Three Generalizations of Davenport-Schinzel Sequences*

Seth Pettie
University of Michigan

October 16, 2018

Abstract

We present new, and mostly sharp, bounds on the maximum length of certain generalizations of Davenport-Schinzel sequences. Among the results are sharp bounds on order-s *double DS* sequences, for all s, sharp bounds on sequences avoiding *catenated permutations* (aka formation free sequences), and new lower bounds on sequences avoiding *zig-zagging* patterns.

1 Introduction

A generalized Davenport-Schinzel (DS) sequence is one over a finite alphabet, say $[n] = \{1, \ldots, n\}$, none of whose subsequences are isomorphic to a fixed forbidden sequence σ or a set of such sequences. (A sparsity criterion is also included in order to prohibit degenerate infinite sequences such as $aaaaa\cdots$.) When σ is the alternating sequence $abab\cdots$ with length $s+2$ this definition reverts to that of standard order-s DS sequences. Whereas standard DS sequences have countless applications in discrete and computational geometry, generalized DS sequences have found fewer applications [29, 24, 6, 20, 17, 3]. Whereas bounding the length of DS sequences is now essentially a closed problem [2, 16, 22], the most basic questions about generalized DS sequences are open, or have received only partial answers.

We are mainly interested in answering two questions about forbidden sequences. A purely quantitative question is to determine the maximum length $\mathrm{Ex}(\sigma, n)$ of a σ-free sequence over an n-letter alphabet, for specific σ or large classes of σ. An equally interesting question, particularly when $\mathrm{Ex}(\sigma, n)$ is superlinear in n, is to characterize the *structure* of σ-free sequences. There are infinitely many forbidden sequences one could study, but some classes of subsequences are more interesting than others, either because of their applications, or their intrinsic structure, or for historical reasons. In this report we focus on forbidden sequences that generalize, in various ways, the idea of an alternating sequence. In order to properly explain our results, in Section 1.4, we need to introduce some notation and terminology and to review the history of DS sequences and their generalizations, in Sections 1.1–1.3. For the moment we can take a high-level tour of the results. Following convention, let $\lambda_s(n) = \mathrm{Ex}(abab\cdots, n)$ be the extremal function for order-s DS sequences, where the alternating pattern has length $s+2$.

*This work is supported by NSF CAREER grant CCF-0746673, NSF grants CCF-1217338 and CNS-1318294, and a grant from the US-Israel Binational Science Foundation.

Double DS sequences. The most modest way to generalize an alternating sequence $abab\cdots$ is simply to *double* each letter, transforming it to $abbaabb\cdots$.[1] Double DS sequences were the first generalized DS sequences to be studied [5, 1, 14]. Let λ_s^{dbl} be the extremal function of order-s double DS sequences. Davenport and Schinzel [5] noted that $\lambda_1^{\mathrm{dbl}}(n)$ is linear (see [13, p. 13]) and Adamec, Klazar, and Valtr [1] proved that $\lambda_2^{\mathrm{dbl}}(n)$ is also linear, matching λ_1 and λ_2 up to constant factors. (The forbidden sequences here are *abba* and *abbaab*.) Klazar and Valtr [14] claimed without proof that $\lambda_3^{\mathrm{dbl}}(n) = \Theta(n\alpha(n))$, which would match λ_3 asymptotically [9]. However, this claim was later retracted [13]. Here $\alpha(n)$ is the inverse-Ackermann function. We prove that $\lambda_3^{\mathrm{dbl}}(n)$ is, in fact, $\Theta(n\alpha(n))$, and more generally, that λ_s^{dbl} and λ_s are asymptotically equivalent for every order s.

Perm-free Sequences. Take any $s+1$ permutations over $\{a, b\}$. Regardless of one's choice, the concatenation of these permutations necessarily contains an alternating subsequence of length $s+2$: the first permutation contributes two symbols and every subsequent permutation at least one. Define $\mathrm{Perm}_{r,s+1}$ to be the set of all sequences obtained by concatenating $s+1$ permutations over an r-letter alphabet, and let $\Lambda_{r,s}$ be the extremal function of $\mathrm{Perm}_{r,s+1}$-free sequences.[2] The argument above shows that order-s DS sequences are $\mathrm{Perm}_{2,s+1}$-free, which implies that $\lambda_s(n) \leqslant \Lambda_{2,s}(n)$. Klazar [10] introduced $\mathrm{Perm}_{r,s+1}$-free sequences as a "universal" method for finding upper bounds on $\mathrm{Ex}(\sigma, n)$. If there exist r, s (and there always do) such that σ is contained in every member of $\mathrm{Perm}_{r,s+1}$, then $\mathrm{Ex}(\sigma, n) = O(\Lambda_{r,s}(n))$.

It is straightforward to show that $\lambda_s(n)$ and $\Lambda_{2,s}(n)$ are asymptotically equivalent. A natural hypothesis, given [16, 22], is that λ_s and $\Lambda_{r,s}$ are asymptotically equivalent, for all r. We prove that this hypothesis is false, which is quite surprising. One upshot of [2, 16, 22] is that when $s \geqslant 7$ is odd, $\lambda_s(n)$ and $\lambda_{s-1}(n)$ are essentially indistinguishable, and that $\lambda_5(n)$ and $\lambda_4(n)$ are asymptotically distinguishable, but very similar. In contrast, we prove that, in general, $\Lambda_{r,s}(n)$ behaves very differently at odd and even s. The extremal functions λ_s and $\Lambda_{r,s}$ are asymptotically equivalent *only* when $s \leqslant 3$, or $s \geqslant 4$ is even, or $r = 2$.

Just as DS sequences can be generalized to double DS sequences, $\mathrm{Perm}_{r,s+1}$ can be transformed into a set $\mathrm{Perm}_{r,s+1}^{\mathrm{dbl}}$ by "doubling" it. Let $\Lambda_{r,s}^{\mathrm{dbl}}(n)$ be the extremal function of $\mathrm{Perm}_{r,s+1}^{\mathrm{dbl}}$-free sequences. The function $\Lambda_{r,s}^{\mathrm{dbl}}$ was studied in a different, but essentially equivalent form by Cibulka and Kynčl [3]. We prove that $\Lambda_{r,s}^{\mathrm{dbl}}$ is asymptotically equivalent to $\Lambda_{r,s}$ for all r, s. This fact is not surprising, but what *is* surprising is how many new techniques are needed to prove it when $s = 3$.

Zig-zagging Patterns. One way to view the alternating sequence $abab\cdots$ with length $s+2$ is as a *zig-zagging* pattern with $s+1$ *zigs* and *zags*. Generalized to larger alphabets, we obtain the N-shaped sequences, of the form $ab\cdots zy\cdots ab\cdots z$, when $s = 2$, the M-shaped sequences $ab\cdots zy\cdots ab\cdots zy\cdots a$, when $s = 3$, the $N\!\!N$-shaped sequences $ab\cdots zy\cdots ab\cdots zb\cdots ab\cdots z$, when $s = 4$, and so on. Klazar and Valtr [14] (see also [20]) proved that the extremal function of each N-shaped forbidden sequence is linear, matching $\lambda_2(n)$. See Valtr [29] for an application of N-shaped sequences to bounding the size of geometric graphs and Pettie [20] for an application of M-shaped sequences to bounding the complexity of the union of fat triangles.

Given [14, 20], one is tempted to guess that the extremal function for a zig-zagging forbidden sequence is, if not asymptotically equivalent to the corresponding order-s DS sequence, at least

[1] It is straightforward to show that repeating letters more than twice, or repeating the first and last at all, can affect the extremal function by at most a constant factor. See [1].

[2] These were called $(r, s+1)$-*formation-free* sequences by Nivasch [16].

close to it. We give lower bounds showing that for each t, there is an M-shaped forbidden sequence with extremal function $\Omega(n\alpha^t(n))$ and an $\mathcal{N}$-shaped forbidden sequence with extremal function $\Omega(n\cdot 2^{(1+o(1))\alpha^t(n)/t!})$. Put a different way, in terms of their extremal functions M-shaped sequences may be similar to *ababa* but $\mathcal{N}$-shaped sequences bear no resemblance to *ababab*.

Our results on zig-zagging patterns are the least conclusive, and therefore offer the most opportunities for future research. They are based on a general, parameterized method for constructing non-linear sequences.

1.1 Sequence Notation and Terminology

Let $|\sigma|$ be the length of a sequence $\sigma = (\sigma_i)_{1\leqslant i\leqslant|\sigma|}$ and let $\|\sigma\|$ be the size of its alphabet $\Sigma(\sigma) = \{\sigma_i\}$. Two equal length sequences are *isomorphic* if they are the same up to a renaming of their alphabets. We say σ is a *subsequence* of σ' if σ can be obtained by deleting symbols from σ'. The predicate $\sigma \prec \sigma'$ asserts that σ is isomorphic to a subsequence of σ'. If $\sigma \nprec \sigma'$ we say σ' is *σ-free*. If P is a *set* of sequences, $\sigma \prec P$ holds if $\sigma \prec \sigma'$ for every $\sigma' \in P$ and $P \nprec \sigma$ holds if $\sigma' \nprec \sigma$ for every $\sigma' \in P$. The assertion that σ *appears in* or *occurs in* or *is contained in* σ' means $\sigma \prec \sigma'$. The *projection* of a sequence σ onto $G \subseteq \Sigma(\sigma)$ is obtained by deleting all non-G symbols from σ. A sequence σ is *k-sparse* if whenever $\sigma_i = \sigma_j$ and $i \neq j$, then $|i-j| \geqslant k$. A *block* is a sequence of distinct symbols. If σ is understood to be partitioned into a sequence of blocks, $[\![\sigma]\!]$ is the number of blocks. The predicate $[\![\sigma]\!] = m$ asserts that σ can be partitioned into at most m blocks. The extremal functions for *generalized* Davenport-Schinzel sequences are defined to be

$$\begin{aligned}\mathrm{Ex}(\sigma,n,m) &= \max\{|S| \;:\; \sigma\nprec S,\ \|S\|=n,\ \text{and } [\![S]\!]\leqslant m\}\\ \mathrm{Ex}(\sigma,n) &= \max\{|S| \;:\; \sigma\nprec S,\ \|S\|=n,\ \text{and } S \text{ is } \|\sigma\|\text{-sparse}\}\end{aligned}$$

where σ may be a single sequence or a set of sequences. The conditions "$[\![S]\!] \leqslant m$" and "S is $\|\sigma\|$-sparse" guarantee that the extremal functions are finite. Note that $\mathrm{Ex}(\sigma,n,m)$ has no sparseness criterion. The extremal functions for order-s DS sequences are defined to be

$$\lambda_s(n) = \mathrm{Ex}(\overbrace{abab\cdots}^{\text{length } s+2},n) \qquad \text{and} \qquad \lambda_s(n,m) = \mathrm{Ex}(\overbrace{abab\cdots}^{\text{length } s+2},n,m).$$

Since $\|abab\cdots\| = 2$, the sparseness criterion forbids only immediate repetitions.

1.2 Davenport, Schinzel, Ackermann, Tarjan

Davenport and Schinzel [4] observed that $\lambda_1(n) = n$ and $\lambda_2(n) = 2n-1$. It took several decades for all the other orders to be understood. The following theorem synthesizes results of Hart and Sharir [9], Agarwal, Sharir, and Shor [2], Klazar [12], Nivasch [16], and Pettie [22].

Theorem 1.1. *Let $\lambda_s(n)$ be the maximum length of a repetition-free sequence over an n-letter* *alphabet avoiding subsequences isomorphic to* $abab\cdots$ *(length* $s+2$*). Then* λ_s *satisfies:*

$$\lambda_s(n) = \begin{cases} n & s = 1 \\ 2n-1 & s = 2 \\ 2n\alpha(n) + O(n) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ \Theta(n\alpha(n)2^{\alpha(n)}) & s = 5 \\ n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))} & s \geqslant 6,\ t = \lfloor \frac{s-2}{2} \rfloor. \end{cases}$$

Here $\alpha(n)$ is the functional inverse of Ackermann's function discovered by Tarjan [28], defined as follows.

$$\begin{aligned} a_{1,j} &= 2^j & j \geqslant 1 \\ a_{i,1} &= 2 & i \geqslant 2 \\ a_{i,j} &= w \cdot a_{i-1,w} & i,j \geqslant 2 \\ &\quad \text{where } w = a_{i,j-1} \end{aligned}$$

One may check that in the table $(a_{i,j})$, the first column is constant and the second column merely exponential: $a_{i,1} = 2$ and $a_{i,2} = 2^i$. Ackermann-type growth only appears at the third column, motivating the following definition of the inverse functions.

$$\begin{aligned} \alpha(n,m) &= \min\{i \mid a_{i,j} \geqslant m,\ \text{ where } j = \max\{\lceil n/m \rceil, 3\}\} \\ \alpha(n) &= \alpha(n,n) \end{aligned}$$

There are numerous variants of Ackermann's function in the literature, all of which are equivalent inasmuch as their inverses differ by at most a constant. Observe that Theorem 1.1 is robust to perturbations of $\alpha(n)$ by $O(1)$, so it does not depend on any particular definition of Ackermann's function or its inverse.[3]

1.3 Generalizations of DS Sequences

Certain classes of forbidden sequences have received significant attention. We review three systems for generalizing (standard) DS sequences, then mention some miscellaneous results in the area.

Double DS Sequences. Let $\mathrm{dbl}(\sigma)$ be obtained from σ by doubling each letter except for the first and last, for example, $\mathrm{dbl}(abcabc) = abbccaabbc$. The extremal functions for order-s *double DS* sequences are $\lambda_s^{\mathrm{dbl}}(n) = \mathrm{Ex}(\mathrm{dbl}(abab\cdots), n)$ and $\lambda_s^{\mathrm{dbl}}(n,m) = \mathrm{Ex}(\mathrm{dbl}(abab\cdots), n, m)$, where the alternating sequence has length $s{+}2$. It is known that $\lambda_1^{\mathrm{dbl}}(n)$ and $\lambda_2^{\mathrm{dbl}}(n)$ are linear, matching λ_1 and λ_2 asymptotically. See Davenport and Schinzel [5], Adamec, Klazar, and Valtr [1], and Klazar [11, 13, p. 13]. Pettie [19, 20] proved that $\lambda_3^{\mathrm{dbl}}(n) = O(n\alpha^2(n))$ and $\mathrm{Ex}(\{abbaabba, ababab\}, n) = \Theta(n\alpha(n))$, and that for $s \geqslant 4$, $\lambda_s^{\mathrm{dbl}}(n)$ matched what were the best upper bounds on $\lambda_s(n)$ at the time [16], namely $\lambda_s^{\mathrm{dbl}}(n) < n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))}$, for even s, and $\lambda_s^{\mathrm{dbl}}(n) < n \cdot 2^{\alpha^t(n)(\log(\alpha(n))+O(1))/t!}$, for odd s.

[3] See Pettie [22, p. 4] for a discussion of this notion of "Ackermann-invariance."

