The value of the $t^{th}$ symbol of $\bar{w}_{[i,j]}$ is the value of the symbol at position $i + t - 1$ of $\bar{w}$. By this definition, given $\bar{u} = \bar{w}_{[i,j]}$, the value of $\bar{u}_t$ is the $i + t - 1^{th}$ symbol of $\bar{w}$ and the length of $\bar{u}$ is $|\bar{u}| = j - i + 1$. The notation $\bar{u} \sqsubseteq \bar{w}$ denotes that $\bar{u}$ is a subword of $\bar{w}$. Further, $\bar{u} \sqsubseteq_i \bar{w}$ denotes that $\bar{u}$ is a subword of $\bar{w}$ of length $i$.

The *rotation* of a word $\bar{w} \in \Sigma^n$ by $r \in [0, n-1]$ returns the word $\bar{w}_{[r+1,n]} : \bar{w}_{[1,r]}$, and is denoted by $\langle \bar{w} \rangle_r$, i.e. $\langle \bar{w}_1\bar{w}_2 \dots \bar{w}_n \rangle_r = \bar{w}_{r+1} \dots \bar{w}_n\bar{w}_1 \dots \bar{w}_r$. Under the rotation operation, the word $\bar{u}$ is equivalent to the word $\bar{v}$ if $\bar{v} = \langle \bar{u} \rangle_r$ for some $r$. A word $\bar{w}$ is *periodic* if there is a subword $\bar{u} \sqsubseteq \bar{w}$ and integer $t \geq 2$ such that $\bar{u}^t = \bar{w}$. Equivalently, word $\bar{w}$ is *periodic* if there exists some rotation $0 < r < |\bar{w}|$ where $\bar{w} = \langle \bar{w} \rangle_r$. A word is *aperiodic* if it is not periodic. The *period* of a word $\bar{w}$ is the aperiodic word $\bar{u}$ such that $\bar{w} = \bar{u}^t$.

A *necklace* is an equivalence class of words under the rotation operation. The notation $\tilde{\mathbf{w}}$ is used to denote that the variable $\tilde{\mathbf{w}}$ is a necklace. Given a necklace $\tilde{\mathbf{w}}$, the *canonical representation* of $\tilde{\mathbf{w}}$ is the lexicographically smallest element of the set of words in the equivalence class $\tilde{\mathbf{w}}$. The canonical representation of $\tilde{\mathbf{w}}$ is denoted by $\langle \tilde{\mathbf{w}} \rangle$, and the $r^{th}$ shift of the canonical representation is denoted by $\langle \tilde{\mathbf{w}} \rangle_r$. Given a word $\bar{w}$, $\langle \bar{w} \rangle$ denotes the canonical representation of the necklace containing $\bar{w}$, i.e. the canonical representation of the necklace $\tilde{\mathbf{u}}$ where $\bar{w} \in \tilde{\mathbf{u}}$. The set of necklaces of length $n$ over an alphabet of size $q$ is denoted by $\mathcal{N}_q^n$. Let $\bar{w} \in \mathcal{N}_q^n$ denote that the word $\bar{w}$ is the canonical representation of some necklace $\tilde{\mathbf{w}} \in \mathcal{N}_q^n$. An aperiodic necklace, known as a *Lyndon word*, is a necklace representing the equivalence class of some aperiodic word. Note that if a word is aperiodic, then every rotation of the word is also aperiodic. The set of Lyndon words of length $n$ over an alphabet of size $q$ is denoted by $\mathcal{L}_q^n$.

As both necklaces and Lyndon words are classical objects, there are many fundamental results regarding each objects. The first results for these objects were equations determining the number of necklaces or Lyndon words of a given length. The number of (1D) necklaces is given by the equation $|\mathcal{N}_q^n| = \frac{1}{n}\sum_{d|n} \phi\left(\frac{n}{d}\right) q^d$ where $\phi(n)$ is Euler's totient function [6]. Similarly the number of Lyndon words is given with the equation $|\mathcal{L}_q^n| = \sum_{d|n} \mu\left(\frac{n}{d}\right) |\mathcal{N}_q^d|$, where $\mu(x)$ is the Möbius function [6]. The *rank* of a word $\bar{w}$ in the set of necklaces $\mathcal{N}_q^n$ is the number of necklaces with a canonical representation smaller than $\bar{w}$.

## 2.1 Unlabelled Necklaces

An *unlabelled necklace* is a generalisation of the set of necklaces. At a high level, two words $\bar{v}, \bar{u} \in \Sigma^n$ belong to the same unlabelled necklace class $\tilde{\mathbf{w}}$ if there exists some labelling function $\psi(x) : \Sigma \mapsto \Sigma$ and rotation $r \in [n]$ such that $(\langle \bar{v} \rangle_r)_i = \psi(\bar{u}_i)$ for every $i \in [n]$. More formally, let $\psi(x)$ be a bijection from $\Sigma$ into $\Sigma$, i.e. a function taking as input some symbol in $\Sigma$ and returning a symbol in $\Sigma$ such that $\{\psi(x) | \forall x \in \Sigma\} = \Sigma$. For notation $\psi(\bar{w})$ is used to denote the word constructed by applying $\psi(x)$ to every symbol in $\bar{w}$ in order, formally $\psi(\bar{w}) = \psi(\bar{w}_1)\psi(\bar{w}_2)\dots\psi(\bar{w}_n)$. Similarly, the notation $\psi(\tilde{\mathbf{w}})$ is used to denote the necklace class constructed by applying $\psi(\bar{w})$ to every word $\bar{w} \in \tilde{\mathbf{w}}$. Further, let $\Psi(\Sigma)$ be the set of all such functions. The unlabelled necklace $\tilde{\mathbf{w}}$ with a canonical representation $\bar{w}$ contains every word $\bar{v} \in \Sigma^n$ where $\psi(\langle \bar{v} \rangle_r) = \bar{w}$ for some $\psi(x) \in \Psi(\Sigma)$ and $r \in [n]$. As in the labelled case, the canonical representation of an unlabelled necklace $\tilde{\mathbf{w}}$, denoted $\langle \tilde{\mathbf{w}} \rangle$, is the lexicographically smallest word in the equivalence class. The set of unlabelled $q$-arry necklaces of length $n$ is denoted $\hat{\mathcal{N}}_q^n$, and the set of $q$-arry Lyndon words of length $n$ $\hat{\mathcal{L}}_q^n$.

In this paper we study *binary unlabelled necklaces*, in other words unlabelled necklaces restricted to a binary alphabet. In this case $\Sigma = \{0, 1\}$ and $\Psi(\Sigma)$ contains the identity function $I(x)$, where $I(x) = x$, and the swapping function $S(x)$ where $S(x) = \begin{cases} 0 & x = 1 \\ 1 & x = 0 \end{cases}$. Gilbert and Riordan [5] provide the following equations for computing the sizes of $\hat{\mathcal{N}}_2^n$ and $\hat{\mathcal{L}}_2^n$:

$$|\hat{\mathcal{N}}_2^n| = \sum_{\text{odd } d|n} \phi(d)2^{n/d}$$
$$|\hat{\mathcal{L}}_2^n| = \sum_{\text{odd } d|n} \mu(d)2^{n/d}$$

In this paper we introduce two subclasses of unlabelled necklaces, the class of *symmetric unlabelled necklaces* and the class of *enclosing unlabelled necklaces* for some given word $\bar{w}$. Observe that a binary unlabelled necklace $\tilde{\mathbf{w}}$ may correspond to either one or two (labelled) necklaces. Informally, a symmetric unlabelled necklaces is such an unlabelled necklaces that corresponds to only a single necklace. An enclosing unlabelled necklace relative to a word $\bar{w}$ is a non-symteric unlabelled necklace corresponding to a pair of necklaces $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{u}}$ such that $\tilde{\mathbf{v}} < \bar{w} < \tilde{\mathbf{u}}$. Any Lyndon word that is a symmetric unlabelled necklace is a *symmetric unlabelled Lyndon word*, and any unlabelled Lyndon word that encloses a word $\bar{w}$ is an *enclosing unlabelled Lyndon word* of $\bar{w}$.

**Definition 1 (Symmetric Necklaces).** *A binary necklace* $\tilde{\mathbf{w}}$ *is* symmetric *if and only if* $\tilde{\mathbf{w}} = S(\tilde{\mathbf{w}})$*.*

**Definition 2 (Enclosing Unlabelled Necklaces).** *An unlabelled necklace* $\tilde{\mathbf{u}}$ encloses *a word* $\bar{w}$ *if* $\langle\tilde{\mathbf{u}}\rangle < \bar{w} < \langle S(\tilde{\mathbf{u}})\rangle$*. An unlabelled necklace* $\tilde{\mathbf{u}}$ *is an* enclosing unlabelled necklace *of* $\bar{w}$ *if* $\tilde{\mathbf{u}}$ *encloses* $\bar{w}$*.*

### 2.2 Bounding Subwords

One important tool that is used in the ranking of unlabelled necklaces are *bounding subwords*, introduced in [2]. Informally, bounding subwords of length $l \leq n$ provide a means to partition $\Sigma^l$ into $n+2$ sets based on the subwords of some $\bar{w} \in \Sigma^n$ of length $l$. Given two subwords $\bar{v}, \bar{u} \sqsubseteq_l \bar{w}$ such that $\bar{v} < \bar{u}$ the set $S(\bar{v}, \bar{u})$ contains all words in $\Sigma^l$ that are between the value of $\bar{v}$ and $\bar{u}$, formally $S(\bar{v}, \bar{u}) = \{\bar{x} \in \Sigma^l | \bar{v} \leq \bar{x} < \bar{u}\}$. In this paper we are only interested in sets between pairs $\bar{v}, \bar{u} \sqsubseteq_l \bar{w}$ where there exists no $\bar{s} \sqsubseteq_l \bar{w}$ such that $\bar{v} < \bar{s} < \bar{u}$. As such, we define a subword of $\bar{w}$ as *bounding* some word $\bar{v}$ if it is the lexicographically largest subword of $\bar{w}$ that is smaller than $\bar{v}$.

**Definition 3 (Bounding Subwords).** *Let* $\bar{w}, \bar{v} \in \Sigma^*$ *where* $|\bar{w}| \leq |\bar{v}|$*. The word* $\bar{w}$ *is* bounded *(resp. strictly bounded) by* $\bar{s} \sqsubseteq_{|\bar{w}|} \bar{v}$ *if* $\bar{s} \leq \bar{w}$ *(resp.* $\bar{s} < \bar{w}$*) and there is no* $\bar{u} \sqsubseteq_{|\bar{w}|} \bar{v}$ *such that* $\bar{s} < \bar{u} \leq \bar{w}$*.*

**Proposition 1 ([2]).** *Let* $\bar{v} \in \Sigma^n$*. The array* $WX[\bar{s} \sqsubseteq \bar{v}, x \in \Sigma]$*, such that* $WX[\bar{s}, x]$ *strictly bounds* $\bar{w} : x$ *for every* $\bar{w}$ *strictly bounded by* $\bar{s}$*, can be computed in* $O(k \cdot n^3 \cdot \log(n))$ *time where* $|\Sigma| = k$*.*

For the remainder of this paper, we can assume that the array $WX$ has been precomputed for every $\bar{s} \sqsubseteq \bar{v}, x \in \Sigma$. Note that in our case $k = 2$, therefore the process of computing $WX$ requires only $O(n^3 \cdot \log(n))$ time.

## 3 Ranking

In this section we present our ranking algorithm. For the remainder of this section, we assume that we are ranking the word $\bar{w}$ that is the canonical representation of the binary unlabelled necklace $\tilde{\mathbf{w}}$. We first provide an overview of the main idea behind our ranking algorithm.

**Theorem 1.** *Let $RankAN(\bar{w}, m)$ be the rank of the word $\bar{w} \in \Sigma^n$ within the set of non-symmetric unlabelled necklaces of length $n$ that do not enclose $\bar{w}$, let $RankSN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of symmetric necklaces of length $m$ and let $RankEN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of necklaces of length $m$ that enclose $\bar{w}$. The rank of any necklace $\tilde{\mathbf{w}}$ represented by the word $\bar{w}$ within the set of binary unlabelled necklaces of length $m$ is given by $RankAN(\bar{w}, m) + RankSN(\bar{w}, m) + RankEN(\bar{w}, m)$. Further the rank can be found in $O(n^6 \log^2 n)$ time for any $m \leq n$.*

*Proof.* Observe that every unlabelled necklace must be one of the above classes. Therefore the rank of $\bar{w}$ within the set of all binary unlabelled necklaces of length $m$ is given by $RankAN(\bar{w}, m) + RankSN(\bar{w}, m) + RankEN(\bar{w}, m)$. Lemma 1 shows that the rank of $\bar{w}$ within the set of non-symmetric unlabelled necklaces of length $m$ that do not enclose $\bar{w}$ can be found in $O(n^6 \log^2(n))$ time. Theorem 2 shows that the rank of $\bar{w}$ within the set of symmetric necklaces can be found in $O(n^6 \log^2 n)$ time. Theorem 3 shows that the rank of $\bar{w}$ within the set of necklaces enclosing $\bar{w}$ can be found in $O(n^6 \log n)$ time.

**Lemma 1.** *Let $RankAN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of non-symmetric unlabelled necklaces of length $m$ that do not enclose $\bar{w}$, and let $RankN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of all necklaces of length $m$. Then $RankAN(\bar{w}, m) = (RankN(\bar{w}, m) - RankSN(\bar{w}, m) - RankEN(\bar{w}, m))/2$. Further, this rank can be found in $O(n^6 \log^2 n)$ time for any $m \leq n$.*

*Proof.* Note that any asymmetric unlabelled necklace appears exactly twice in the set of necklaces smaller than $\bar{w}$. Further, any enclosing or symmetric necklace appears exactly once in the same set. Therefore $RankAN(\bar{w}, m) = \frac{RankN(\bar{w},m) - RankSN(\bar{w},m) - RankEN(\bar{w},m)}{2}$. As the value of $RankN(\bar{w}, m)$ can be found in $O(n^2)$ time using the algorithm due to Sawada and Williams [11], the value of $RankSN(\bar{w}, m)$ found in $O(n^6 \log^2 n)$ time from Theorem 2, and the of $RankEN(\bar{w}, m)$ found in $O(n^6 \log n)$ time from Theorem 3, the total time complexity is $O(n^6 \log^2 n)$.

## 4 Symmetric Necklaces

In this section we show how to rank a word $\bar{w}$ within the set of symmetric necklaces of length $m$. Before presenting our computational tools, we first introduce the key theoretical results that form the basis for our ranking approach. The key observation is that any symmetric necklace $\tilde{\mathbf{v}}$ must have a period of length $2 \cdot r$ where $r$ is the smallest rotation such that $\langle \tilde{\mathbf{v}} \rangle_r = S(\langle \tilde{\mathbf{v}} \rangle)$. This is formally proven in Proposition 2, and restated in Observation 1 in terms of Lyndon words.