Catenated Permutations. Recall that $\mathrm{Perm}_{r,s+1}$ is defined to be the set of sequences obtained by concatenating $s+1$ permutations over an r-letter alphabet. For example, $abcd\ cbad\ badc \in \mathrm{Perm}_{4,3}$. Let $\Lambda_{r,s}(n) = \mathrm{Ex}(\mathrm{Perm}_{r,s+1}, n)$ to be the extremal function for $\mathrm{Perm}_{r,s+1}$-free sequences, with $\Lambda_{r,s}(n,m)$ defined analogously.[4] It is straightforward to show that if σ is contained in every member of $\mathrm{Perm}_{r,s+1}$ then

$$\mathrm{Ex}(\sigma, n, m) \leqslant \Lambda_{r,s+1}(n,m) \qquad \text{and} \qquad \mathrm{Ex}(\sigma, n) = O(\Lambda_{r,s+1}(n)).$$

Nivasch [16] proved that any σ is contained in every member of $\mathrm{Perm}_{\|\sigma\|,|\sigma|-\|\sigma\|+1}$. Very recently Geneson, Prasad, and Tidor [8] showed that it suffices to consider a subset $\mathrm{Bin}_{r,s+1} \subset \mathrm{Perm}_{r,s+1}$ consisting of *binary* patterns, where each of the $s+1$ permutations is either $12\cdots(r-1)r$ or $r(r-1)\cdots 21$. By repeated application of the Erdős-Szekeres theorem, they showed that every member of $\mathrm{Perm}_{r',s+1}$ contains a member of $\mathrm{Bin}_{r,s+1}$, where $r' = (r-1)^{2^s}+1$. Consequently, if σ is contained in every member of $\mathrm{Bin}_{r,s+1}$ then $\mathrm{Ex}(\sigma,n) = O(\Lambda_{r',s}(n))$.

Nivasch [16], improving [10], gave the following upper bounds on $\Lambda_{r,s}$, for any $r \geqslant 2, s \geqslant 1$, where $t = \lfloor \frac{s-2}{2} \rfloor$. The lower bounds follow from previous [9, 2] and subsequent [22] constructions of order-s DS sequences.

$$\Lambda_{r,s}(n) = \begin{cases} \Theta(n) & s \leqslant 2 \\ \Theta(n\alpha(n)) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ \Omega\big(n\alpha(n)2^{\alpha(n)}\big) \text{ and } O\big(n2^{\alpha(n)(\log\alpha(n)+O(1))}\big) & s = 5 \\ n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))} & \text{even } s \geqslant 6 \\ \Omega\Big(n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))}\Big) \text{ and } O\Big(n \cdot 2^{\alpha^t(n)(\log\alpha(n)+O(1))/t!}\Big) & \text{odd } s \geqslant 7 \end{cases}$$

Note that $\Lambda_{r,s}$ matches the behavior of λ_s when $s \leqslant 3$ or s is even.

Cibulka and Kynčl [3] studied a problem on 0-1 matrices that is essentially equivalent to the following generalization of Perm-free sequences. Define $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$ to be the set of all sequences over $[r] = \{1,\ldots,r\}$ that can be written $\sigma_1 \ldots \sigma_{s+1}$, where σ_1 and σ_{s+1} are permutations of $[r]$ and $\sigma_2,\ldots,\sigma_s$ are sequences containing two copies of each symbol in $[r]$. Define $\Lambda^{\mathrm{dbl}}_{r,s}(n)$ and $\Lambda^{\mathrm{dbl}}_{r,s}(n,m)$ to be the extremal functions of $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequences. Cibulka and Kynčl only considered $\Lambda^{\mathrm{dbl}}_{r,s}(n,m)$. For consistency we state the bounds on $\Lambda^{\mathrm{dbl}}_{r,s}(n)$ they *would* have obtained using the

[4] The "$s+1$" here is chosen to highlight the parallels with order-s DS sequences. Recall that every $\sigma \in \mathrm{Perm}_{2,s+1}$ contains an alternating sequence $abab\cdots$ with length $s+2$, hence $\lambda_s(n) \leqslant \Lambda_{2,s}(n)$.

available reductions from r-sparse to blocked sequences [16].[5] For any $r \geqslant 2, s \geqslant 1$, and $t = \lfloor \frac{s-2}{2} \rfloor$,

$$\Lambda^{\mathrm{dbl}}_{r,s}(n) = \begin{cases} \Theta(n) & s = 1 \\ \Omega(n) \text{ and } O(n\alpha(n)) & s = 2 \\ \Omega(n\alpha(n)) \text{ and } O(n\alpha^2(n)) & s = 3 \\ \Omega(n2^{\alpha(n)}) \text{ and } O(n\alpha^2(n)2^{\alpha(n)}) & s = 4 \\ \Omega\big(n\alpha(n)2^{\alpha(n)}\big) \text{ and } O\big(n2^{\alpha(n)(\log\alpha(n)+O(1))}\big) & s = 5 \\ n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))} & \text{even } s \geqslant 6 \\ \Omega\Big(n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))}\Big) \text{ and } O\Big(n \cdot 2^{\alpha^t(n)(\log\alpha(n)+O(1))/t!}\Big) & \text{odd } s \geqslant 7 \end{cases}$$

The definition of $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$ may at first seem unnatural. Surely $\mathrm{dbl}(\mathrm{Perm}_{r,s+1}) = \{\mathrm{dbl}(\sigma) \mid \sigma \in \mathrm{Perm}_{r,s+1}\}$ would be a more useful way to "double" the set $\mathrm{Perm}_{r,s+1}$. For example, it is known that $abcacbc \prec \mathrm{Perm}_{4,4}$, and therefore that $\mathrm{dbl}(abcacbc) \prec \mathrm{dbl}(\mathrm{Perm}_{4,4})$, but we cannot immediately conclude, as we would like, that $\mathrm{Ex}(\mathrm{dbl}(abcacbc), n) \leqslant \Lambda^{\mathrm{dbl}}_{4,3}(n)$. It turns out that the maximum length of $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequences and $\mathrm{dbl}(\mathrm{Perm}_{r,s+1})$-free sequences are the same asymptotically. The proof of Lemma 1.2 appears in the appendix.

Lemma 1.2. *The following bounds hold for any $r \geqslant 2, s \geqslant 1$.*

$$\mathrm{Ex}(\mathrm{dbl}(\mathrm{Perm}_{r,s+1}), n, m) \leqslant r \cdot \Lambda^{\mathrm{dbl}}_{r,s}(n,m) + 2rn$$
$$\mathrm{Ex}(\mathrm{dbl}(\mathrm{Perm}_{r,s+1}), n) = O(\Lambda^{\mathrm{dbl}}_{r,s}(n)).$$

Zig-zagging Patterns. Klazar and Valtr [14] introduced the N-shaped zig-zagging patterns $\{N_k\}$, where

$$N_k = 1\ 2\ \cdots\ (k+1)\ k\ \cdots\ 1\ 2\ \cdots\ (k+1).$$

Note that N_k-free sequences generalize order-2 DS sequences since $N_1 = abab$. (The vertical placement of the symbols in N_k carries no meaning. It is only intended to improve readability.) It was shown [14, 20] that $\mathrm{Ex}(\mathrm{dbl}(N_k), n) = O(n)$, which matches $\lambda_2(n)$ asymptotically. Pettie [20] proved that $\mathrm{Ex}(\{M_k, ababab\}, n) = \Theta(n\alpha(n))$, matching $\lambda_3(n)$, where M_k is the kth M-shaped sequence,

$$M_k = 1\ 2\ \cdots\ (k+1)\ k\ \cdots\ 1\ 2\ \cdots\ (k+1)\ k\ \cdots\ 1.$$

See [29, 24, 6, 20] for applications of N- and M-shaped sequences.

A different way to view even-length alternating patterns $abab\cdots$ with length $s+2$ is as a sequence of $(s+2)/2$ zigs, without corresponding zags. When generalized to an r-letter alphabet we get the sequence $(12\cdots r)^{(s+2)/2}$, which is contained in every member of $\mathrm{Bin}_{r,s+1}$ since at least $\lceil \frac{s+1}{2} \rceil$ of the constituent permutations must be identical. It follows from [8, 2, 16] that $\mathrm{Ex}((1\cdots r)^{(s+2)/2}, n) = \Theta(\Lambda_{r',s}(n)) = n \cdot 2^{(1+o(1))\alpha^t(n)/t!}$, where $r' = (r-1)^{2^s} + 1$ and $t = \lfloor \frac{s-2}{2} \rfloor$.

[5] The only notable case here is $s = 4$. Cibulka and Kynčl proved that $\Lambda^{\mathrm{dbl}}_{r,1}(n,m) = O(n+m)$, $\Lambda^{\mathrm{dbl}}_{r,2}(n,m) = O((n+m)\alpha(n,m))$ and $\Lambda^{\mathrm{dbl}}_{r,4}(n,m) = O((n+m)\alpha(n,m)2^{\alpha(n,m)})$, which imply, by [16, Lem. 5.7], that $\Lambda^{\mathrm{dbl}}_{r,2}(n) = O(n\alpha(n))$ and $\Lambda^{\mathrm{dbl}}_{r,4}(n) = O(n\alpha^2(n)2^{\alpha(n)})$.

Other Forbidden Patterns. Much of the research on generalized DS sequences [1, 14, 13, 20, 21, 19, 18] has focussed on delineating linear and non-linear forbidden sequences. A σ is *linear* if $\mathrm{Ex}(\sigma, n) = O(n)$. It is known that *ababa* and *abcacbc* are the only 2-sparse minimally non-linear sequences over three letters [14, 19, 20]. There are only a few varieties of sequences known to be linear. We have already seen that doubled N-shaped sequences ($\mathrm{dbl}(N_k)$) are in this category. Pettie [20, 18] proved that *abcbbccac* is linear, and showed that if π_1, π_2 are two permutations on the same alphabet, then $\pi_1\,\mathrm{dbl}(\pi_2)$ is linear. For example, $\mathrm{Ex}(abcde\,aacceebbd, n) = O(n)$. More linear sequences can be generated via Klazar and Valtr's [14] splicing operation. If $\sigma = \sigma_1 aa\sigma_2$ and σ' are linear, where $\Sigma(\sigma) \cap \Sigma(\sigma') = \varnothing$, then $\sigma_1 a\sigma' a\sigma_2$ is also linear.

Other research has focussed on identifying cofinal sets of forbidden sequences, with respect to the total order on extremal functions.[6] Klazar's general upper bounds [10] imply that standard DS sequences $\{(ab)^k\}$ are cofinal. Pettie [19], answering a question of Klazar [13], proved that the set of *ababa*-free forbidden sequences is also cofinal. This fact is witnessed by the two-sided *comb*-shaped sequences $\{D_k\}$, which generalize $D_1 = abacacbc$. Here D_k is defined to be

$$D_k = 1\ 2\ 1\ 3\ 1\ 4\ \cdots\ 1\ (k+2)\ 1\ (k+2)\ 2\ (k+2)\ 3\ (k+2)\ 4\ (k+2)\ \cdots\ (k+1)\ (k+2).$$

1.4 New Results

In prior work [22] we showed that λ_s behaves very similarly at the odd and even orders. In this paper we prove, quite unexpectedly, that $\Lambda_{r,s}$ matches λ_s *only* when $s \leqslant 3$, or $s \geqslant 4$ is even, or $r = 2$. When $s \geqslant 5$ is odd and $r \geqslant 3$, $\Lambda_{r,s}$ and λ_s diverge. Moreover, we prove that λ_s and λ_s^{dbl} are essentially equivalent, and that $\Lambda_{r,s}$ and $\Lambda_{r,s}^{\mathrm{dbl}}$ are essentially equivalent.

Theorem 1.3. *(Omnibus Bounds) For all $s \geqslant 1$ and $r = 2$, λ_s, λ_s^{dbl}, $\Lambda_{r,s}$, and $\Lambda_{r,s}^{\mathrm{dbl}}$ are asymptotically equivalent, namely,*

$$\lambda_s(n), \lambda_s^{\mathrm{dbl}}(n), \Lambda_{2,s}(n), \Lambda_{2,s}^{\mathrm{dbl}}(n) = \begin{cases} \Theta(n) & s \leqslant 2 \\ \Theta(n\alpha(n)) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ \Theta(n\alpha(n)2^{\alpha(n)}) & s = 5 \\ n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))} & s \geqslant 6, \text{ where } t = \lfloor \frac{s-2}{2} \rfloor. \end{cases}$$

However, the behavior of $\Lambda_{r,s}$ and $\Lambda_{r,s}^{\mathrm{dbl}}$ changes when $r \geqslant 3$. In particular,

$$\Lambda_{r,s}(n), \Lambda_{r,s}^{\mathrm{dbl}}(n) = \begin{cases} \Theta(n) & s \leqslant 2 \\ \Theta(n\alpha(n)) & s = 3 \\ \Theta(n2^{\alpha(n)}) & s = 4 \\ n \cdot 2^{\alpha^t(n)(\log\alpha(n) + O(1))/t!} & \text{odd } s \geqslant 5 \\ n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))} & \text{even } s \geqslant 6. \end{cases}$$

[6] A set $\mathcal{A}$ is cofinal if, for any σ, there is a $\sigma' \in \mathcal{A}$ such that $\mathrm{Ex}(\sigma, n) = o(\mathrm{Ex}(\sigma', n))$.

The new parts of Theorem 1.3 not covered by previous work [9, 2, 16, 3, 22] are

(i) upper bounds on λ_s^{dbl}, for $s \geqslant 4$, which also cover $\Lambda_{2,s}^{\mathrm{dbl}}$,

(ii) lower bounds on $\Lambda_{r,s}$ for $r \geqslant 3$ and odd $s \geqslant 5$,

(iii) a linear upper bound on $\Lambda_{r,2}^{\mathrm{dbl}}$,

(iv) an $O(n2^{\alpha(n)})$ upper bound on $\Lambda_{r,4}^{\mathrm{dbl}}$, and

(v) an $O(n\alpha(n))$ upper bound on $\Lambda_{r,3}^{\mathrm{dbl}}$, which also covers λ_3^{dbl}.

For task (i) we generalize (and simplify) the recent analysis of [22] to work for double DS sequences. This analysis *only* achieves tight bounds for $s \geqslant 4$. For task (ii) we give a construction of sequences that are $\mathrm{Perm}_{3,s+1}$-free (but necessarily not $\mathrm{Perm}_{2,s+1}$-free) with length $n \cdot 2^{\alpha^t(n)(\log\alpha(n)+O(1))/t!}$. Task (iii) requires no proof. It follows from the linearity of $\mathrm{dbl}(N_k)$-free sequences. For task (iv) we give a single analysis of $\Lambda_{r,s}^{\mathrm{dbl}}$ that is tight for all $r \geqslant 3, s \geqslant 4$, but not $s = 3$. Task (v) is far and away the most difficult to prove. It requires the development of techniques new to the analysis of generalized DS sequences.

Zig-zagging Patterns. Recall that the N- and M-shaped sequences $\{N_k, M_k\}$ generalize $abab = N_1$ and $ababa = M_1$. Define Z_k to be the corresponding generalization of $ababab = Z_1$, that is,

$$Z_k = 1\ 2\ \cdots\ (k+1)\ k\ \cdots\ 1\ 2\ \cdots\ (k+1)\ k\ \cdots\ 1\ 2\ \cdots\ (k+1).$$

We give a flexible new way to construct (and succinctly encode) nonlinear sequences that subsumes nearly all prior constructions [9, 2, 15, 16, 21, 19, 22]. Using the new constructions we are able to show that for any t, there exists a k such that $\mathrm{Ex}(M_k, n) = \Omega(n\alpha^t(n))$ and an l such that $\mathrm{Ex}(Z_l, n) = \Omega(n \cdot 2^{(1+o(1))\alpha^t(n)/t!})$. The bounds on M_k-free sequences are perhaps not too surprising, but they demonstrate that the extremal function for a *set* of forbidden sequences can be different than any member. (Recall that $\mathrm{Ex}(\{M_k, ababab\}, n) = \Theta(n\alpha(n))$ for any k [20].) The new bounds on Z_l show definitively that, in general, zig-zagging sequences are not closely tied to the corresponding DS sequences. In fact, the set $\{Z_l\}$ is cofinal among all forbidden sequences, the other known cofinal sets being $\{(ab)^k\}$ and two-sided combs $\{D_k\}$. Our new sequence constructions also let us show that the one-sided combs $\{C_k\}$ behave differently than $C_1 = abcacbc$, where

$$C_k = 1\ 2\ 3\ \cdots\ (k+2)\ 1\ (k+2)\ 2\ (k+2)\ 3\ (k+2)\ \cdots\ (k+1)\ (k+2).$$

We prove $\mathrm{Ex}(C_k, n) = \Omega(n\alpha^k(n))$.

1.5 Organization

In Section 2 we present sharp lower bounds on $\mathrm{Perm}_{r,s+1}$-free sequences. In Section 3 we review a number of standard sequence transformations and review the linear upper bounds on $\lambda_s, \lambda_s^{\mathrm{dbl}}, \Lambda_{r,s}$, and $\Lambda_{r,s}^{\mathrm{dbl}}$ when $s \in \{1, 2\}$. In Section 4 we establish sharp upper bounds on $\Lambda_{r,s}^{\mathrm{dbl}}$-free sequences, for all $s \geqslant 4$. Section 5 reviews the *derivation tree* structure introduced in [22], which is used in Sections 6 and 7. In Section 6 we present sharp upper bounds on $\Lambda_{r,3}^{\mathrm{dbl}}$ (and λ_3^{dbl}) and in Section 7 we give sharp upper bounds on λ_s^{dbl} for all $s \geqslant 4$. Section 8 is devoted to a new, generalized construction of nonlinear sequences. We prove that, under appropriate parameterization, they are M_k-free, Z_k-free, and C_k-free. Some open problems are discussed in Section 9.

2 Lower Bounds on Perm-Free Sequences

2.1 Composition and Shuffling

We consider sequences made up of blocks, each of which is designated *live* or *dead*. To distinguish the two we use parentheses to indicate live blocks and angular brackets for dead blocks. The number of live blocks in T is $\llparenthesis T \rrparenthesis$ and the number of both types is $\llbracket T \rrbracket$. Our sequences are constructed through *composition* and two types of *shuffling* operations. These operations were implicit in all constructions since Hart and Sharir [9] but were usually presented in an ad hoc manner.

Composition A sequence T over the alphabet $\{1, \ldots, \|T\|\}$ is in *canonical form* if symbols are ordered according to their first appearance in T. All sequences encountered in our construction are assumed to be in canonical form. To *substitute* T for a block $B = (a_1, \ldots, a_{\|T\|})$ means to replace B with a copy of $T(B)$ under the alphabet mapping $k \mapsto a_k$. If T_{mid} is a sequence with $\|T_{\text{mid}}\| = j$ and T_{top} a sequence in which live blocks have length j, $T_{\text{sub}} = T_{\text{top}} \circ T_{\text{mid}}$ is obtained by substituting for each live block B in T_{top} a copy $T_{\text{mid}}(B)$. The live/dead status of a block in T_{sub} is inherited from its status in T_{top} or T_{mid}, hence $\llparenthesis T_{\text{sub}} \rrparenthesis = \llparenthesis T_{\text{top}} \rrparenthesis \cdot \llparenthesis T_{\text{mid}} \rrparenthesis$ and $\llbracket T_{\text{sub}} \rrbracket = \llbracket T_{\text{top}} \rrbracket + \llparenthesis T_{\text{top}} \rrparenthesis (\llbracket T_{\text{mid}} \rrbracket - 1)$. If all symbols appear in μ_{top} live blocks and ν_{top} dead blocks in T_{top}, and μ_{mid} live blocks and ν_{mid} dead blocks in T_{mid}, then the corresponding multiplicities in T_{sub} are $\mu_{\text{top}} \cdot \mu_{\text{mid}}$ and $\nu_{\text{top}} + \mu_{\text{top}} \cdot \nu_{\text{mid}}$.

Shuffling Let $T_{\text{bot}} = (L_1)\ \langle D_1 \rangle\ (L_2)\ \langle D_2 \rangle \cdots (L_l)\ \langle D_l \rangle$ be a sequence with l live blocks $L_1, \ldots, L_l$ and $T_{\text{sub}} = (L'_1)\ \langle D'_1 \rangle\ (L'_2)\ \langle D'_2 \rangle \cdots \left(L'_{l'}\right)\ \left\langle D'_{l'} \right\rangle$ be a sequence whose live blocks $L'_1, \ldots, L'_{l'}$ have length precisely $l = \llparenthesis T_{\text{bot}} \rrparenthesis$. The Ds here represents zero or more dead blocks appearing between live blocks. The *post*shuffle $T_{\text{sh}} = T_{\text{sub}} \oslash T_{\text{bot}}$ is obtained by first forming the concatenation T^*_{bot} of l' copies of T_{bot}, each over an alphabet disjoint from the other copies. A copy of T_{sub} is shuffled into T^*_{bot} as follows. Let $L'_q = (a_1 a_2 \cdots a_l)$ be the qth live block of T_{sub} and $T^{(q)}_{\text{bot}} = \left(L^{(q)}_1\right)\ \left\langle D^{(q)}_1 \right\rangle \cdots \left(L^{(q)}_l\right)\ \left\langle D^{(q)}_l \right\rangle$ be the qth copy of T_{bot} in T^*_{bot}. We substitute the following for $T^{(q)}_{\text{bot}}$, for all q, yielding T_{sh}.

$$\left(L^{(q)}_1 a_1\right)\ \left\langle D^{(q)}_1 \right\rangle \cdots \left(L^{(q)}_l a_l\right)\ \left\langle D^{(q)}_l D'_q \right\rangle$$

In other words, we insert a_p at the end of the pth live block in $T^{(q)}_{\text{bot}}$ and insert all the dead blocks D'_q following L'_q in T_{sub} immediately after $T^{(q)}_{\text{bot}}$. See Figure 1. The *pre*shuffle $T_{\text{sh}} = T_{\text{sub}} \otimes T_{\text{bot}}$ is formed in exactly the same way except that we insert a_p at the *beginning* of the block, that is, we substitute for $T^{(q)}_{\text{bot}}$ the sequence $\left(a_1 L^{(q)}_1\right)\ \left\langle D^{(q)}_1 \right\rangle \cdots \left(a_l L^{(q)}_l\right)\ \left\langle D^{(q)}_l D'_q \right\rangle$. In this section we consider only postshuffling whereas both pre- and postshuffling are used in Section 8.

2.2 Construction of the Sequences

Our $\text{Perm}_{r,s+1}$-free sequences are constructed inductively, beginning with $\text{Perm}_{r,4}$-free sequences $\{T_\rho(i,j)\}_{i \geqslant 1, j \geqslant 0, \rho \geqslant 2}$. Each $T_\rho(i,j)$ consists of a mixture of live and dead blocks. The parameters i and j control the multiplicity of symbols and the length of live blocks, respectively. The length of dead blocks are guaranteed to be a multiple of ρ. This construction is essentially the same as [19], and, ignoring the role of ρ, essentially the same as [9, 15, 30, 21].

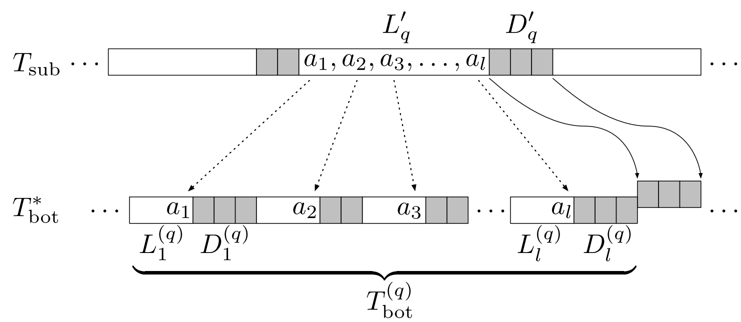

Figure 1: Here $L'_q = (a_1 \cdots a_l)$ is the qth live block of T_{sub} and $T_{\mathrm{mid}}^{(q)}$ is the qth copy of T_{mid} in T^*_{mid}. The sequence $T_{\mathrm{sub}} \oslash T_{\mathrm{mid}}$ is obtained by shuffling L'_q into the live blocks of $T_{\mathrm{mid}}^{(q)}$ and inserting D'_q after $T_{\mathrm{mid}}^{(q)}$.

$$
\begin{aligned}
V(j) &= (1 \cdots j)\ \langle j \cdots 1 \rangle && \text{one live block, one dead}\\
T_\rho(1,j) &= V(j)\\
T_\rho(i,0) &= (\,)^\rho && \rho \geqslant 2 \text{ empty live blocks, for } i \geqslant 2\\
T_\rho(i,j) &= T_{\mathrm{sub}} \oslash T_{\mathrm{bot}} = (T_{\mathrm{top}} \circ T_{\mathrm{mid}}) \oslash T_{\mathrm{bot}}\\
\text{where } T_{\mathrm{bot}} &= T_\rho(i, j-1)\\
T_{\mathrm{mid}} &= V(\llparenthesis T_{\mathrm{bot}} \rrparenthesis)\\
T_{\mathrm{top}} &= T_\rho(i-1, \llparenthesis T_{\mathrm{bot}} \rrparenthesis)
\end{aligned}
$$

Lemma 2.1 identifies some simple properties of $T_\rho(i,j)$ that let us analyze its length and forbidden substructures.

Lemma 2.1. *Let $T = T_\rho(i,j)$ for some $\rho \geqslant 2$.*

1. *Live blocks of T consist solely of first occurrences and all first occurrences appear in live blocks.*
2. *Live blocks of T have length j.*
3. *All symbols appear $i+1$ times in T.*
4. *When $i \geqslant 2$, the number of live blocks and the length of dead blocks are both multiples of ρ.*
5. *As a consequence of Parts 1–3, $|T| = (i+1)\|T\| = (i+1)j\llparenthesis T \rrparenthesis$.*

Proof. All the claims trivially hold in the base cases, when $i = 1$ or $j = 0$. Assume the claim holds inductively for pairs lexicographically smaller than (i,j). Note that Part 1 holds for $T_{\mathrm{mid}} = V(\cdot)$. If it holds for T_{top} and T_{mid} it clearly holds for T_{sub}, and if it holds for T_{bot} as well then it also holds for $T_\rho(i,j) = T_{\mathrm{sub}} \oslash T_{\mathrm{bot}}$.

Part 2 follows since, by the inductive hypothesis, live blocks in $T_{\mathrm{bot}} = T_\rho(i, j-1)$ have length $j-1$ and exactly one symbol gets shuffled into each live block when forming $T_\rho(i,j) = T_{\mathrm{sub}} \oslash T_{\mathrm{bot}}$.

Part 3 follows since the multiplicity of symbols in T_{top} is i, by the induction hypothesis, and the multiplicity in $V(\cdot)$ is 2, so the multiplicity of symbols in T_{sub} is $i+1$. The multiplicity of symbols in T_{bot} is already $i+1$, by the induction hypothesis, so all symbols occur in T with multiplicity $i+1$.

Turning at last to Part 4, the claim is vacuous when $i = 1$ and clearly holds when $i \geqslant 2, j = 0$. In general, if $(\!| T_{\text{bot}} |\!) = (\!| T_\rho(i, j-1) |\!)$ is a multiple of ρ then $(\!| T_\rho(i,j) |\!)$ is also a multiple of ρ. All dead blocks in $T_\rho(i,j)$ are either (i) inherited from T_{bot}, or (ii) inherited from T_{top}, or (iii) are first introduced in T_{sub} as the second block in a copy of $T_{\text{mid}} = V((\!| T_{\text{bot}} |\!))$. The inductive hypothesis implies that the length of category (i) blocks are multiples of ρ. When $i \geqslant 3$ the inductive hypothesis also implies the length of category (ii) blocks are multiples of ρ. When $i = 2$ we have $T_{\text{top}} = T_\rho(1, (\!| T_{\text{bot}} |\!)) = V((\!| T_{\text{bot}} |\!))$. By virtue of $(\!| T_{\text{bot}} |\!)$ being a multiple of ρ, the length of the lone dead block in T_{top} is a multiple of ρ. Category (iii) blocks satisfy the property for the same reason, since $T_{\text{mid}} = V((\!| T_{\text{bot}} |\!))$ and $(\!| T_{\text{bot}} |\!)$ is a multiple of ρ. □

Lemma 2.2. *$T_\rho(i,j)$ is an order-3 DS sequence, and hence $\mathrm{Perm}_{r,4}$-free for all $r \geqslant 2$.*

Proof. The claim clearly holds in all base cases, so we can assume $T = T_\rho(i,j)$ was formed from $T_{\text{top}}, T_{\text{mid}}$, and T_{bot}. Any occurrence of *ababa* could not have arisen from a shuffling event. If $a \in \Sigma(T_{\text{top}})$ and $b \in \Sigma(T^*_{\text{bot}})$, the projection of T onto $\{a, b\}$ is $|b^*ab^*|\, a^*$, where the bars mark the boundary of b's copy of T_{bot}. (The live block of T_{sub} shuffled into b's T_{bot} contains the first occurrence of a. All other as in T_{sub} are inserted after this copy of T_{bot}.) We could also not create an occurrence of *ababa* during a composition event, where a and b shared a live block in T_{top}. The projections of T_{top} and T_{sub} onto $\{a, b\}$ would be, respectively, of the form $(ab)a^*b^*$ and $(ab)\,\langle ba\rangle\, a^*b^*$, the latter being *ababa*-free. □