**Proposition 2.** *A necklace* $\tilde{\mathbf{w}}$ *represented by the word* $\bar{w} \in \Sigma^n$ *is symmetric if and only if there exists some* $r \in [n]$ *s.t.* $\bar{w}_i = S(\bar{w}_{i+r \bmod n})$ *for every* $i \in [n]$*. Further, the period of* $\bar{w}$ *equals* $2 \cdot r$ *where* $r \in [n]$ *is the smallest rotation such that* $\langle \bar{w} \rangle_r = S(\bar{w})$*.*

*Proof.* As $\tilde{\mathbf{w}}$ is symmetric, $S(\bar{w})$ must belong to the necklace class $\tilde{\mathbf{w}}$. Therefore, there must be some rotation $r$ such that $\langle \bar{w} \rangle_r = S(\bar{w})$. We now claim that $r \leq \frac{n}{2}$. Assume for the sake of contradiction that $r > \frac{n}{2}$. Then $\bar{w}_i = S(\bar{w}_{i+r \bmod n}) = \bar{w}_{i+2r \bmod n} = \ldots = \bar{w}_{i+2 \cdot k \cdot r \bmod n} = S(\bar{w}_{i+(2 \cdot k+1)r \bmod n})$. As $r > \frac{n}{2}$ this sequence must imply that either $\bar{w}_i = S(\bar{w}_i)$, an obvious contradiction, or that there exists some smaller value $p = GCD(n, r) \leq \frac{n}{2}$ such that $\bar{w}_i = S(\bar{w}_{i+p \bmod n})$. Further, $\bar{w}$ must have a period of at most $2 \cdot r$.

Assume now that $r$ is the smallest rotation such that $\langle \bar{w} \rangle_r = S(\bar{w})$ and for the sake of contradiction further assume that the period of $\bar{w}$ is $p < r$. Then, as $\bar{w}_i = \bar{w}_{i+p \bmod n}$ for every $i \in [n]$, $\bar{w}_{i+r \bmod n} = \bar{w}_{i+r-p \bmod n}$, hence $\bar{w}_i = S(\bar{w}_{i+r-p \bmod n})$, contradicting the initial assumption. The period can not be equal to the value of $r$ as by definition $\bar{w}_i = S(\bar{w}_{i+r \bmod n})$. Assume now that the period $p$ of $\bar{w}$ is between $r$ and $2 \cdot r$. As $\bar{w}_i = \bar{w}_{i+c \cdot p+2k \cdot r \bmod n}$ for every $c, k \in \mathbb{N}$ and $i \in [n]$. Further both $r$ and $p$ must be less than $\frac{n}{2}$. Therefore $\bar{w}_i = \bar{w}_{i+((n/p)-1)p+2 \cdot r \bmod n} = i + 2 \cdot r - p \bmod n$ and hence $\bar{w}$ is periodic in $2 \cdot r - p$. As $p > r, 2 \cdot r - p < r$, however as no such period can exist, this leads to a contradiction. Therefore, $2 \cdot r$ is the smallest period of $\bar{w}$.

**Lemma 2.** *Let* $\mathbf{RA}(\bar{w}, m, S, r)$ *contain the set of words belonging to an symmetric necklace smaller than* $\bar{w}$ *such that* $\bar{v}_i = S(\bar{v}_{i+r \bmod m})$ *for every* $\bar{v} \in \mathbf{RA}(\bar{w}, m, S, r)$*. Further let* $\mathbf{RB}(\bar{w}, m, S, r) \subseteq \mathbf{RA}(\bar{w}, m, S, r)$ *contain the set of words belonging to an symmetric Lyndon word smaller than* $\bar{w}$ *such that* $r$ *is the smallest value for which* $\bar{v}_i = S(\bar{v}_{i+r \bmod m})$ *for every* $\bar{v} \in \mathbf{RB}(\bar{w}, m, S, r)$*. The size of* $\mathbf{RB}(\bar{w}, m, S, r)$ *is given by:*

$$|\mathbf{RB}(\bar{w}, m, S, r)| = \sum_{p|r} \mu\left(\frac{m}{p}\right) |\mathbf{RA}(\bar{w}, m, S, p)|$$

*Proof.* Observe that every word in $\mathbf{RA}(\bar{w}, m, S, r)$ must have a unique period which is a factor of $2 \cdot r$. Therefore, the size of $\mathbf{RA}(\bar{w}, m, S, r)$ can be expressed as $\sum_{d|r} |\mathbf{RB}(\bar{w}, m, S, r)|$. Applying the Möbius inversion formula to this equation gives $|\mathbf{RB}(\bar{w}, m, S, r)| = \sum_{p|r} \mu\left(\frac{m}{p}\right) |\mathbf{RA}(\bar{w}, m, S, p)|$.

**Observation 1.** *Observe that any symmetric Lyndon word* $\tilde{\mathbf{v}}$ *must have length* $2 \cdot r$*, where* $r$ *is the smallest rotation such that* $\langle \bar{v} \rangle_r = S(\langle \bar{v} \rangle)$*.*

**Lemma 3.** *Let* $RankSL(\bar{w}, 2 \cdot r)$ *be the rank of* $\bar{w}$ *within the set of symmetric Lyndon words of length* $2 \cdot r$*. The value of* $RankSL(\bar{w}, r)$ *is given by* $\frac{|\mathbf{RB}(\bar{w}, 2 \cdot r, S, r)|}{2 \cdot r}$*.*

*Proof.* Observe that any symmetric Lyndon word has exactly $2 \cdot r$ unique translations. Further, as any word in $\mathbf{RB}(\bar{w}, 2 \cdot r, S, r)$ must correspond to an aperiodic word, following Observation 1, the size of $\mathbf{RB}(\bar{w}, 2 \cdot r, S, r)$ can be used to find $RankSL(\bar{w}, 2\dot{r})$ by dividing the cardinality of $\mathbf{RB}(\bar{w}, 2 \cdot r, S, r)$ by $2 \cdot r$.

**Lemma 4.** *Let $RankSN(\bar{w}, m, r)$ be the rank of $\bar{w}$ within the set of symmetric necklaces of length $m$ such that for each such necklace $\tilde{\mathbf{v}}$, $r$ is the smallest rotation such that $\langle \tilde{\mathbf{v}} \rangle_r = S(\langle \tilde{\mathbf{v}} \rangle)$. The value of $RankSN(\bar{w}, m, r)$ is given by $\sum\limits_{d|2r} RankSL(\bar{w}, d)$.*

*Proof.* Following the same arguments as in Lemma 2, observe that every necklace counted by $RankSN(\bar{w}, m, r)$ must have a period that is a factor of $2 \cdot r$. Therefore, the value of $RankSN(\bar{w}, m, r)$ is given by $\sum\limits_{d|2r} RankSL(\bar{w}, d)$.

**Lemma 5.** *Let $RankSN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of symmetric necklaces of length $m$ and let $RankSN(\bar{w}, m, r)$ be the rank of $\bar{w}$ within the set of symmetric necklaces of length $m$ such that for each such necklace $\tilde{\mathbf{v}}$, $r$ is the smallest rotation such that $\langle \tilde{\mathbf{v}} \rangle_r = S(\langle \tilde{\mathbf{v}} \rangle)$. The value of $RankSN(\bar{w}, m)$ is given by $\sum\limits_{r|(m/2)} RankSN(\bar{w}, m, r)$.*

*Proof.* Observe that every necklace counted by $RankSN(\bar{w}, m)$ must have a unique translation that is the minimal translation under which it is symmetric. Further this translation must be a factor of $\frac{m}{2}$. Therefore $RankSN(\bar{w}, m) = \sum\limits_{r|(m/2)} RankSN(\bar{w}, m, r)$.