The $U_s(i,j)$ sequences defined below have the property that all blocks are live and have length exactly j and all symbols occur $\mu_{s,i}$ times, where the μ-values are defined below. This contrasts with $T_\rho(i,j)$, where there is a mixture of live and dead blocks having non-uniform lengths. We define $U_3(i,j)$ to be identical to $T_j(i,j)$ as a sequence, but we interpret it as a sequence of live blocks of length exactly j. This is possible since, in $T_j(i,j)$, the length of live blocks is j and the length of all dead blocks a multiple of j. Since all blocks in U_s are live we can use the identities $[\![U_s(i,j)]\!] = (\!| U_s(i,j) |\!)$ and $|U_s(i,j)| = \mu_{s,i}\|U_s(i,j)\| = j[\![U_s(i,j)]\!]$. Sequences essentially the same as $\{U_s\}$ were used in [19] to prove lower bounds on $\mathrm{Ex}(D_k, n)$, where $\{D_k\}$ are the two-sided combs defined in Section 1.3.

$$
\begin{aligned}
&U_2(i,j) = (1\cdots j)\,(j\cdots 1) && \text{two blocks, for all } i\\
&U_s(i,1) = (1)^{\mu_{s,i}} && \mu_{s,i} \text{ identical blocks, for } i \geqslant 1, s \geqslant 3\\
&U_s(0,j) = (1\cdots j) && \text{one block, for } s \geqslant 3\\
&U_3(i,j) = T_j(i,j) \;\; \text{(reinterpreted)} && \text{for } i \geqslant 1, \text{ where } \rho = j \geqslant 2\\
&U_s(i,j) = U_{\text{sub}} \circledcirc U_{\text{bot}} = (U_{\text{top}} \circ U_{\text{mid}}) \circledcirc U_{\text{bot}} &&\\
&\quad \text{where } U_{\text{bot}} = U_s(i, j-1) &&\\
&\qquad\quad U_{\text{mid}} = U_{s-2}(i, [\![U_{\text{bot}}]\!]) &&\\
&\qquad\quad U_{\text{top}} = U_s(i-1, \|U_{\text{mid}}\|) &&
\end{aligned}
$$

The multiplicities $\{\mu_{s,i}\}$ are defined as follows.

$$\begin{aligned}
\mu_{2,i} &= 2 && \text{for all } i\\
\mu_{3,i} &= i+1 && \text{for all } i\\
\mu_{s,0} &= 1 && \text{for all } s \geqslant 4\\
\mu_{s,i} &= \mu_{s,i-1}\mu_{s-2,i} && \text{for } s \geqslant 4 \text{ and } i \geqslant 1
\end{aligned}$$

Lemma 2.3. *Let $U = U_s(i,j)$, where $s \geqslant 2, i \geqslant 1, j \geqslant 1$.*

1. *All symbols appear in U with multiplicity precisely $\mu_{s,i}$.*
2. *All blocks in U have length precisely j.*
3. *If a and b share a common block and $a < b$ according to the canonical ordering of U, then the projection of U onto $\{a,b\}$ has the form either $a^*b^*(ba)b^*a^*$ or $a^*(ab)a^*b^*$. Moreover, unless $s = 2$, every pair of symbols appear in at most one common block.*

Proof. Parts 1 and 2 hold in the base cases and follow easily by induction on s, i, and j. For Part 3, if b precedes a in their common block then, in some shuffling event, $a \in \Sigma(U_{\text{sub}})$ was *post*shuffled into b's copy of U_{bot} and all other copies of a were placed before or after this copy of U_{bot}, hence U's projection onto $\{a,b\}$ is $a^*b^*(ba)b^*a^*$. If a precedes b in their common block then this must be the *first* occurrence of b in U (otherwise $b < a$ in the canonical ordering). By the same reasoning as above the projection of U onto $\{a,b\}$ must be of the form $a^*(ab)a^*b^*$. □

In Lemma 2.4 we analyze the subsequences avoided by U_s and in Lemma 2.6 we lower bound the length of U_s.

Lemma 2.4. *When $s = 3$ or $s \geqslant 2$ is even, U_s is an order-s DS sequence and hence* $\mathrm{Perm}_{2,s+1}$*-free. When $s \geqslant 5$ is odd and $r \geqslant 3$, U_s is* $\mathrm{Perm}_{r,s+1}$*-free.*

Proof. The claim is clearly true for $s = 2$ and Lemma 2.2 takes care of $s = 3$. Observe that $ababab$ can never be introduced by a shuffling event. If $a \in \Sigma(U_{\text{sub}})$ and $b \in \Sigma(U^*_{\text{bot}})$, only one copy of a can appear between two bs; all others precede or follow b's copy of U_{bot} in U^*_{bot}. Thus any alternating subsequence $ab \cdots ab$ of length $s + 2 \geqslant 6$ must be introduced in $U_{\text{sub}} = U_{\text{top}} \circ U_{\text{mid}}$ by composition. The projection of U_{top} onto $\{a,b\}$ is of the form $a^*b^*(ba)b^*a^*$. Since $U_{\text{mid}} = U_{s-2}(\cdot,\cdot)$ has order $s-2$ and b precedes a in the canonical ordering of U_{mid}, its longest alternating subsequence is $bab \cdots ab$ (length $s-1$), hence the longest alternating subsequence in U_{sub} has length $s+1$.

We now consider $U = U_s(i,j)$, where $s \geqslant 5$ is odd. Recall that U is regarded as a sequence over the alphabet $\{1, \ldots, \|U\|\}$ in canonical form. Generalizing our previous terminology, we will say U is σ-free, where $\Sigma(\sigma) = \{1, \ldots, \|\sigma\|\}$, if U contains no subsequences *order-isomorphic* to σ, that is, that are both isomorphic to σ and preserve the relative order of symbols in σ.[7] Define P_{s+1} to be

[7] For example, 5678 5678 contains several subsequences isomorphic to 2121, but none are order-isomorphic. It contains many subsequences order-isomorphic to 1212 such as 6868. We should point out that the concepts of *canonical form* and *order-isomorphic* were introduced by none other than Davenport and Schinzel [4, p. 691], who noted that order-s DS sequences in canonical form are $(3(12)^{s/2})$-free, for even s, and $(31(21)^{(s-1)/2})$-free, for odd s.

the set of $\sigma \in \{1,2,3\}^*$ such that $\mathrm{dbl}(\sigma)$ contains a subsequence $\sigma_1\sigma_2\cdots\sigma_{s+1}$, where σ_1 and σ_{s+1} are permutations of $\{2,3\}$ and $\sigma_2,\ldots,\sigma_s$ are permutations of $\{1,2,3\}$.[8]

We will prove that $U_s(i,j)$ (in canonical form) is P_{s+1}-free by induction, which implies that $U_s(i,j)$ is $\mathrm{Perm}_{r,s+1}$-free for all $r \geqslant 3$. The claim holds at $s = 3$ since all members of P_4 contain *ababa* as a subsequence, on the alphabet $\{2,3\}$. For $s \geqslant 5$, P_{s+1} could not have arisen from a shuffling event since every member of P_{s+1} contains a sequence isomorphic to *ababab*. It also could not have arisen from a composition even in which some *strict* subset of $\{1,2,3\}$ appears in one block. Whether this subset is $\{1,2\}$ or $\{2,3\}$ or $\{1,3\}$, the 1s can only be involved in two permutations whereas they must be involved in at least four, namely $\sigma_2,\ldots,\sigma_s$.

We can therefore assume that any P_{s+1} sequence over the alphabet $\{a,b,c\}$ arises from a composition event, where a,b,c share a common block B in U_{top}. (For reasons that will become clear shortly, it is better to use symbols a,b,c rather than integers $1,2,3$.) To obtain U_{sub} we substitute for B a copy $U_{\mathrm{mid}}(B)$ of $U_{\mathrm{mid}} = U_{s-2}(\cdot,\cdot)$. Without loss of generality $a < b < c$ according to the canonical ordering of U_{top}. According to Lemma 2.3(3) the projection of U_{top} onto $\{a,b,c\}$, ignoring immediate repetitions, is either

(i) $abc(cba)cba$, or

(ii) $ab(bca)bca$, or

(iii) $ab(bac)bac$, or

(iv) $a(abc)abc$.

That is, in cases (ii)–(iv) B contains the first c in U_{top} and in case (iv) B also contains the first b in U_{top}. In case (i) $c < b < a$ according to the canonical ordering of $U_{\mathrm{mid}}(B)$. In order for U_{sub} to contain a P_{s+1} sequence we would need $U_{\mathrm{mid}}(B)$ to contain

$$\{ab\}\ \overbrace{\{abc\}\cdots\{abc\}}^{s-3}\ \{ab\},$$

where the curly brackets indicate arbitrary permutations of the enclosed sequences. (The $\{ab\}$ permutations on either end can be extended to permutations on $\{abc\}$ by borrowing the cs adjacent to B in U_{top}.) In cases (ii) and (iii), $b < a, c$ according to the canonical ordering of $U_{\mathrm{mid}}(B)$, so for U_{sub} to contain a P_{s+1} sequence, $U_{\mathrm{mid}}(B)$ must contain

$$\{c\}\ \overbrace{\{abc\}\cdots\{abc\}}^{s-2}\ \{ac\}.$$

Once again, the permutations on $\{c\}$ and $\{ac\}$ on either end can be extended to $\{bc\}$ and $\{abc\}$ by borrowing the bs on either side of B. In case (iv) we have $a < b < c$ according to the canonical ordering of $U_{\mathrm{mid}}(B)$, which, by the same reasoning, would need to contain

$$\{bc\}\ \overbrace{\{abc\}\cdots\{abc\}}^{s-2}\ \{bc\}$$

None of cases (i)–(iv) is possible since $U_{\mathrm{mid}} = U_{s-2}$ is P_{s-1}-free, by the induction hypothesis. $\square$

[8] For example, $23\,21\,23\,2 \in P_4$ since doubling the first and second 3 and the first 1 yields a sequence of the desired form.

Remark 2.5. Notice that in the proof of Lemma 2.4, "P_{s-1}-freeness" is defined with respect to the canonical ordering on $\{a,b,c\}$ in U_{mid}, which is identical to their ordering in B. Although $a<b<c$ with respect to U_{top}, identifying a,b, and c with 1,2, and 3 would be confusing as their canonical ordering is typically different in U_{mid}.

We have established that U_s is $\text{Perm}_{r,s+1}$-free and now need to lower bound its length.

Lemma 2.6. *Fix s and let $t=\lfloor (s-2)/2\rfloor$.*

1. *For even s, $\mu_{s,i}=2^{\binom{i+t-1}{t}}=2^{i^t/t!+O(i^{t-1})}$.*
2. *For odd s, $\mu_{s,i}=\prod_{l=0}^{i}(i+1-l)^{\binom{l+t-1}{t-1}}=2^{i^t(\log i)/t!+O(i^t)}$.*

Proof. Consider the even case first. When $i=0$ we have $\mu_{s,0}=1=2^{\binom{0+t-1}{t}}$ and when $s=2,t=0$ we have $\mu_{2,i}=2^{\binom{i+0-1}{0}}=2$. The claim holds for all even $s\geqslant 4$ since, by Pascal's identity, $\mu_{s,i}=\mu_{s,i-1}\cdot\mu_{s-2,i}=2^{\binom{(i-1)+t-1}{t}+\binom{i+(t-1)-1}{t-1}}=2^{\binom{i+t-1}{t}}$. Clearly $2^{\binom{i+t-1}{t}}\geqslant 2^{i^t/t!}$.

For odd s the base case $i=0$ is trivial. When $s=5,t=1$ we have $\mu_{5,i}=\mu_{3,i}\mu_{3,i-1}\cdots\mu_{3,0}=(i+1)!$, which can be expressed as $\prod_{l=0}^{i}(i+1-l)^{\binom{l+t-1}{t-1}}$ since $t=1$ and $\binom{l+0}{0}=1$ for all l. For odd $s\geqslant 7$ the bound follows by induction.

$$
\begin{aligned}
&\mu_{s,i}=\mu_{s,i-1}\cdot\mu_{s-2,i}\\
&=\prod_{l=0}^{i-1}((i-1)+1-l)^{\binom{l+t-1}{t-1}}\cdot\prod_{l'=0}^{i}(i+1-l')^{\binom{l'+t-2}{t-2}}\\
&=\prod_{l''=0}^{i}(i+1-l'')^{\binom{l''+t-2}{t-1}}\cdot\prod_{l'=0}^{i}(i+1-l')^{\binom{l'+t-2}{t-2}}\qquad\{l''\overset{\text{def}}{=}l+1.\text{ When }l''=0,\ (i+1)^{\binom{t-2}{t-1}}=1.\}\\
&=\prod_{l=0}^{i}(i+1-l)^{\binom{l+t-2}{t-1}+\binom{l+t-2}{t-2}}\ =\ \prod_{l=0}^{i}(i+1-l)^{\binom{l+t-1}{t-1}}
\end{aligned}
$$

When s is odd, it is simpler to obtain asymptotic bounds on $\log_2(\mu_{s,i})$ directly, without analyzing the closed-form expression above. Assuming inductively that $\log_2(\mu_{s-2,i})=i^{t-1}(\log i)/(t-1)!+O(i^{t-2})$, where the constant hidden in the second term depends on $s-2$, we have

$$
\begin{aligned}
\log_2(\mu_{s,i})=\log_2(\mu_{s-2,i})+\log_2(\mu_{s,i-1})&=\sum_{x=1}^{i}\log_2(\mu_{s-2,x})\\
&=\sum_{x=1}^{i}\left[\frac{x^{t-1}\log x}{(t-1)!}+O(x^{t-2})\right]\\
&=\frac{i^t\log i}{t!}+O(x^{t-1}).
\end{aligned}
$$

Note that the sum is faithfully approximated by the integral $\int_0^i x^{t-1}(\log x)/(t-1)!+O(x^{t-2})\,\mathrm{d}x=i^t(\log i)/t!+O(i^{t-1})$ as the two differ by $O(i^{t-1})$. □

It is a tedious exercise to show that for $n=\|U_s(i,j)\|$ and $m=[\![U_s(i,j)]\!]$, $i=\alpha(n,m)+O(1)$ and $i=\alpha(n)+O(1)$ when $j=O(1)$. (See [16, 19] for several examples of such calculations.) Lemmas 2.2, 2.4, and 2.6 establish all the lower bounds of Theorem 1.3, with the exception of $\lambda_5(n)=\Omega(n\alpha(n)2^{\alpha(n)})$, which is proved in [22].

Remark 2.7. It should be possible to improve the lower bounds on $\Lambda_{3,s}$, for odd $s \geqslant 5$, by substituting Nivasch's construction of order-3 DS sequences [16, §6] for $T_j(i,j)$ in the definition of $U_3(i,j)$. Nivasch's sequences are roughly twice as long as $T_j(i,j)$, which would lead to a $2^{\binom{i+O(1)}{t}}$ factor improvement in $\mu_{s,i}$, for odd $s \geqslant 5$. The only technical issue is to deal with non-uniform block lengths. In the [16] construction there is no straightforward way to force dead blocks to have lengths that are multiples of some ρ. As a consequence, the block lengths in $U_s(i,j)$ would also be non-uniform, but upper bounded by j.

3 Sequence Transformations and Decompositions

This section reviews some basic results and notation that is used throughout the article, sometimes without direct reference.