Following Lemmas 2, 3, 4, and 5 the main challenge in computing $RankSN(\bar{w}, m)$ is computing the size of $\mathbf{RA}(\bar{w}, m, S, r)$. In order to do so, $\mathbf{RA}(\bar{w}, m, S, r)$ is partitioned into two sets, $\alpha(\bar{w}, r, j)$ and $\beta(\bar{w}, r, j)$ where $j \in [r]$. Let $\bar{v}$ be some arbitrary word in the set $\mathbf{RA}(\bar{w}, m, S, r)$. The set $\alpha(\bar{w}, r, j)$ contains the word $\bar{v}$ if $j$ is the smallest rotation under which $\langle \bar{v} \rangle_j \leq \bar{w}$. The set $\beta(\bar{w}, r, j)$ contains $\bar{v}$ if $j$ is the smallest rotation under which $\langle \bar{v} \rangle_j \leq \bar{w}$ and $\langle \bar{v} \rangle_t > \bar{w}$ for every $t \in [r + 1, 2 \cdot r]$. Note that by this definition, $\beta(\bar{w}, r, j) \subseteq \alpha(\bar{w}, r, j)$.

**Observation 2.** *Given any word $\bar{v} \in \mathbf{RA}(\bar{w}, m, S, r)$ such that $\bar{v} \notin \alpha(\bar{w}, r, j)$ for any $j \in [r]$, there exists some $j' \in [r]$ for which $\langle \bar{v} \rangle_r \in \beta(\bar{w}, r, j')$.*

*Proof.* As $\bar{v} \in \mathbf{RA}(\bar{w}, m, S, r)$, there must be some rotation $t$ such that $\langle \bar{v} \rangle_t < \bar{w}$. As $\bar{v} \notin \alpha(\bar{w}, r, j)$, $t$ must be greater than $r$. Therefore, $\langle \bar{v} \rangle_r$ must belong to $\beta(\bar{w}, r, t - r)$ confirming the observation.

**Observation 3.** *For any $\bar{v} \in \beta(\bar{w}, r, j)$, $\langle \bar{v} \rangle_r \notin \alpha(\bar{w}, r, j')$ for any $j' \in [r]$.*

*Proof.* As $\bar{v} \in \beta(\bar{w}, r, j)$, for any rotation $t > r, \langle \bar{v} \rangle_t > \bar{w}$. Therefore $\langle \bar{v} \rangle_t \notin \alpha(\bar{w}, r, j')$ for any $j' \in [r]$.

Combining Observations 2 and 3, the size of $\mathbf{RA}(\bar{w}, m, S, r)$ can be given in terms of the sets $\alpha(\bar{w}, r, j)$ and $\beta(\bar{w}, r, j)$ as $\sum_{j \in [r]} |\alpha(\bar{w}, r, j)| + |\beta(\bar{w}, r, j)|$. The remainder of this section is laid out as follows. We first provide a high level overview of how to compute the size of $\alpha(\bar{w}, r, j)$. Then we provide a high level overview on computing the size of $\beta(\bar{w}, r, j)$. Finally, we state Theorem 2, summarising the main contribution of this section and showing that $RankSN(\bar{w}, m)$ can be computed in at most $O(n^6 \log^2 n)$ time.

*Computing the size of* $\alpha(\bar{w}, r, j)$*.* We begin with a formal definition of $\alpha(\bar{w}, r, j)$. Let $\alpha(\bar{w}, r, j) \subseteq \mathbf{RA}(\bar{w}, m, S, r)$ be the subset of words in $\mathbf{RA}(\bar{w}, m, S, r)$ such that for every word $\bar{v} \in \alpha(\bar{w}, r, j)$, $j$ is the smallest rotation for which $\langle \bar{v} \rangle_j \leq \bar{w}$. Note that if $j$ is the smallest rotation such that $\langle \bar{v} \rangle_j \leq \bar{w}$, the first $j$ symbols of $\bar{v}$ must be such that for every $j' \in [j-1], \bar{v}_{[j', 2r]} > \bar{w}$. Let $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r) \subseteq \alpha(\bar{w}, r, j)$ be the set of words of length $2 \cdot r$ such that every word $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$:

1. $\langle \bar{v} \rangle_s > \bar{w}$ for every $s \in [j-1]$.
2. $\langle \bar{v} \rangle_j < \bar{w}$.
3. $\bar{v}_{[1,r]} = S(\bar{v}_{[r+1, 2 \cdot r]})$.
4. The subword $\bar{v}_{[r+1, r+i]}$ is strictly bound by $\bar{B} \sqsubseteq_i \bar{w}$.
5. The subword $\bar{v}_{[i-p, i]} = \bar{w}_{[1,p]}$.

Rather than computing the size of $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ directly, we are instead interested in the number of unique suffixes of length $r - i$ of the words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$. Note that as every word in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ belongs to a symmetric necklace, the number of possible suffixes on length $r - i$ of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ equals the number of unique subwords of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ between position $i + 1$ and $r$. Let $SA(\bar{w}, p, \bar{B}, i, j, r)$ be a function returning the number of unique suffixes of length $r - i$ of the words within $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$. The value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is computed in a dynamic manner relaying on a key structural proposition regarding $\mathbf{A}(\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r))$.

**Proposition 3.** *Given* $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$*, such that* $\bar{v}_{[i-s, i+1]} \geq \bar{w}_{[1,s]}$ *for every* $s \in [i]$*,* $\bar{v}$ *also belongs to* $\mathbf{A}(\bar{w}, p', WX[\bar{B}, \bar{v}_{i+1}], i+1, j, r)$ *where* $p' = p + 1$ *if* $\bar{v}_{i+1} = \bar{w}_{p+1}$ *and* $0$ *otherwise.*

*Proof.* By definition, if $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ then there must exists some $p' \in [i+1]$, and $\bar{B} \sqsubseteq_i \bar{w}$ such that $\bar{v} \in \mathbf{A}(\bar{w}, p', \bar{B}', i, j, r)$. From Proposition 1, the value of $\bar{B}' = WX[\bar{B}, S(\bar{v}_{i+1})]$. Further $\bar{v}_{i+1} \geq \bar{w}_{p+1}$ as otherwise $\bar{v}_{[i-p, i+1]} < \bar{w}_{[1, p+1]}$, contradicting the original assumption. If $\bar{v}_{i+1} = \bar{w}_{p+1}$ then $p' = p + 1$ by definition. Otherwise $p' = 0$ as $\bar{v}_{[i-s, i+1]} > \bar{w}_{[1, s+1]}$.

**Corollary 1.** *Let* $\bar{v}, \bar{u} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ *be a pair of words and let* $\bar{v}' = \bar{u}_{[1,i]} : \bar{v}_{[i+1,r]} : S(\bar{v}_{[1,i]} : \bar{u}_{[i+1,r]})$*. Then* $\bar{v}' \in \mathbf{A}(\bar{w}, p', WX[\bar{B}, \bar{v}_{i+1}], i+1, j, r)$ *if and only if* $\bar{v} \in \mathbf{A}(\bar{w}, p', WX[\bar{B}, \bar{v}_{i+1}], i+1, j, r)$*.*

Proposition 3 and Corollary 1 provide the basis for computing the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$. This is done by considering 4 cases based on the value of $i$ relative to the values of $j$ and $r$ which we will sketch bellow. The key observation behind this partition is that the value of the symbol at position $i+1$ is restricted differently depending on the values of $i, j, r$ and $p$.

If $i < j$, then $\bar{v}_{i+1}$ must be greater than or equal to $\bar{w}_{p+1}$ for every $v \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$, to avoid a contradiction caused by there being a rotation smaller than $j$ for which $\bar{v}$ is smaller than $\bar{w}$. This gives two cases. If $\bar{w}_p = 1$ then the only possible value of $\bar{v}_{i+1}$ is 1 and therefore the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is equal to the value $SA(\bar{w}, p+1, WX[\bar{B}, 0], i+1, j, r)$. Alternatively, if $\bar{w}_{p+1} = 0$, then $\bar{v}_{i+1}$ can be either 0 or 1. The number of suffixes of length $r-i$ of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ where the symbol at position $i+1$ is 0 equals the value of $SA(\bar{w}, p+1, WX[\bar{B}, 1], i+1, j, r)$. The number of suffixes of length $r-i$ of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ where the symbol at position $i+1$ is 1 equals the value of $SA(\bar{w}, 0, WX[\bar{B}, 0], i+1, j, r)$. Therefore, if $i < j$ and $\bar{w}_{p+1} = 0$, the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is $SA(\bar{w}, p+1, WX[\bar{B}, 1], i+1, j, r) + SA(\bar{w}, 0, WX[\bar{B}, 0], i+1, j, r)$.

If $i = j$ then the value of $\bar{v}_{i+1}$ depends on the value of $p$ for every $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$. In order for $j$ to be the smallest rotation for which $\bar{v}$ is smaller than $\bar{w}$, the value of $p$ must be 0, as otherwise the rotation by $j-p$ would be a smaller rotation for which $\bar{v}$ is smaller than $\bar{w}$. Hence, if $p > 0$, $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r) = \emptyset$ and by extension $SA(\bar{w}, p, \bar{B}, i, j, r) = 0$. If $p = 0$ and $i = j$, then the value of $\bar{v}_{i+1}$ must be 0, as otherwise the rotation by $r$ leads to a word that is greater than $\bar{w}$. Therefore, when $p = 0$ and $i = r$, the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is exactly equal to the value $SA(\bar{w}, 1, WX[\bar{B}, 1], i+1, j, r)$ of length $r-i-1$.

If $j < i < r$ and $p < i - j$, then the rotation of $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ by $j$ leads to a word smaller than $\bar{w}$ regardless of the value of $\bar{v}_{i+1}$, and hence $SA(\bar{w}, p, \bar{B}, i, j, r) = 2^{r-i}$, corresponding to the set of all possible words of length $i-r$ over the binary alphabet. If $j < i < r$ and $p = i - j$, then the symbol at position $i+1$ must be less than or equal to $\bar{w}_{p+1}$, therefore the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ of length $r-i$ is determined by the value of $\bar{w}_{p+1}$. If $\bar{w}_{p+1} = 0$ then the value of $\bar{v}_{i+1}$ must be 0 to avoid a contradiction, and hence the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ equals the value of $SA(\bar{w}, p+1, WX[\bar{B}, 1], i+1, j, r)$. Otherwise, if $\bar{w}_{p+1} = 1$ then the value of $\bar{v}_{i+1}$ can be either 0 or 1. Any word in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ where the symbol at position $i+1$ is 0 will be smaller than $\bar{w}$ after being rotated by $j$ regardless of the value of the symbols at position $i+2$ to $r$. Therefore, the number of suffixes of length $r-i$ of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ where the symbol at position $i+1$ is 0 is $2^{r-i-2}$. Further, the number of suffixes of length $r-i$ of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ where the symbol at position $i+1$ is 1 is equal to the value of $SA(\bar{w}, p+1, WX[\bar{B}, 0], i+1, j, r)$.

Finally, if $i = r$ then the number of unique zero length suffixes of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ is determined by the value of $p$ and $\bar{B}$. If $p < i - j$, then for every $\bar{v} \in \mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$, the rotation of $\bar{v}$ by $j$ is less than $\bar{w}$ regardless of the value of $\bar{B}$. Therefore the number of possible suffixes of length 0 of words in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ is 1, representing the empty word. On the other hand, if

$p = i - j$, then the number of possible suffixes of length 0 can be determined by the value of $\bar{B}$. Note that the rotation of any word in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ by $j$ is less than $\bar{w}$ if and only if $\bar{w}_{[1,p]} : \bar{B} < \bar{w}_{[1,p+r]}$. Therefore the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is either 1, if $\bar{w}_{[1,p]} : \bar{B} < \bar{w}_{[1,p+r]}$, or 0 otherwise.

**Lemma 6.** *The value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ can be computed in $O(n^3)$ time.*

*Proof (Sketch).* Following the outline given above, the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is computed in a dynamic manner, starting with $i = r$ as a base case, and progressing in descending value of $i$. For each value of $i$, the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ is computed for every $\bar{B} \sqsubseteq_i \bar{w}$, and $p \in [1, i]$ if $i \leq j$, or $p = i - j$ if $i > j$. For $i = r$, the value of $SA(\bar{w}, i - j, \bar{B}, i, j, r)$ can be computed in $O(n)$ time for every $\bar{B} \sqsubseteq_i \bar{w}$. For $i < r$ the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ can be computed in $O(1)$ time provided the value of $SA(\bar{w}, p', \bar{B}', i + 1, j, r)$ has been precomputed for every $p' \in \{p + 1, 0\}$ and $\bar{B}' \sqsubseteq_{i+1} \bar{w}$. As there are only $n$ values of $\bar{B} \sqsubseteq_r \bar{w}$ to consider in the base case, and at most $O(n^3)$ total possible value of $i, p \in [r], \bar{B} \sqsubseteq_i \bar{w}$, the total complexity of this process is $O(n^3)$.

**Lemma 7.** *The size of $\alpha(\bar{w}, j, r)$ can be computed in $O(n^4)$ time.*

*Proof.* From Lemma 6, the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$ can be computed in $O(n^3)$ time for any value of $i, p \in [n]$ and $\bar{B} \sqsubseteq_i \bar{w}$. Note that $SA(\bar{w}, 0, \emptyset, 0, j, r)$ allows us to count the number of words $\bar{v} \in \alpha(\bar{w}, j, r)$ where $\bar{v}_{[r+1,r+i]} \not\sqsubseteq \bar{w}$ for every $i \in [r]$, or equivalently, where $S(\bar{v}_{[1,i]}) \not\sqsubseteq \bar{w}$. To compute the remaining words, let $\bar{u} \sqsubseteq_{i-1} \bar{w}$ and let $x \in \{0, 1\}$ be a symbol such that $\bar{u} : x \not\sqsubseteq \bar{w}$. Further let $\bar{B} \sqsubseteq_i \bar{w}$ be the subword of $\bar{w}$ strictly bounding $\bar{u} : x$ and let $p$ be the length of the longest suffix of $S(\bar{u} : x)$ that is a prefix of $\bar{w}$, i.e. the largest value such that $S(\bar{u} : x)_{[i-p:i]} = \bar{w}_{[1,p]}$. Observe that $S(\bar{u} : x)_{[1,p]}$ is the prefix of some word $\bar{v} \in \alpha(\bar{w}, j, r)$ if and only if one of the following holds:

- if $i < r$ then $(\bar{u} : x)_{[i-s,i]} > \bar{w}_{[1,s]}$ for every $s \in [p + 1, i]$.
- if $i = r$ then $p = 0$.
- if $i > r$ then $p = i - r$.

As each condition can be checked in at most $O(n)$ time, and there are at most $O(n^2)$ subwords of $\bar{w}$, it is possible to check for every such subword if it is a prefix of some word in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ in $O(n^3)$ time. Following Corollary 1, the number of suffixes of each word in $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$ is equal to the value of $SA(\bar{w}, p, \bar{B}, i, j, r)$. By precomputing $SA(\bar{w}, p, \bar{B}, i, j, r)$, the number of words in $\alpha(\bar{w}, j, r)$ with $\bar{u} : x$ as a prefix can be computed in $O(1)$ time. Therefore the total complexity of computing the size of $\alpha(\bar{w}, j, r)$ is $O(n^3)$.