3.1 Sparse Versus Blocked Sequences

An m-block sequence can easily be converted to an r-sparse one by removing up to $r-1$ symbols in each block, except the first. This shows, for example, that $\lambda_s(n,m) \leqslant \lambda_s(n) + m - 1$ and $\Lambda^{\mathrm{dbl}}_{r,s}(n,m) \leqslant \Lambda^{\mathrm{dbl}}_{r,s}(n) + (r-1)(m-1)$. However, converting an r-sparse sequence into one with $O(n)$ blocks is, in general, not known to be possible without suffering some asymptotic loss. The following lemma generalizes reductions of Sharir [23] and Pettie [22] to $\lambda^{\mathrm{dbl}}_s, \Lambda_{r,s}$, and $\Lambda^{\mathrm{dbl}}_{r,s}$. In the interest of completeness we include a proof in Appendix A.

Lemma 3.1. *(Cf. Sharir [23], Füredi and Hajnal [7], and Pettie [22].) Define* $\gamma_s, \gamma^{\mathrm{dbl}}_s, \gamma_{r,s}, \gamma^{\mathrm{dbl}}_{r,s} : \mathbb{N} \to \mathbb{N}$ *to be non-decreasing functions bounding the leading factors of* $\lambda_s(n), \lambda^{\mathrm{dbl}}_s(n), \Lambda_{r,s}(n)$, *and* $\Lambda^{\mathrm{dbl}}_{r,s}(n)$, *e.g.,* $\Lambda^{\mathrm{dbl}}_{r,s} \leqslant \gamma^{\mathrm{dbl}}_{r,s}(n) \cdot n$. *The following bounds hold.*

$$\lambda_s(n) \leqslant \gamma_{s-2}(n) \cdot \lambda_s(n,2n) \qquad \lambda^{\mathrm{dbl}}_s(n) \leqslant (\gamma^{\mathrm{dbl}}_{s-2}(n) + 4) \cdot \lambda^{\mathrm{dbl}}_s(n,2n)$$
$$\lambda_s(n) \leqslant \gamma_{s-2}(\gamma_s(n)) \cdot \lambda_s(n,3n) \qquad \lambda^{\mathrm{dbl}}_s(n) \leqslant (\gamma^{\mathrm{dbl}}_{s-2}(\gamma^{\mathrm{dbl}}_s(n)) + 4) \cdot \lambda^{\mathrm{dbl}}_s(n,3n)$$
$$\Lambda_{r,s}(n) \leqslant \gamma_{r,s-2}(n) \cdot \Lambda_{r,s}(n,2n) + 2n \qquad \Lambda^{\mathrm{dbl}}_{r,s}(n) \leqslant (\gamma^{\mathrm{dbl}}_{r,s-2}(n) + O(1)) \cdot \Lambda^{\mathrm{dbl}}_s(n,2n))$$
$$\Lambda_{r,s}(n) \leqslant \gamma_{r,s-2}(\gamma_{r,s}(n)) \cdot \Lambda_{r,s}(n,3n) + 2n \qquad \Lambda^{\mathrm{dbl}}_{r,s}(n) \leqslant (\gamma^{\mathrm{dbl}}_{r,s-2}(\gamma^{\mathrm{dbl}}_{r,s}(n)) + O(1)) \cdot \Lambda^{\mathrm{dbl}}_s(n,3n)),$$

where the $O(1)$ *term in the last two inequalities depends on* r *and* s.

3.2 Reductions Between Perm-Free Sequences and DS Sequences

It is not immediate from the definitions that $\lambda_s(n) = \Theta(\Lambda_{2,s}(n))$ and $\lambda^{\mathrm{dbl}}_s(n) = \Theta(\Lambda^{\mathrm{dbl}}_{2,s}(n))$. These functions are, in fact, asymptotically equivalent. Refer to Appendix A for proof of Lemma 3.2.

Lemma 3.2. *The extremal functions for order-s (double) Davenport-Schinzel sequences and* $\mathrm{Perm}_{2,s+1}$-*free (*$\mathrm{Perm}^{\mathrm{dbl}}_{2,s+1}$-*free) sequences are equivalent up to constant factors. In particular,*

$$\begin{aligned}
\lambda_s(n) &\leqslant \Lambda_{2,s}(n) < 3 \cdot \lambda_s(n) + 2n \\
\lambda_s(n,m) &\leqslant \Lambda_{2,s}(n,m) < 2 \cdot \lambda_s(n,m) + n \\
\lambda^{\mathrm{dbl}}_s(n) &\leqslant \Lambda^{\mathrm{dbl}}_{2,s}(n) < 5 \cdot \lambda^{\mathrm{dbl}}_s(n) + 4n \\
\lambda^{\mathrm{dbl}}_s(n,m) &\leqslant \Lambda^{\mathrm{dbl}}_{2,s}(n,m) < 3 \cdot \lambda^{\mathrm{dbl}}_s(n,m) + 2n
\end{aligned}$$

Given these equivalences, we will only prove upper bounds on λ^{dbl}_s and not discuss $\Lambda^{\mathrm{dbl}}_{2,s}$.

3.3 Linearity at Orders 1 and 2

We bound the length of sequences inductively through the use of recurrences. The induction bottoms out when $s \in \{1,2\}$, so we need to handle these two orders directly. Lemma 3.3 summarizes known linear bounds on $\lambda_s, \lambda_s^{\mathrm{dbl}}, \Lambda_{r,s}$, and $\Lambda_{r,s}^{\mathrm{dbl}}$ when $s \leqslant 2$. A proof of Lemma 3.3 appears in Appendix A.

Lemma 3.3. *At orders $s = 1$ and $s = 2$, the extremal functions $\lambda_s, \lambda_s^{\mathrm{dbl}}, \Lambda_{r,s}$, and $\Lambda_{r,s}^{\mathrm{dbl}}$ obey the following.*

$\lambda_1(n) = n$	$\lambda_1(n,m) = n+m-1$	
$\lambda_2(n) = 2n-1$	$\lambda_2(n,m) = 2n+m-2$	*(Davenport-Schinzel [4])*
$\lambda_1^{\mathrm{dbl}}(n) = 3n-2$	$\lambda_1^{\mathrm{dbl}}(n,m) = 2n+m-2$	*(Dav.-Sch. [5],Klazar [13])*
$\lambda_2^{\mathrm{dbl}}(n) < 8n$	$\lambda_2^{\mathrm{dbl}}(n,m) < 5n+m$	*(Klazar [11], Füredi-Hajnal [7])*
$\Lambda_{r,1}(n) = \Lambda_{r,1}^{\mathrm{dbl}}(n) < rn$	$\Lambda_{r,1}(n,m) = \Lambda_{r,1}^{\mathrm{dbl}}(n,m) < n+(r-1)m$	*(Klazar [10])*
$\Lambda_{r,2}(n) < 2rn$	$\Lambda_{r,2}(n,m) < 2n+(r-1)m$	*(Klazar [10])*
$\Lambda_{r,2}^{\mathrm{dbl}}(n) < 6^r rn$	$\Lambda_{r,2}^{\mathrm{dbl}}(n,m) < 2\cdot 6^{r-1}(n+m/3)$	*(Pettie [20], cf. [14])*

The linear bound on $\Lambda_{r,2}^{\mathrm{dbl}}$ is a consequence of bounds on $\mathrm{dbl}(N_{r-1})$-free sequences [14, 20], though this connection was not noted earlier [3].

3.4 Sequence Decomposition

We adopt and extend the sequence decomposition notation from [22]. This style of decomposition goes back to Hart and Sharir [9] and Agarwal, Sharir, and Shor [2], and has been used many times since then [10, 16, 19, 3]. This notation is used liberally throughout Sections 4–7.

Let S be a sequence over an $n = \|S\|$ letter alphabet consisting of $m = [\![S]\!]$ blocks. (It may be that S avoids some forbidden sequences, but this has no bearing on the decomposition.) A partition of S into $\hat{m}$ intervals $S_1 \cdots S_{\hat{m}}$ is called *uniform* if $m_1 = \cdots = m_{\hat{m}-1}$ are equal powers of two and $m_{\hat{m}}$ may be smaller, where $m_q = [\![S_q]\!]$ is the number of blocks in the qth interval. A symbol is *global* if it appears in multiple intervals and *local* otherwise. Let $\check{S} = \check{S}_1 \cdots \check{S}_{\hat{m}}$ and $\hat{S} = \hat{S}_1 \cdots \hat{S}_{\hat{m}}$ be the projections of S onto local and global symbols, so $|S| = |\check{S}| + |\hat{S}|$. Define $\hat{n} = \|\hat{S}\|$ to be the size of the global alphabet and $\hat{n}_q = \|\hat{S}_q\|$ and $\check{n}_q = \|\check{S}_q\|$ to be number of global and local symbols in $\Sigma(S_q)$, so $n = \hat{n} + \sum_{1 \leqslant q \leqslant \hat{m}} \check{n}_q$.

A global symbol $a \in \Sigma(\hat{S}_q)$ is classified as *first*, *last*, or *middle* if no as appear before S_q, no as appear after S_q, or as appear both before and after S_q.[9] Let $\acute{S}_q, \grave{S}_q, \bar{S}_q \prec \hat{S}_q$ be the projections of $\hat{S}_q$ onto symbols classified as first, last, and middle in $\hat{S}_q$; let $\acute{n}_q, \grave{n}_q$, and $\bar{n}_q$ be the sizes of the alphabets $\Sigma(\acute{S}_q), \Sigma(\grave{S}_q)$, and $\Sigma(\bar{S}_q)$. Define $\acute{S}, \grave{S}$, and $\bar{S}$ to be subsequences of first, last, and middle occurrences, namely

$$\begin{aligned}
\acute{S} &= \acute{S}_1\, \acute{S}_2 \,\cdots\, \acute{S}_{\hat{m}-1} \\
\grave{S} &= \quad\;\; \grave{S}_2 \,\cdots\, \grave{S}_{\hat{m}-1}\, \grave{S}_{\hat{m}} \\
\bar{S} &= \quad\;\; \bar{S}_2 \,\cdots\, \bar{S}_{\hat{m}-1}
\end{aligned}$$

Note that $\hat{S}_1 = \acute{S}_1$ consists solely of first occurrences and $\hat{S}_{\hat{m}} = \grave{S}_{\hat{m}}$ consists solely of last occurrences, so $\bar{S}$ is empty if $\hat{m} = 2$. These notational conventions will be applied to sequences and other objects

[9]Note that if $a \in \Sigma(\hat{S}_q)$ is classified as first, all of the possibly *many* occurrences of a in S_q are "first" occurrences.

defined later. For example, the diacritical marks ˇ, ˆ, ´, `, and ¯ will be applied to objects pertaining to local, global, first, last, and middle symbols, respectively. Moreover, whenever we define a new subsequence of S_q, say $\tilde{S}_q$, quantities and objects pertaining to $\tilde{S}_q$ will be indicated with the same diacritical mark, such as $\tilde{n}_q = \|\tilde{S}_q\|$.

The global contracted sequence $\hat{S}' = B_1 \cdots B_{\hat{m}}$ is obtained by *contracting* each interval $\hat{S}_q$ to a single block B_q consisting of some permutation of $\Sigma(\hat{S}_q)$. Unless specified otherwise, the symbols in B_q are ordered according to their *first* occurrence in $\hat{S}_q$. It follows that $\hat{S}' \prec \hat{S}$, so $\hat{S}'$ inherits any forbidden sequences of $\hat{S}$.

4 Upper Bounds on $\mathrm{Perm}^{\mathrm{dbl}}_{r,s}$-free Sequences

In this section we give recurrences for the extremal functions of $\mathrm{Perm}_{r,s+1}$-free sequences and $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequences. Lemmas 4.4 and 4.5 give closed-form upper bounds on the length of such sequences in terms of Ackermann's function. These bounds on $\Lambda_{r,s}$ and $\Lambda^{\mathrm{dbl}}_{r,s}$ are sharp, except for $\Lambda_{2,s}$ and $\Lambda^{\mathrm{dbl}}_{2,s}$, when $s \geqslant 5$ is odd, and $\Lambda^{\mathrm{dbl}}_{r,3}$, for any $r \geqslant 2$. These exceptions are addressed in Sections 6 and 7.

4.1 A Recurrence for $\Lambda_{r,s}$

In reading the proofs of Recurrences 4.1 and 4.3 one should keep in mind that all extremal functions are superadditive. For example,

$$\Lambda_{r,s}(n_1, m_1) + \Lambda_{r,s}(n_2, m_2) \leqslant \Lambda_{r,s}(n_1 + n_2, m_1 + m_2)$$

Recurrence 4.1. *Define n and m to be the alphabet size and block count parameters. For any $\hat{m} \geqslant 2$, any block partition $\{m_q\}_{1 \leqslant q \leqslant \hat{m}}$, and any alphabet partition $\{\hat{n}\} \cup \{\check{n}_q\}_{1 \leqslant q \leqslant \hat{m}}$, $\Lambda_{r,s}$ obeys the following recurrences, for any fixed $r \geqslant 2, s \geqslant 3$.*

When $\hat{m} = 2$,

$$\Lambda_{r,s}(n, m) \leqslant \sum_{q \in \{1,2\}} \Lambda_{r,s}(\check{n}_q, m_q) + \Lambda_{r,s-1}(2\hat{n}, m)$$

and when $\hat{m} > 2$,

$$\Lambda_{r,s}(n, m) \leqslant \sum_{q=1}^{\hat{m}} \Lambda_{r,s}(\check{n}_q, m_q) \;+\; 2 \cdot \Lambda_{r,s-1}(\hat{n}, m) \;+\; \Lambda_{r,s-2}(\Lambda_{r,s}(\hat{n}, \hat{m}) - 2\hat{n}, m).$$

Proof. We adopt the sequence decomposition notation from Section 3.4. The contribution of local symbols is $\sum_q |\check{S}_q| \leqslant \sum_q \Lambda_{r,s}(\check{n}_q, m_q)$. As each symbol in $\acute{S}_q$ appears at least once after S_q, each $\acute{S}_q$ is a $\mathrm{Perm}_{r,s}$-free sequence, it follows that

$$\sum_{q=1}^{\hat{m}-1} |\acute{S}_q| \leqslant \sum_{q=1}^{\hat{m}-1} \Lambda_{r,s-1}(\acute{n}_q, m_q) \leqslant \Lambda_{r,s-1}\left(\sum_{q=1}^{\hat{m}-1} \acute{n}_q, \sum_{q=1}^{\hat{m}-1} m_q\right) = \Lambda_{r,s-1}(\hat{n}, m - m_{\hat{m}}).$$

A symmetric statement is true for each $\grave{S}_q$, hence the contribution of last occurrences is $\sum_q |\grave{S}_q| \leqslant \Lambda_{r,s-1}(\hat{n}, m - m_1)$. If $\hat{m} = 2$ then we have accounted for all symbols, and by superadditivity $\Lambda_{r,s-1}(\hat{n}, m_1) + \Lambda_{r,s-1}(\hat{n}, m_2) \leqslant \Lambda_{r,s-1}(2\hat{n}, m)$.