*Computing the Size of $\beta(\bar{w}, r, j)$.* We start by subdividing $\beta(\bar{w}, r, j)$ into the subsets $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$. Let $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r) \subseteq \beta(\bar{w}, r, j)$ be the subset of $\beta(\bar{w}, r, j)$ containing every word $\bar{v} \in \beta(\bar{w}, r, j)$ where $\bar{v}$ satisfies:

1. $\bar{v}_{[1,r]} = S(\bar{v}_{[r+1,2\cdot r]})$.
2. The first $i$ symbols of $\bar{v}$ are strictly bound by $\bar{B}_f \sqsubseteq_i \bar{w}$ ($\bar{B}_f$ standing for **b**ounding the **f**ront).

3. The subword $\bar{v}_{[r+1,r+i]}$ is strictly bound by $\bar{B}_b \sqsubseteq_i \bar{w}$ ($\bar{B}_b$ standing for **b**ounding the **b**ack).
4. The subword $\bar{v}_{[i-p_f,i]} = \bar{w}_{[p_f]}$ ($p_f$ standing for the **f**ront **p**refix).
5. the subword $\bar{v}_{[r+i-p_b,r+i]} = \bar{w}_{[p_b]}$ ($p_b$ standing for the **b**ack **p**refix).

**Proposition 4.** *Given $\bar{v} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$, where $\bar{v}_{[i-s,i+1]} \geq \bar{w}_{[1,s]}$ for every $s \in [i]$, $\bar{v}$ also belongs to $\mathbf{B}(\bar{w}, p'_f, p'_b, XW[\bar{B}_f, \bar{v}_{i+1}], XW[\bar{B}_b, S(\bar{v}_{i+1})], i, j, r)$ where $p'_f = p_f + 1$ if $\bar{v}_{i+1} = w_{p_f+1}$ or 0 otherwise, and $p'_b = p_b + 1$ if $S(\bar{v}_{i+1}) = \bar{w}_{p_b+1}$, and 0 otherwise.*

*Proof.* Following the same arguments as Proposition 3, observe that $\bar{v}_{[1,i+1]}$ is bound by $XW[\bar{B}_f, \bar{v}_{i+1}]$ and $S(\bar{v}_{[1,i+1]})$ is bound by $XW[\bar{B}_b, S(\bar{v}_{i+1})]$. Similarly, the value of $p'_f$ is $p_f + 1$ if and only if $\bar{v}_{i+1} = \bar{w}_{p_f+1}$, and must be 0 otherwise. Further the value of $p'_b$ is $p_b + 1$ if and only if $S(\bar{v}_{i+1}) = \bar{w}_{p_b+1}$, and 0 otherwise.

**Corollary 2.** *Let $\bar{v}, \bar{u} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ be a pair of words and let $\bar{v}' = \bar{u}_{[1,i]} : \bar{v}_{[i+1,r]} : S(\bar{v}_{[1,i]} : \bar{u}_{[i+1,r]})$. Then $\bar{v}' \in \mathbf{B}(\bar{w}, p'_f, p'_b, \bar{B}_f{}', \bar{B}_b{}', i, j, r)$ if and only if $\bar{v} \in \mathbf{B}(\bar{w}, p'_f, p'_b, \bar{B}_f{}', \bar{B}_b{}', i, j, r)$.*

Proposition 4 and Corollary 2 are used in an analogous manner the Proposition 3. As before, the goal is not to directly compute the size of $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$, but rather to compute the number of suffixes of length $r-i$ of the words therein. To that end, let $SB(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ be the number of unique suffixes of length $r-i$ of words in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$. Note that the number of suffixes of length $r-i$ of words in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ equals the number of unique subwords between positions $i+1$ and $r$ of the words $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$. Additionally, note that following Corollary 2, the size of $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ can be computed by taking the product of the number of unique prefixes of words in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$, and the number of unique suffixes of words in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$. The process of computing the number of such suffixes is divided into four cases based on the values of $i, j$ and $r$.

When $i < j$, for every word $\bar{v} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$, $v_{i+1}$ must be greater than or equal to $\bar{w}_{p_f+1}$ to avoid there being a rotation less than $j$ for which $\bar{v}$ is less than $\bar{w}$. Further, the value of the relabelling of $\bar{v}_{i+1}$ must be greater than or equal to $\bar{w}_{p_b+1}$ to avoid any rotation in $[r+1, 2 \cdot r]$ being less than $\bar{w}$. Therefore, the symbol at position $i+1$ can be 0 if and only if $\bar{w}_{p_f+1} = 0$, and can be 1 if and only if $\bar{w}_{p_b+1} = 0$. The number of suffixes of length $r-i$ of words in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ where the symbol at position $i+1$ is 0 is equal to the value of $SB(\bar{w}, p'_f, p'_b, XW[\bar{B}_f, 0], XW[\bar{B}_b, 1], i, j, r)$, and the number of suffixes where the symbol at position $i+1$ is 1 is equal to the value of $SB(\bar{w}, p'_f, p'_b, XW[\bar{B}_f, 1], XW[\bar{B}_b, 0], i, j, r)$.

When $i = j$, then the value of $SB(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$ depends primarily on the value of $p_f$. If $p_f > 0$, then as $\langle v \rangle_j < \bar{w}$ for every $v \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B}_f, \bar{B}_b, i, j, r)$, $\langle v \rangle_{j-p_f} < \bar{w}$, contradicting the assumption that $j$ is the smallest rotation for which $\bar{v}$ is smaller than $\bar{w}$. Hence, if $p_f > 0$, then the set $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ must be empty and by extension have no suffixes of length $r-i$. If $p_f = 0$ then as $\bar{w}_1 = 0$, the symbol $\bar{v}_{i+1}$ must be 0 for every $\bar{v} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$. Therefore, the value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ is exactly equal to the value of $SB(\bar{w}, 1, p'_b, , XW[\bar{B_f}, 0], WX[\bar{B_b}, 1], i+1, j, r)$ of length $r-i-1$.

To count the number of suffixes of length $r-i$ when $i > j$, an auxiliary, technical set $\mathbf{Y}(\bar{w}, i, p_b, \bar{B_f})$ is introduced. Informally, $\mathbf{Y}(\bar{w}, i, p_b, \bar{B_f})$ contains the set of words of length $i$ such that every pair of words $\bar{u} \in \mathbf{Y}(\bar{w}, r-i, p_b, \bar{B_f})$ and $\bar{v} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, r-i, j, r)$, every suffix of the word $S(\bar{v}_{[1,r-i]} : u) : \bar{B_f}$ of length at least $r$ is greater than the prefix of $\bar{w}$ of the same length. In other words, $\mathbf{Y}(\bar{w}, i, p_b, \bar{B_f})$ contains the set of words that can be appended to prefixes of $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, r-i, j, r)$ while maintaining the condition that any rotation by more than $r$ results in a word strictly greater than $\bar{w}$. Treating the method of counting $\mathbf{Y}(\bar{w}, i, p_b, \bar{B_f})$ as a black box, the number of suffixes of length $i-r$ in $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ when $p_f < i-j$ is exactly the size of $\mathbf{Y}(\bar{w}, i, p_b, \bar{B_f})$. If $p_f = i-j$, then the observe that every word $\bar{v} \in \mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ must satisfy the conditions that $\bar{v}_{i+1} \leq \bar{w}_{p_f+1}$ and $S(\bar{v}_{i+1}) \geq \bar{w}_{p_b+1}$. If $\bar{w}_{p_f+1} = 1$ and $\bar{w}_{p_b+1} = 1$ then $\bar{v}_{i+1}$ must be 0, giving a total of $|\mathbf{Y}(\bar{w}, r-i-1, p_b+1, WX[\bar{B_f}, 0])|$ suffixes of length $r-i$. If $\bar{w}_{p_f+1} = 1$ and $\bar{w}_{p_b+1} = 0$ then there are $|\mathbf{Y}(\bar{w}, r-i-1, 0, WX[\bar{B_f}, 0])|$ suffixes of length $r-i$ where the first symbol is 0, and the number of $r-i$ length suffixes where the first symbol equals 1 is equal to the value of $SB(\bar{w}, p_f+1, p_b+1, WX[\bar{B_f}, 1], WX[\bar{B_b}, 0], i+1, j, r)$. If $\bar{w}_{p_f+1} = 0$ then $\bar{v}_{i+1}$ must be 0, and hence the value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ is $SB(\bar{w}, p_f+1, p'_b, WX[\bar{B_f}, 0], WX[\bar{B_b}, 1], i+1, j, r)$.

When $i = r$, the number of zero length suffixes of $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ is either 0, if $\bar{w}_{[1,p_f]} : \bar{B_f} \geq \bar{w}_{[1,p_f+r]}$, or 1 otherwise.

**Lemma 8.** *The size of $\beta(\bar{w}, j, r)$ can be computed in $O(n^5)$ time.*

*Proof (Sketch).* The size of $\beta(\bar{w}, j, r)$ is computed in an analogous manner to the size of $\alpha(\bar{w}, j, t)$ as shown in Lemma 7. This is done by computing the size value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ using the layout given above.

The value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ can be computed in $O(n)$ time if $i = r$, and $O(1)$ time if $i < r$ and the size of $SB(\bar{w}, p'_f, p'_b, \bar{B_f}', \bar{B_b}', i+1, j, r)$ has been precomputed for every $p'_f \in \{0, p_f+1\}, p_b \in \{0, p_b+1\}$ and $\bar{B_b}', \bar{B_f}' \sqsubseteq_i \bar{w}$. As there are at most $O(n^4)$ possible values of $p_f, p_b \in [r]$ and $\bar{B_b}, \bar{B_f} \sqsubseteq_r \bar{w}$, the value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, r, j, r)$ can be computed for every $p_f, p_b \in [r]$ and $\bar{B_b}, \bar{B_f} \sqsubseteq_r \bar{w}$ in $O(n^5)$ time. Similarly, as there are at most $O(n^5)$ possible values of $i \in [r], p_f, p_b \in [i]$ and $\bar{B_b}, \bar{B_f} \sqsubseteq_i \bar{w}$, the value of $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ can be computed in $O(n^5)$ time for every value of $i \in [r], p_f, p_b \in [i]$ and $\bar{B_b}, \bar{B_f} \sqsubseteq_i \bar{w}$.

Note that the set $\mathbf{B}(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ does not include the words in $\beta(\bar{w}, r, j)$ with a prefix that is a subword of $\bar{w}$. The number of such words can be computed in a brute force manner by finding the length of the longest prefix that is a subword of $\bar{w}$, and determining the number of possible suffixes. The number of such suffixes are counted in using $SB(\bar{w}, p_f, p_b, \bar{B_f}, \bar{B_b}, i, j, r)$ in a manner analogous to the way $SA(\bar{w}, p, i, j, r)$ is used in Lemma 7, to count the number of words in $\alpha(\bar{w}, j, r)$ with a prefix that is a subword of $\bar{w}$.

**Theorem 2.** *The value of $RankSN(\bar{w}, m)$ can be computed in $O(n^6 \log^2 n)$ time for any $m \leq n$.*

*Proof.* Following Lemmas 2, 3, and 4, the value of $RankSN(\bar{w}, m, r)$ is:

$$RankSN(\bar{w}, m, r) = \sum_{d|r} \left( \frac{1}{2 \cdot r} \sum_{p|d} \mu \left( \frac{d}{p} \right) |\mathbf{RA}(\bar{w}, m, S, p)| \right)$$

From Observations 3 and 2 , the size of $\mathbf{RA}(\bar{w}, m, S, r)$ equals $\sum_{j \in [m]} |\alpha(\bar{w}, r, j)| + |\beta(\bar{w}, r, j)|$. Following Lemma 7, the size of $\alpha(\bar{w}, r, j)$ can be computed in $O(n^3)$ time. Following Lemma 8, the size of $\beta(\bar{w}, r, j)$ can be computed in $O(n^5)$ time. As there are at most $O(n)$ values of $j$, the total time complexity for determining the size of $\mathbf{RA}(\bar{w}, m, S, r)$ is $O(n^6)$. As there are at most $O(\log n)$ possible divisors of $r$, the size of $\mathbf{RA}(\bar{w}, m, S, p)$ needs to be evaluated at most $O(\log n)$ times, giving a total time complexity of $O(n^6 \log n)$. The value of $RankSN(\bar{w}, m, r)$ can then be computed in at most $O(\log^2 n)$ time once the size of $\mathbf{RA}(\bar{w}, m, S, p)$ has been precomputed for every factor $p$ of $r$. Finally, following Lemma 5, the value of $RankSN(\bar{w}, m)$ can be computed from the value of $RankSN(\bar{w}, m, r)$ for at most $O(\log n)$ values of $r$. Therefore the total time complexity of computing $RankSN(\bar{w}, m)$ is $O(n^6 \log^2 n)$.

## 5 Enclosing Necklaces

This section shows how to rank a word $\bar{w}$ within the set of binary unlabelled necklaces enclosing $\bar{w}$. Note that the rank of $\bar{w}$ within this set is equivalent to the number of binary unlabelled necklaces enclosing $\bar{w}$. As with the ranking approach to symmetric necklaces, we start with the key theoretical results that inform our approach.

**Lemma 9.** *Let $RankEN(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of necklaces of length $m$ that enclose $\bar{w}$ and let $RankEL(\bar{w}, m)$ be the rank of $\bar{w}$ within the set of Lyndon words of length $m$ that enclose $\bar{w}$. $RankEN(\bar{w}, m) = \sum_{d|m} RankEL(\bar{w}, d)$.*

*Proof.* Observe that every necklace counted by $RankEN(\bar{w}, m)$ must have a unique period that is a factor of $m$, hence $RankEN(\bar{w}, m) = \sum_{d|m} RankEL(\bar{w}, d)$.

**Lemma 10.** *Let $\mathbf{EL}(\bar{w}, m)$ be the set of words of length $m$ belonging to a Lyndon word that encloses $\bar{w}$. $RankEL(\bar{w}, m) = \frac{|\mathbf{EL}(\bar{w}, m))}{m}$.*

*Proof.* Following the same arguments as in Lemma 3, every aperiodic necklace counted by $RankEL(\bar{w}, m)$ must have exactly $m$ words in $\mathbf{EL}(\bar{w}, m)$ representing it. Therefore $RankEL(\bar{w}, m) = \frac{|\mathbf{EL}(\bar{w}, m))}{m}$.

**Lemma 11.** *Let* $\mathbf{EL}(\bar{w}, m)$ *be the set of words of length* $m$ *belonging to a Lyndon word that encloses* $\bar{w}$ *and let* $\mathbf{EN}(\bar{w}, m)$ *be the set of words of length* $m$ *belonging to a necklace that encloses* $\bar{w}$. *The size of* $\mathbf{EL}(\bar{w}, m)$ *equals* $\sum_{d|m} \mu\left(\frac{m}{d}\right) |\mathbf{EN}(\bar{w}, d)|$.