If $\hat{m} > 2$ then we must also count middle symbols. Each symbol in $\bar{S}_q$ appears at least once before $\bar{S}_q$ and at least once afterward. This implies that $\bar{S}_q$ is $\mathrm{Perm}_{r,s-1}$-free, hence

$$\begin{aligned}
\sum_q |\bar{S}_q| &\leqslant \sum_q \Lambda_{r,s-2}(\bar{n}_q, m_q) \\
&\leqslant \Lambda_{r,s-2}\left(\sum_q \bar{n}_q, \sum_q m_q\right) && \text{superadditivity} \\
&= \Lambda_{r,s-2}(|\hat{S}'| - 2\hat{n}, m - m_1 - m_{\hat{m}}) && (1) \\
&< \Lambda_{r,s-2}(\Lambda_{r,s}(\hat{n}, \hat{m}) - 2\hat{n}, m) && \hat{S}' \text{ is } \mathrm{Perm}_{r,s+1}\text{-free}
\end{aligned}$$

Equality (1) follows since $\sum_q \bar{n}_q$ counts the number of middle occurrences of symbols in $\hat{S}'$, that is, the length of $\hat{S}'$ less $2\hat{n}$ for first and last occurrences. $\square$

4.2 A Recurrence for $\Lambda^{\mathrm{dbl}}_{r,s}$

Recall that $\Lambda^{\mathrm{dbl}}_{r,s}(n, m)$ was defined to be the extremal function for $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free, m-block sequences over an n-letter alphabet. Here $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$ is the set of sequences over the alphabet $[r] = \{1, \ldots, r\}$ of the form $\sigma_1 \cdots \sigma_{s+1}$, where σ_1 and σ_{s+1} contain one occurrence of each symbol in $[r]$ and $\sigma_2, \ldots, \sigma_s$ contain exactly two occurrences of each symbol in $[r]$.

Remark 4.2. The definition of $\Lambda^{\mathrm{dbl}}_{r,s}$ has one annoying property. Suppose S is a sequence and S' a contracted version of it in which each occurrence of a symbol represents two or more occurrences in S. We would like to say that if S is $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free then S' is $\mathrm{Perm}_{r,s+1}$-free, but this is not strictly true. For example, suppose S' contained the $\mathrm{Perm}_{2,4}$ sequence $ab \mid b(a \mid b)a \mid ab$, where the bars separate the four constituent permutations over $\{a, b\}$ and the parentheses mark the boundaries of one block B in S'. If we substitute aa and bb for all as and bs outside B, and substitute $abab$ for B, we find that S may only contain $aabb\, bb\, (abab)\, aa\, aabb$, which contains no $\mathrm{Perm}^{\mathrm{dbl}}_{2,4}$ sequence. On the other hand, if occurrences in S' represent at least *three* occurrences in S, and symbols in the blocks of S' are sorted according to the *2nd* occurrence in the corresponding subsequence of S, then S' is $\mathrm{Perm}_{r,s+1}$ free if S is $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free.

We can easily "force" blocks in S' to represent at least three corresponding occurrences in the original sequence. Suppose we are given an initial $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequence $S^\star$. Obtain S from $S^\star$ by retaining every other occurrence of each symbol, so S is also $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free and $|S| \geqslant |S^\star|/2$. When bounding $|S|$ inductively we may construct a contracted version S' whose occurrences represent at least two occurrences in S, and hence at least three occurrences in $S^\star$. (One subtlety here is that S' will be a subsequence of $S^\star$, not necessarily S, since we order symbols in the blocks of S' according to their position in $S^\star$.)

In Recurrence 4.3 (and Recurrences 6.1 and 7.4 later on) we use the inference [S is $\mathrm{dbl}(\sigma)$-free] $\rightarrow$ [S' is σ-free], knowing that the bounds we obtain on the given extremal function may be off by a factor of two.

Recurrence 4.3. *Define n and m to be the alphabet size and block count parameters. For any $\hat{m} \geqslant 2$, block partition $\{m_q\}_{1\leqslant q\leqslant \hat{m}}$, and alphabet partition $\{\hat{n}\} \cup \{\check{n}_q\}_{1\leqslant q\leqslant\hat{m}}$, $\Lambda^{\mathrm{dbl}}_{r,s}$ obeys the following recurrences, for any fixed $r \geqslant 2, s \geqslant 3$.*

When $\hat{m} = 2$,

$$\Lambda^{\mathrm{dbl}}_{r,s}(n,m) \;\leqslant\; \sum_{q\in\{1,2\}} \Lambda^{\mathrm{dbl}}_{r,s}(\check{n}_q, m_q) \;+\; \Lambda^{\mathrm{dbl}}_{r,s-1}(2\hat{n}, m) \;+\; 2\hat{n}$$

and when $\hat{m} > 2$,

$$\begin{aligned}\Lambda^{\mathrm{dbl}}_{r,s}(n,m) \;\leqslant\; &\sum_{q=1}^{\hat{m}} \Lambda^{\mathrm{dbl}}_{r,s}(\check{n}_q, m_q) \;+\; \Lambda^{\mathrm{dbl}}_{r,s}(\hat{n},\hat{m}) \;+\; 2\cdot\Lambda^{\mathrm{dbl}}_{r,s-1}(\hat{n}, m)\\ &+\; \Lambda^{\mathrm{dbl}}_{r,s-2}(\Lambda_{r,s}(\hat{n},\hat{m}) - 2\hat{n}, m) \;+\; 2\cdot\Lambda_{r,s}(\hat{n},\hat{m})\end{aligned}$$

Proof. We consider the case when $\hat{m} > 2$ first. Let S be a $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequence. The contribution of local symbols is $\sum_q |\check{S}_q| \leqslant \sum_q \Lambda^{\mathrm{dbl}}_{r,s}(\check{n}_q, m_q)$. If a global symbol appears exactly once in some $\hat{S}_q$ that occurrence is called a *singleton*. Let $\dot{S}$ be the subsequence of $\hat{S}$ consisting of singletons. Clearly $\dot{S}$ can be partitioned into $\hat{m}$ blocks, hence $|\dot{S}| \leqslant \Lambda^{\mathrm{dbl}}_{r,s}(\hat{n},\hat{m})$. Remove all singleton occurrences from $\hat{S}$ and let $\ddot{S}$ be what remains. Classify occurrences in $\ddot{S}_q$ as *first*, *middle*, and *last* according to whether they do not occur before, do not occur after, or occur both before and after interval q in $\hat{S}$ (not in $\ddot{S}$.) Let $\acute{S}, \grave{S}, \bar{S} \prec \ddot{S}$ be the subsequences of first, last, and middle occurrences. Obtain $\acute{S}^-_q$ (and $\grave{S}^-_q$) from $\acute{S}_q$ (and $\grave{S}_q$) by removing the last (and first) occurrence of each symbol, and obtain $\bar{S}^-_q$ from $\bar{S}_q$ by removing both the first and last occurrence of each symbol. It follows that both $\acute{S}^-_q$ and $\grave{S}^-_q$ are $\mathrm{Perm}^{\mathrm{dbl}}_{r,s}$-free, and that $\bar{S}^-_q$ is $\mathrm{Perm}^{\mathrm{dbl}}_{r,s-1}$-free. The contribution of first and last non-singleton occurrences in $\ddot{S}$ is therefore at most

$$\sum_q \Big[\Lambda^{\mathrm{dbl}}_{r,s-1}(\acute{n}_q, m_q) + \acute{n}_q + \Lambda^{\mathrm{dbl}}_{r,s-1}(\grave{n}_q, m_q) + \grave{n}_q\Big] \;\leqslant\; 2\cdot\Big[\Lambda^{\mathrm{dbl}}_{r,s-1}(\hat{n}, m) + \hat{n}\Big].$$

Form $\ddot{S}'$ from $\ddot{S}$ by contracting each interval into a single block. Since $\ddot{S}$ is $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free, $\ddot{S}'$ must be $\mathrm{Perm}_{r,s+1}$. (See Remark 4.2.) Therefore, the contribution of middle non-singleton occurrences is at most

$$\begin{aligned}\sum_q \Big[\Lambda^{\mathrm{dbl}}_{r,s-2}(\bar{n}_q, m_q) + 2\bar{n}_q\Big] &\leqslant \Lambda^{\mathrm{dbl}}_{r,s-2}\left(\sum_q \bar{n}_q, \sum_q m_q\right) + 2\cdot\sum_q \bar{n}_q\\ &= \Lambda^{\mathrm{dbl}}_{r,s-2}(|\ddot{S}'| - 2\hat{n}, m) + 2(|\ddot{S}'| - 2\hat{n})\\ &\leqslant \Lambda^{\mathrm{dbl}}_{r,s-2}(\Lambda_{r,s}(\hat{n},\hat{m}) - 2\hat{n}, m) + 2\cdot\Lambda_{r,s}(\hat{n},\hat{m}) - 4\hat{n}.\end{aligned}$$

When $\hat{m} = 2$ there are no middle occurrences and, in the worst case, no singletons. The total number of first and last occurrences is $(\Lambda^{\mathrm{dbl}}_{r,s-1}(\hat{n}, m_1)+\hat{n})+(\Lambda^{\mathrm{dbl}}_{r,s-1}(\hat{n}, m_2)+\hat{n}) \leqslant \Lambda^{\mathrm{dbl}}_{r,s-1}(2\hat{n}, m)+2\hat{n}$. This concludes the proof of the recurrence. □

Lemma 4.4 gives explicit upper bounds on $\Lambda_{r,s}$ and $\Lambda^{\mathrm{dbl}}_{r,s}$ in terms of inductively defined coefficients $\{\pi_{s,i}, \pi^{\mathrm{dbl}}_{s,i}\}$ and the ith row-inverse of Ackermann's function. One should keep in mind, when reading this lemma and similar lemmas, that we will ultimately substitute $\alpha(n,m) + O(1)$ for i, and that this choice makes the dependence on the block count m negligible.

Lemma 4.4. *Fix parameters $i \geqslant 1$, $r \geqslant 2$, $s \geqslant 3$, and $c \geqslant s-2$. Let n, m be the alphabet size and block count and let j be minimal such that $m \leqslant (a_{i,j})^c$. Then $\Lambda_{r,s}$ and $\Lambda^{\mathrm{dbl}}_{r,s}$ are bounded as follows.*

$$\begin{aligned}\Lambda_{r,s}(n,m) &\leqslant \pi_{s,i}\big(n + O((cj)^{s-2}m)\big)\\ \Lambda^{\mathrm{dbl}}_{r,s}(n,m) &\leqslant \pi^{\mathrm{dbl}}_{s,i}\big(n + O((cj)^{s-2}m)\big),\end{aligned}$$

where the asymptotic notation hides a constant depending only on r. The coefficients $\{\pi_{s,i}, \pi^{\mathrm{dbl}}_{s,i}\}$ are defined as follows.

$$\begin{aligned}\pi_{1,i} = \pi^{\mathrm{dbl}}_{1,i} &= 1\\ \pi_{2,i} &= 2\\ \pi_{s,1} &= 2\pi_{s-1,1} = 2^{s-1}\\ \pi_{s,i} &= 2\pi_{s-1,i} + \pi_{s-2,i}(\pi_{s,i-1} - 2) \qquad (2)\\ \pi^{\mathrm{dbl}}_{2,i} &= 2\cdot 6^{r-1}\\ \pi^{\mathrm{dbl}}_{s,1} &= 2\pi^{\mathrm{dbl}}_{s-1,1} + 1 < (6^{r-1}+1)2^s\\ \pi^{\mathrm{dbl}}_{s,i} &= \pi^{\mathrm{dbl}}_{s,i-1} + 2\pi^{\mathrm{dbl}}_{s-1,i} + (\pi^{\mathrm{dbl}}_{s-2,i} + 2)\pi_{s,i-1} \qquad (3)\end{aligned}$$

The proof is by induction over tuples (s, i, j), where c and r are regarded as fixed. (The base cases when $s \in \{1, 2\}$ follow from Lemma 3.3.) At the base case $i = 1$ we let j be minimal such that $m \leqslant a_{1,j}$. By invoking Recurrence 4.1 with $\hat{m} = 2$ is it easy to show that $\Lambda_{r,s}(n,m) \leqslant \pi_{s,1}(n + O(j^{s-2}m))$, where the constant hidden by the asymptotic notation does not depend on s or c. This also implies that $\Lambda_{r,s}(n,m) \leqslant \pi_{s,1}(n + O((cj)^{s-2}m))$ when j is defined to be minimal such that $m \leqslant a^c_{1,j}$, since $a^c_{1,j} = a_{1,cj} = 2^{cj}$. In the general case, when $i > 1$, we apply Recurrence 4.1 using a uniform block partition with width $w^c = a^c_{i,j-1}$, so

$$\hat{m} = \lceil m/w^c \rceil \leqslant (a_{i,j})^c/(a_{i,j-1})^c = (a_{i-1,w})^c.$$

We invoke the inductive hypothesis with parameters $i, j-1$ on sequences with w^c blocks (namely $\{\check{S}_q\}$). On sequences with m blocks (such as $\acute{S}, \grave{S}$) we invoke the inductive hypothesis with i, j and on sequences with $\hat{m}$ blocks we invoke it with $i-1, w$. The induction goes through smoothly so long as the coefficients $\{\pi_{s,i}, \pi^{\mathrm{dbl}}_{s,i}\}$ are defined as in Lemma 4.4, Eqns. (2,3). See [22, Appendices B and C] for several examples of such proofs in this style.[10]

Lemma 4.5. (Closed Form Bounds) *The ensemble $\{\pi_{s,i}, \pi^{\mathrm{dbl}}_{s,i}\}_{s\geqslant 3, i\geqslant 1}$ satisfies the following,*

[10]For an alternative approach see Nivasch [16, §3]. It differs in two respects. First, it refers to the slowly growing row-inverses of Ackermann's function rather than using the 'j' parameter of Ackermann's function. Second, there is no equivalent to our 'c' parameter in [16], which leads to a system of *two* recurrences, one for the leading factor of the n term, and one for the leading factor of the $j^{s-2}m$ term. For yet another style of analysis, which leads to the same recurrences for $\pi_{s,i}$ and $\pi^{\mathrm{dbl}}_{s,i}$, see Nivasch [16, §4], Cibulka and Kynčl [3, §2], or Sundar [25].

where $t = \lfloor \frac{s-2}{2} \rfloor$.

$$\begin{aligned}
\pi_{3,i} &= 2i+2 \\
\pi_{3,i}^{\text{dbl}} &= \Theta(i^2) \\
\pi_{4,i}, \pi_{4,i}^{\text{dbl}} &= \Theta(2^i) \\
\pi_{5,i}, \pi_{5,i}^{\text{dbl}} &\leqslant 2^i(i+O(1))! \\
\pi_{s,i}, \pi_{s,i}^{\text{dbl}} &\leqslant 2^{\binom{i+O(1)}{t}} && \textit{for even } s > 4 \\
\pi_{s,i}, \pi_{s,i}^{\text{dbl}} &\leqslant 2^{\binom{i+O(1)}{t}\log(2(i+1)/e)} && \textit{for odd } s > 5
\end{aligned}$$

Proof. First consider the case when $s \in \{3,4\}$. Eqn. (2) simplifies to

$$\begin{aligned}
\pi_{3,i} &= 2 + \pi_{3,i-1} \\
\pi_{4,i} &= 2\pi_{3,i} + 2(\pi_{4,i-1} - 2)
\end{aligned}$$

One proves by induction that $\pi_{3,i} = 2i+2$ and $\pi_{4,i} = 10 \cdot 2^i - 4(i+2)$. Using these identities, Eqn. (3) can be simplified to

$$\begin{aligned}
\pi_{3,i}^{\text{dbl}} &= \pi_{3,i-1}^{\text{dbl}} + 2 \cdot (2 \cdot 6^{r-1}) + (1+2)(2i-2) \\
\pi_{4,i}^{\text{dbl}} &\leqslant \pi_{4,i-1}^{\text{dbl}} + 2 \cdot \pi_{3,i}^{\text{dbl}} + (2 \cdot 6^{r-1} + 2)(10 \cdot 2^{i-1} - 4(i+1)).
\end{aligned}$$

A short proof by induction shows $\pi_{3,i}^{\text{dbl}} \leqslant 6\binom{i+1}{2} + 4 \cdot 6^{r-1}(i+1)$ and that $\pi_{4,i}^{\text{dbl}} \leqslant 20(6^{r-1}+2)2^i$. In the general case we have, for $s \geqslant 5$,

$$\begin{aligned}
\pi_{s,i} &\leqslant 2\pi_{s-1,i} + \pi_{s-2,i}\pi_{s,i-1} \\
&= 2\pi_{s-1,i} + \pi_{s-2,i}(2\pi_{s-1,i-1} + \pi_{s-2,i-1}(2\pi_{s-1,i-2} + \pi_{s-2,i-2}(\cdots + \pi_{s-2,2}\pi_{s,1})\cdots)) \\
&= \sum_{l=0}^{i-2} 2\pi_{s-1,i-l} \cdot \prod_{k=0}^{l-1} \pi_{s-2,i-k} \;+\; \pi_{s,1} \cdot \prod_{k=0}^{i-2} \pi_{s-2,i-k} \qquad (4)
\end{aligned}$$

When $s = 5$ we have $\pi_{s-1,i} = \Theta(2^i)$ and $\pi_{s-2,i} = 2(i+1)$, so (4) can be written

$$\begin{aligned}
&= \sum_{l=0}^{i-2} \Theta(2^{i-l}) \cdot 2(i+1)2i \cdots 2(i+2-l) \;+\; \pi_{s,1} \cdot 2(i+1)2i2(i-1)\cdots 2(3) \\
&= \Theta(2^i \cdot (i+1)!) \;=\; 2^{(i+O(1))\log(2(i+1)/e)}
\end{aligned}$$

We prove that there are constants $\{C_s\}$ such that $\pi_{s,i} \leqslant 2^{\binom{i+C_s}{t}}$ when s is even and $\pi_{s,i} \leqslant 2^{\binom{i+C_s}{t}\log(2(i+1)/e)}$ when s is odd. The analysis above shows that C_4 and C_5 exist. When $s > 4$ is even, (4) is bounded by

$$\leqslant \sum_{l=0}^{i-2} 2^{\binom{i-l+C_{s-1}}{t-1}\log(2(i-l+1)/e)} \cdot \prod_{k=0}^{l-1} 2^{\binom{i-k+C_{s-2}}{t-1}} \;+\; \pi_{s,1} \cdot \prod_{k=0}^{i-2} 2^{\binom{i-k+C_{s-2}}{t-1}} \qquad (5)$$

By Pascal's identity $\sum_{k=0}^{x}\binom{i-k+C_{s-2}}{t-1} = \binom{i+1+C_{s-2}}{t} - \binom{i-x+C_{s-2}}{t}$, so (5) is bounded by

$$\begin{aligned}
&\leqslant 2^{\binom{i+1+C_{s-2}}{t}} \cdot \left(\sum_{l=0}^{i-2} 2^{\binom{i-l+C_{s-1}}{t-1}\log(2(i-l+1)/e) - \binom{i-l+1+C_{s-2}}{t}} \ + \ \pi_{s,1}\right) && (6)\\
&\leqslant 2^{\binom{i+1+C_s}{t}}, \ \text{ for some sufficiently large } C_s.
\end{aligned}$$

The sum in (6) clearly converges as $i \to \infty$, though for some constant values of $i-l$ (depending on C_{s-1} and C_{s-2}), $\binom{i-l+C_{s-1}}{t-1}\log(2(i-l+1)/e)$ may be significantly larger than $\binom{i-l+1+C_{s-2}}{t}$. When $s>5$ is odd the calculations are similar. By the inductive hypothesis, (4) is bounded by

$$\begin{aligned}
&\leqslant \sum_{l=0}^{i-2} 2^{\binom{i-l+C_{s-1}}{t}} \cdot \prod_{k=0}^{l-1} 2^{\binom{i-k+C_{s-2}}{t-1}\log(2(i-k+1)/e)} + \pi_{s,1} \cdot \prod_{k=0}^{i-2} 2^{\binom{i-k+C_{s-2}}{t-1}\log(2(i-k+1)/e)} && (7)\\
&\leqslant 2^{\binom{i+1+C_{s-2}}{t}\log(2(i+1)/e)} \cdot \left(\sum_{l=0}^{i-2} 2^{\binom{i-l+C_{s-1}}{t} - \binom{i-l+1+C_{s-2}}{t}\log(2(i+1)/e)} + \pi_{s,1}\right)\\
&\leqslant 2^{\binom{i+1+C_s}{t}\log(2(i+1)/e)}, \ \text{ for some sufficiently large } C_s.
\end{aligned}$$

Turning to $\pi^{\mathrm{dbl}}_{s,i}$, we have

$$\begin{aligned}
\pi^{\mathrm{dbl}}_{s,i} &= \pi^{\mathrm{dbl}}_{s,i-1} \ + \ 2\pi^{\mathrm{dbl}}_{s-1,i} + (\pi^{\mathrm{dbl}}_{s-2,i} + 2)\pi_{s,i-1}\\
&= \pi^{\mathrm{dbl}}_{s,1} \ + \ \sum_{l=0}^{i-2}\left[2\pi^{\mathrm{dbl}}_{s-1,i-l} + (\pi^{\mathrm{dbl}}_{s-2,i-l} + 2)\pi_{s,i-1-l}\right] && (8)
\end{aligned}$$

It is straightforward to show that when $s \geqslant 4$, the bounds on $\pi_{s,i}$ also hold for $\pi^{\mathrm{dbl}}_{s,i}$ with respect to different constants $\{D_s\}$. When $s=5$, Eqn. (8) becomes

$$\begin{aligned}
\pi^{\mathrm{dbl}}_{5,i} &= \pi^{\mathrm{dbl}}_{5,1} \ + \ \sum_{l=0}^{i-2}\Big(2\cdot\Theta(2^{i-l}) \ + \ (\Theta(i-l)^2) + 2)\cdot\Theta(2^{i-1-l}(i-l)!)\Big)\\
&= \Theta(2^i(i+2)!) \ \leqslant \ 2^{(i+D_5)\log(2(i+1)/e)}, \ \text{ for a sufficiently large } D_5.
\end{aligned}$$

When $s>4$ is even, Eqn. (8) implies, by the inductive hypothesis, that

$$\begin{aligned}
\pi^{\mathrm{dbl}}_{s,i} &\leqslant \pi^{\mathrm{dbl}}_{s,1} \ + \ \sum_{l=0}^{i-2}\left[2^{\binom{i-l+D_{s-1}}{t-1}\log(2(i-l+1)/e)+1} \ + \ (2^{\binom{i-l+D_{s-2}}{t-1}} + 2)2^{\binom{i-1-l+C_s}{t}}\right]\\
&\leqslant 2^{\binom{i+l+D_s}{t}}, \ \text{ for a sufficiently large } D_s.
\end{aligned}$$

When $s>5$ is odd,

$$\begin{aligned}
\pi^{\mathrm{dbl}}_{s,i} &\leqslant \pi^{\mathrm{dbl}}_{s,1} \ + \ \sum_{l=0}^{i-2}\left[2^{\binom{i-l+D_{s-1}}{t}+1} \ + \ (2^{\binom{i-l+D_{s-2}}{t-1}\log(2(i-l+1)/e)} + 2)2^{\binom{i-1-l+C_s}{t}\log(2(i-l)/e)}\right]\\
&\leqslant 2^{\binom{i+D_s}{t}\log(2(i+1)/e)}, \ \text{ for a sufficiently large } D_s.
\end{aligned}$$

$\square$

Given that Lemma 4.5 holds for all i, one chooses i to be minimum such that the 'm' term does not dominate, that is, the minimum i for which $j \leqslant 3$ or $(cj)^{s-2} \leqslant n/m$. It is straightforward to show that $i = \alpha(n,m) + O(1)$ is optimal, which immediately gives bounds on $\Lambda_{r,s}(n,m)$ and $\Lambda^{\mathrm{dbl}}_{r,s}(n,m)$ analogous to those claimed for $\Lambda_{r,s}(n)$ and $\Lambda^{\mathrm{dbl}}_{r,s}(n)$ in Theorem 1.3, excluding the case $s = 3$, which is dealt with in Section 6. In order to obtain bounds on $\Lambda_{r,s}(n)$ and $\Lambda^{\mathrm{dbl}}_{r,s}(n)$ we invoke Lemma 3.1. For example, it states that $\Lambda_{r,s}(n) = \gamma_{r,s-2}(\gamma_{r,s}(n)) \cdot \Lambda_{r,s}(n,3n)) + 2n$, where $\gamma_{r,s}(n)$ is a non-decreasing upper bound on $\Lambda_{r,s}(n)/n$. The $\gamma_{r,s-2}(\gamma_{r,s}(n))$ factor may not be constant, but it does not affect the error tolerance already in the bounds of Theorem 1.3.[11]

Remark 4.6. Our lower and upper bounds on $\Lambda_{r,s}(n)$ are tight (when $r \geqslant 3$) inasmuch as they are both of the form $n \cdot 2^{\alpha^t(n)/t! + O(\alpha^{t-1}(n))}$ when $s \geqslant 4$ is even and $n \cdot 2^{\alpha^t(n)(\log\alpha(n) + O(1))/t!}$ when $s \geqslant 5$ is odd. However, it is only when s is even that these bounds are sharp in the Ackermann-invariant sense of [22, Remark 1.1], that is, invariant under $\pm O(1)$ perturbations in the definition of $\alpha(n)$. For example, our lower and upper bounds on $\Lambda_{r,5}(n)$ are $n\cdot(\alpha(n)+O(1))!$ and $n\cdot 2^{\alpha(n)}(\alpha(n)+O(1))!$. The $2^{\alpha(n)}$ factor gap could probably be closed by substituting Nivasch's construction of order-3 DS sequences [16, §6] for $U_3(i,j)$ in Section 2, which would lead to sharp, Ackermann-invariant bounds of $\Lambda_{r,5}(n) = n \cdot 2^{\alpha(n)}(\alpha + O(1))!$. With a more careful analysis of the recurrence for $\pi_{s,i}$ it should be possible to obtain sharp, Ackermann-invariant bounds on $\Lambda_{r,s}(n)$ for all odd s.

5 Derivation Trees

Derivation trees were introduced in [22] to model hierarchical decompositions of sequences. They are instrumental in our analysis of $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free sequences, in Section 6, and of double DS sequences, in Section 7. Throughout this section we use the sequence decomposition notation defined in Section 3.4.

A recursive decomposition of a sequence S can be represented as a rooted *derivation tree* $\mathcal{T} = \mathcal{T}(S)$. Nodes of $\mathcal{T}$ are identified with blocks. The leaves of $\mathcal{T}$ correspond to the blocks of S whereas internal nodes correspond to blocks of derived sequences. Let $\mathcal{B}(v)$ be the block of $v \in \mathcal{T}$, which may be treated as a *set* of symbols if we are indifferent to their permutation in $\mathcal{B}(v)$.

Base Case. Suppose $S = B_1B_2$ is a two block sequence, where each block contains the whole alphabet $\Sigma(S)$. The tree $\mathcal{T}(S)$ consists of three nodes u, u_1, and u_2, where u is the parent of u_1 and u_2, $\mathcal{B}(u_1) = B_1$, $\mathcal{B}(u_2) = B_2$, and $\mathcal{B}(u)$ does not exist. For every $a \in \Sigma(S)$ call u its *crown* and u_1 and u_2 its left and right *heads*, respectively. These nodes are denoted $\mathrm{cr}_{|a}, \mathrm{lh}_{|a}$, and $\mathrm{rh}_{|a}$.

Inductive Case. If S contains $m > 2$ blocks, choose a uniform block partition $\{m_q\}_{1\leqslant q\leqslant \hat{m}}$, that is, one where $m_1, \ldots, m_{\hat{m}-1}$ are equal powers of two and $m_{\hat{m}}$ may be smaller. This block partition induces local sequences $\{\check{S}_q\}_{1\leqslant q\leqslant \hat{m}}$ and an $\hat{m}$-block contracted global sequence $\hat{S}'$. Inductively construct derivation trees $\hat{\mathcal{T}} = \mathcal{T}(\hat{S}')$ and $\{\check{\mathcal{T}}_q\}_{1\leqslant q\leqslant \hat{m}}$, where $\check{\mathcal{T}}_q = \mathcal{T}(\check{S}_q)$. To obtain $\mathcal{T}(S)$, identify the root of $\check{\mathcal{T}}_q$ (which has no block) with the qth leaf of $\hat{\mathcal{T}}$, then place the blocks of S at the leaves of $\mathcal{T}$. This last step is necessary since only local symbols appear in the blocks of $\{\check{\mathcal{T}}_q\}$ whereas the

[11] For example, when $s = 6$, $\gamma_{r,s-2}(\gamma_{r,s}(n)) = O\left(2^{\alpha\left(2^{\alpha^2(n)/2 + O(\alpha(n))}\right)}\right) = O(2^{\alpha(\alpha(n))})$ is non-constant. Nonetheless $O(2^{\alpha(\alpha(n))}) \cdot \Lambda_{r,s}(n,3n) = O(2^{\alpha(\alpha(n))}) \cdot n \cdot 2^{\alpha^2(n)/2 + O(\alpha(n))} = n \cdot 2^{\alpha^2(n)/2 + O(\alpha(n))}$.