*Proof.* Following the same arguments as in Lemma 9, the size of $\mathbf{EN}(\bar{w}, m)$ can be expressed in terms of the size of $\mathbf{EL}(\bar{w}, d)$ for every factor $d$ of $m$ as $|\mathbf{EN}(\bar{w}, m)| = \sum_{d|m} |\mathbf{EL}(\bar{w}, d)|$. Applying the Möbius inversion formula to this equation gives $|\mathbf{EL}(\bar{w}, m)| = \sum_{d|m} \mu\left(\frac{m}{d}\right) |\mathbf{EN}(\bar{w}, d)|$.

As in the Symmetric case, we partition the set of necklaces into a series of subsets for ease of computation. Let $\gamma(\bar{w}, m, r)$ denote the set of words belonging to a necklace which encloses $\bar{w}$ such that $r$ is the smallest rotation for which $\bar{v} \in \gamma(\bar{w}, m, r)$ is smaller than $\bar{w}$, i.e. the smallest value where $\langle \bar{v} \rangle_r < \bar{w}$. We further introduce the set $\mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b) \subseteq \gamma(\bar{w}, m, r)$ as the set of words where every $\bar{v} \in \mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b)$ satisfies the following conditions:

1. $\langle \bar{v} \rangle_s > \bar{w}$ for every $s \in [r-1]$.
2. $\langle S(\bar{v}) \rangle_s > \bar{w}$ for every $s \in [m]$.
3. $\langle \bar{v} \rangle_r < \bar{w}$.
4. $\bar{v}_{[1,i]}$ is bound by $\bar{B_f} \sqsubseteq_i \bar{w}$.
5. $S(\bar{v}_{[1,i]})$ is bound by $\bar{B_b} \sqsubseteq_i \bar{w}$.
6. $p_f$ is the length of the longest suffix of $\bar{v}_{[1,i]}$ that is a prefix of $\bar{w}$, i.e. the largest value such that $\bar{v}_{[i-p_f,i]} = \bar{w}_{[1,p_f]}$.
7. $p_b$ is the length of the longest suffix of $S(\bar{v}_{[1,i]})$ that is a prefix $\bar{w}$, i.e. the largest value such that $S(\bar{v}_{[i-p_b,i]}) = \bar{w}_{[1,p_b]}$.

Note that Conditions 1, 2, and 3 are the necessary conditions for $\bar{v}$ to be in $\gamma(\bar{w}, m, r)$. As before, we break our dynamic programming based approach into several sub cases based on the value of $i$ relative to $r$. As in the symmetric case, we relay upon a technical proposition.

**Proposition 5.** *Given* $\bar{v} \in \mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b)$, $\bar{v}$ *also belongs to* $\mathbf{C}(\bar{w}, i+1, r, WX[\bar{B_f}, \bar{v}_{i+1}], WX[\bar{B_b}, \bar{v}_{i+1}], p'_f, p'_b)$

**Corollary 3.** *Given a pair of words* $\bar{v}, \bar{u} \in \mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b)$ *let* $\bar{v}' = \bar{v}_{[1,i]} : \bar{u}_{[i+1,m]}$. *Then* $\bar{v}' \in \mathbf{C}(\bar{w}, i+1, r, \bar{B_f}', \bar{B_b}', p'_f, p'_b)$ *if and only if* $\bar{v} \in \mathbf{C}(\bar{w}, i+1, r, \bar{B_f}', \bar{B_b}', p'_f, p'_b)$.

**Theorem 3.** *Let* $RankEN(\bar{w}, m)$ *be the rank of* $\bar{w}$ *within the set of necklaces of length* $m$ *which enclose* $\bar{w} \in \Sigma^n$. *The value of* $RankEN(\bar{w}, n)$ *can be computed in* $O(n^6 \log n)$ *time for any* $m \leq n$.

*Proof (Sketch).* The high level idea is to compute the size of $\mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b)$ in a dynamic manner analogous to the computation of the size of $\mathbf{A}(\bar{w}, p, \bar{B}, i, j, r)$. Starting with $i = m$ as the base case and progressing in descending value of $i$, the size of $\mathbf{C}(\bar{w}, i, r, \bar{B_f}, \bar{B_b}, p_f, p_b)$ is computed for every

$\bar{B}_f, \bar{B}_b \sqsubseteq_i \bar{w}$, $p_f, p_b \in [i]$. By showing that the size of $\mathbf{C}(\bar{w}, i, r, \bar{B}_f, \bar{B}_b, p_f, p_b)$ can be computed in $O(1)$ time for any $i < m$, and $O(n)$ time when $i = m$, the size of $\mathbf{C}(\bar{w}, i, r, \bar{B}_f, \bar{B}_b, p_f, p_b)$ for every $i, j \in [m], \bar{B}_f, \bar{B}_b \sqsubseteq_i \bar{w}$, and $p_f, p_b \in [i]$ is computed in $O(n^6)$ time. The additional complexity is due to number of lengths that need to be computed following Lemmas 9, 10 and 11.

## References


1. Adamson, D., Deligkas, A., Gusev, V.V., Potapov, I.: Combinatorial algorithms for multidimensional necklaces. arXiv preprint https://arxiv.org/abs/2108.01990 (2021)
2. Adamson, D., Deligkas, A., Gusev, V.V., Potapov, I.: Ranking bracelets in polynomial time. In: 32nd Annual Symposium on Combinatorial Pattern Matching (CPM 2021). Leibniz International Proceedings in Informatics (LIPIcs), vol. 191, pp. 4:1–4:17 (2021)
3. De Felice, C., Zaccagnino, R., Zizza, R.: Unavoidable sets and circular splicing languages. Theor. Comput. Sci. **658**, 148–158 (2017). Formal languages and automata: models, methods and application. In: Honour of the 70th birthday of Antonio Restivo
4. Filos-Ratsikas, A., Goldberg, P.W.: The complexity of splitting necklaces and bisecting ham sandwiches. In: Charikar, M., Cohen, E. (eds.) Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, 23–26 June 2019, pp. 638–649. ACM (2019)
5. Gilbert, E.N., Riordan, J.: Symmetry types of periodic sequences. Ill. J. Math. **5**(4), 657–665 (1961)
6. Graham, R.L., Knuth, D.E., Patashnik, O.: Concrete Mathematics: A Foundation for Computer Science. Addison-Wesley (1994)
7. Kociumaka, T., Radoszewski, J., Rytter, W.: Computing $k$-th Lyndon word and decoding lexicographically minimal de Bruijn sequence. In: Kulikov, A.S., Kuznetsov, S.O., Pevzner, P. (eds.) CPM 2014. LNCS, vol. 8486, pp. 202–211. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07566-2_21
8. Kopparty, S., Kumar, M., Saks, M.: Efficient indexing of necklaces and irreducible polynomials over finite fields. Theory Comput. **12**(1), 1–27 (2016)
9. Mareš, M., Straka, M.: Linear-time ranking of permutations. In: Arge, L., Hoffmann, M., Welzl, E. (eds.) ESA 2007. LNCS, vol. 4698, pp. 187–193. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75520-3_18
10. Myrvold, W., Ruskey, F.: Ranking and unranking permutations in linear time. Inf. Process. Lett. **79**(6), 281–284 (2001)
11. Sawada, J., Williams, A.: Practical algorithms to rank necklaces, Lyndon words, and de Bruijn sequences. J. Discret. Algorithms **43**, 95–110 (2017)
12. Shimizu, T., Fukunaga, T., Nagamochi, H.: Unranking of small combinations from large sets. J. Discret. Algorithms **29**, 8–20 (2014)
13. Williamson, S.G.: Ranking algorithms for lists of partitions. SIAM J. Comput. **5**(4), 602–617 (1976)