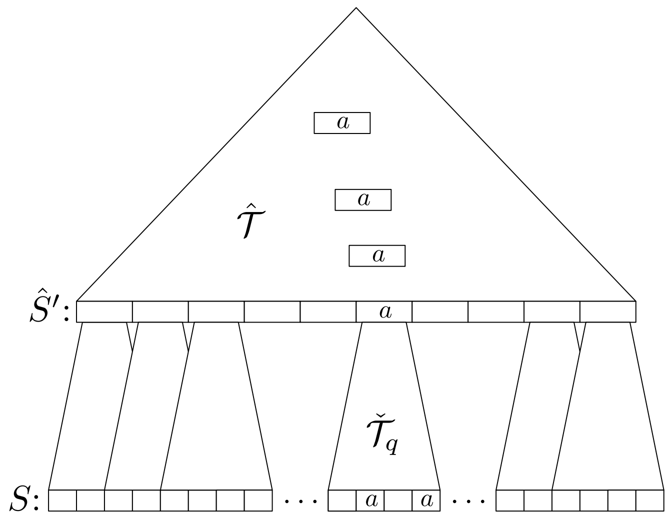

Figure 2: The derivation tree $\mathcal{T}(S)$ is the composition of $\hat{\mathcal{T}} = \mathcal{T}(\hat{S}')$ and $\{\check{\mathcal{T}}_q\}_{1 \leqslant q \leqslant \hat{m}}$, where $\check{\mathcal{T}}_q = \mathcal{T}(\check{S}_q)$. A global symbol $a \in \Sigma(\hat{S})$ appears in blocks at the leaf level of $\mathcal{T}$, at the leaf level of $\hat{\mathcal{T}}$, and possibly at higher levels of $\hat{\mathcal{T}}$.

leaves of $\mathcal{T}$ must be identified with the blocks of S. The crown and heads of each symbol $a \in \Sigma(S)$ are inherited from $\hat{\mathcal{T}}$, if a is global, or some $\check{\mathcal{T}}_q$ if a is local to S_q. See Figure 2 for a schematic.

5.1 Special Derivation Trees

It is useful to constrain $\mathcal{T}$ to use a uniform block partition. Every derivation tree generated in this fashion can be embedded in a full rooted binary tree with height $\lceil \log m \rceil$, though the composition of blocks depends on how block partitions are chosen. We will generate two varieties of derivation trees. At one extreme is the *canonical* derivation tree, where block partitions are chosen in the least aggressive way possible. At the other extreme is one where block partitions are guided by Ackermann's function.

Canonical Derivation Trees. The canonical derivation tree $\mathcal{T}^\star(S)$ of a sequence S is obtained by choosing the uniform block partition with $\hat{m} = \lceil m/2 \rceil$. We form $\mathcal{T}^\star(S)$ by constructing $\mathcal{T}^\star(\hat{S}')$ recursively and composing it with the trivial three-node base case trees $\{\mathcal{T}(\check{S}_q)\}_q$.

Derivation Trees via Ackermann's Function. Given a parameter $i \geqslant 1$, define $j \geqslant 1$ to be minimal such that $m \leqslant a_{i,j}$. If $j = 1$ then $m = a_{i,1} = 2$, meaning $\mathcal{T}(S)$ must be the three-node base case tree. When $j > 1$ we choose a uniform block partition with width $w = a_{i,j-1}$ (which is a power of 2), so $\hat{m} = \lceil m/w \rceil \leqslant a_{i,j}/a_{i,j-1} = a_{i-1,w}$. The global tree $\hat{\mathcal{T}}$ is constructed recursively with parameter[12] $i-1$ and each local tree $\check{\mathcal{T}}_q$ is constructed recursively with parameter i.

[12]Note that when $i = 1$ it does not matter that $i - 1 = 0$ is an invalid parameter. In this case $w = a_{1,j-1} = a_{1,j}/2$ and $\hat{m} = 2$, so $\hat{\mathcal{T}}$ is forced to be a three-node base case tree.

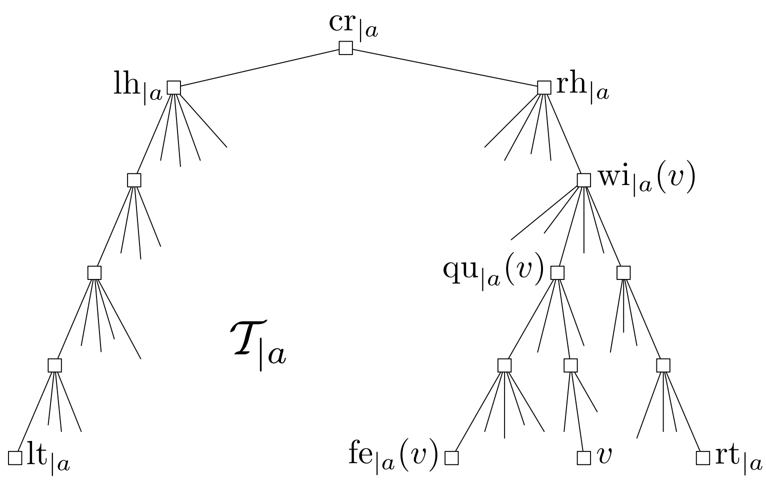

Figure 3: In this example v is a hawk leaf in $\mathcal{T}_{|a}$ since it is a descendant of $\mathrm{rh}_{|a}$. Its wing node $\mathrm{wi}_{|a}(v)$, quill $\mathrm{qu}_{|a}(v)$, and feather $\mathrm{fe}_{|a}(v)$ are indicated.

5.2 Projections of the Derivation Tree

The *projection of* $\mathcal{T}$ *onto* $a \in \Sigma(S)$, written $\mathcal{T}_{|a}$, is the tree rooted at $\mathrm{cr}_{|a}$ on the node set $\{\mathrm{cr}_{|a}\} \cup \{v \in \mathcal{T} \mid a \in \mathcal{B}(v)\}$. The edges of $\mathcal{T}_{|a}$ represent paths in $\mathcal{T}$ passing through blocks that do not contain a.

Definition 5.1. (Anatomy of a projection tree)

- The leftmost and rightmost leaves of $\mathcal{T}_{|a}$ are *wingtips*, denoted $\mathrm{lt}_{|a}$ and $\mathrm{rt}_{|a}$.
- The left and right *wings* are those paths in $\mathcal{T}_{|a}$ extending from $\mathrm{lh}_{|a}$ to $\mathrm{lt}_{|a}$ and from $\mathrm{rh}_{|a}$ to $\mathrm{rt}_{|a}$.
- Descendants of $\mathrm{lh}_{|a}$ and $\mathrm{rh}_{|a}$ in $\mathcal{T}_{|a}$ are called *doves* and *hawks*, respectively.
- A child of a wing node that is not itself on the wing is called a *quill*.
- A leaf is called a *feather* if it is the rightmost descendant of a dove quill or leftmost descendant of a hawk quill.
- Suppose v is a node in $\mathcal{T}_{|a}$. Let $\mathrm{wi}_{|a}(v)$ be the nearest wing node ancestor of v, $\mathrm{qu}_{|a}(v)$ the quill ancestral to v, and $\mathrm{fe}_{|a}(v)$ the feather descending from $\mathrm{qu}_{|a}(v)$. See Figure 3 for an illustration.

If $\mathcal{T}(S)$ is specified, the terms *feather* and *wingtip* can also be applied to individual occurrences in S. For example, an occurrence of a in block $\mathcal{B}(v)$ of S is a feather if v is a feather in $\mathcal{T}_{|a}$.

When $\mathcal{T}(S)$ is constructed according to Ackermann's function, a short proof by induction shows that the height of each projection tree $\mathcal{T}_{|a}$ (distance from $\mathrm{cr}_{|a}$ to a leaf) is at most $i + 1$.

6 Upper Bounds on $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free Sequences

Since order-3 DS sequences are necessarily $\mathrm{Perm}_{2,4}$-free, we have $\Lambda^{\mathrm{dbl}}_{r,3}(n) \geqslant \Lambda_{r,3}(n) \geqslant \lambda_3(n) = \Theta(n\alpha(n))$. In this Section we prove tight upper bounds of $\Lambda^{\mathrm{dbl}}_{r,3}(n) = O(n\alpha(n))$. These bounds imply $\lambda^{\mathrm{dbl}}_3(n)$ is also $O(n\alpha(n))$, resolving one of Klazar's open problems [13].

Our analysis is different in character from all previous analyses of (generalized) Davenport-Schinzel sequences. There are two new techniques used in the proof which are worth highlighting. Previous analyses partition the symbols in a block based on some attributes (first, middle, last, etc.), but do not assign any attributes to the blocks themselves. In our analysis we must treat blocks differently based on their context within the larger sequence, that is, according to properties that are independent of the contents of the block. (See the definition of *roosts* in Section 6.2.) The second ingredient is an accounting scheme for bounding the proliferation of symbols. Rather than count the *number* of occurrences of a symbol, say b, we assign each occurrence of b a *potential* based on its context. If one b in $\hat{S}'$ begets multiple bs in $\hat{S}$, the number of bs increases, but the *aggregate* potential of the bs in S may, in fact, be at most the potential of the originating b in $\hat{S}'$. That is, sometimes proliferating symbols "pay for themselves." We only need to track changes in sequence potential, not sequence length. Amortizing the analysis in this way lets us account for the proliferation of symbols across *many* levels of the derivation tree, not just between $\hat{S}'$ and S.

6.1 A Potential-Based Recurrence

Fix a $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free sequence Z and $i^\star \geqslant 1$. Define $j^\star$ to be minimal such that its block count $[\![Z]\!] \leqslant a_{i^\star,j^\star}$ and let $\mathcal{T} = \mathcal{T}(Z)$ be constructed as in Section 5.1 with parameter $i^\star$. In this section we analyze a sequence S encountered in the recursive decomposition of Z, that is, S is either Z itself or a sequence encountered when recursively decomposing $\hat{Z}'$ and $\{\check{Z}_q\}$. Since $S \prec Z$, it too must be $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free but we can often say something stronger. If each occurrence of a symbol in S represents at least two occurrences in Z then S must be $\mathrm{Perm}_{r,4}$-free.[13] Call an occurrence in S *terminal* if it represents exactly one occurrence in Z and *non-terminal* otherwise. In terms of the derivation tree, an occurrence of a in S is terminal iff it has exactly one leaf descendant in $\mathcal{T}_{|a}$.

Each occurrence of a symbol in S carries a nonnegative integer potential based on its context within S and even within $\mathcal{T}(Z)$. Since the length of S is no more than its aggregate potential, it suffices to upper bound the potential. Define $\Upsilon(n,m)$ to be the maximum potential of an m-block sequence over an n-letter alphabet encountered in decomposing Z. The way potentials are assigned will be discussed shortly. For the time being it suffices to know that the maximum potential is $\phi = O(1)$, all terminals carry unit potential, and all non-terminals carry potential at least three.

Our goal is to prove that Υ obeys the following recurrence.

Recurrence 6.1.

$$\Upsilon(n,m) = \sum_{1\leqslant q\leqslant \hat{m}} \Upsilon(\check{n}_q, m_q) \;+\; 2\cdot\left[\phi\cdot\Lambda_{r,2}(\hat{n},m) \;+\; \Lambda^{\mathrm{dbl}}_{r,2}(\hat{n},m) + \hat{n}\right] \;+\; \Upsilon(\hat{n},\hat{m})$$
$$+\; (r-1)\phi\cdot m \;+\; 2[(r-1)(i^\star-2)]^2\cdot\hat{m}$$

Decomposing S as usual, it follows that the maximum potential of local sequences $\{\check{S}_q\}_q$ is $\sum_q \Upsilon(\check{n}_q, m_q)$, giving the first term of Recurrence 6.1. The sequence $\acute{S}$ of global first occurrences

[13] This is not quite true, but we can make this inference when bounding $\Lambda^{\mathrm{dbl}}_{r,3}$ asymptotically. See Remark 4.2 for a discussion of this issue.

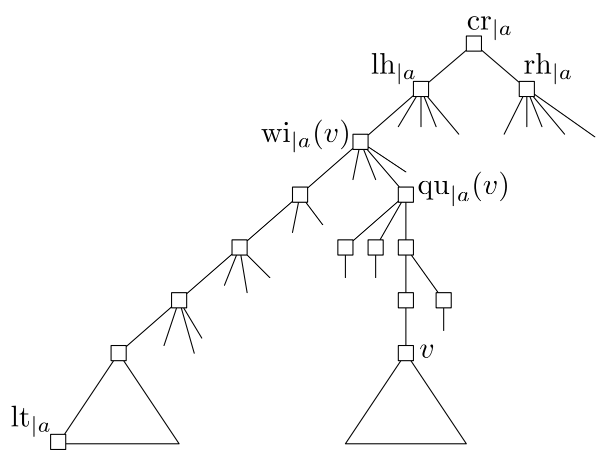

Figure 4: Here v is an internal node of $\mathcal{T}_{|a}$. Between $\mathrm{qu}_{|a}(v)$ and v, a has molted twice: at v's parent it molted one a to the right and at v's grandparent it molted two as to the left.

can be partitioned into terminals $\acute{S}^{\mathrm{t}}$ and non-terminals $\acute{S}^{\mathrm{nt}}$. After removing the last occurrence of each symbol in $\acute{S}^{\mathrm{t}}$, the resulting sequence is $\mathrm{Perm}_{r,3}^{\mathrm{dbl}}$-free, so its length (and potential) is $|\acute{S}^{\mathrm{t}}| \leqslant \Lambda_{r,2}^{\mathrm{dbl}}(\hat{n}, m) + \hat{n}$. We endow each non-terminal in $\acute{S}^{\mathrm{nt}}$ an initial potential at most ϕ. (Note that occurrences of a in $\acute{S}$ correspond to quills in $\mathcal{T}_{|a}$.) Being $\mathrm{Perm}_{r,3}$-free, the potential of $\acute{S}^{\mathrm{nt}}$ is therefore at most $\phi \cdot \Lambda_{r,2}(\hat{n}, m)$. A symmetric analysis is applied to $\grave{S}$, the sequences of last occurrences, which gives the second term of Recurrence 6.1.

The global contracted sequence $\hat{S}'$ begets $\acute{S}, \grave{S}$, and $\bar{S}$, the first two of which we have just accounted for. In general $|\bar{S}|$ may be significantly larger than $|\hat{S}'|$. We account for this proliferation in symbols by showing that the aggregate potential of $\bar{S}$ is nonetheless at most that of $\hat{S}'$ plus $(r-1)\phi \cdot m + 2[(r-1)(i^\star - 1)]^2 \cdot \hat{m}$, which explains the last three terms of Recurrence 6.1. Consider the sequence $\bar{S}_q$ begat by the middle symbols of block B_q in $\hat{S}'$. We decompose $\bar{S}_q$ as follows.

1. Tag any symbol occurring exactly once in $\bar{S}_q$. (Its potential in $\bar{S}_q$ will be at most its potential in $\hat{S}'$.)

2. Tag the first non-terminal occurrence of each symbol in $\bar{S}_q$.

3. Tag the first, second, and last terminal occurrence of each symbol in $\bar{S}_q$.

4. Tag the first $r-1$ untagged occurrences (terminal and non-terminal) in each block of $\bar{S}_q$.

Symbols that are tagged in both of Steps 2 and 3 have *molted*; all others are *unmolted*. We will say that the non-terminal a tagged in Step 2 has *molted* those terminal as tagged in Step 3. See Figure 4 for a schematic.

We claim $\bar{S}_q$ has been completely tagged after Step 4. If this were not so, there must be r symbols $a_1, \ldots, a_r$ in some block B in $\bar{S}_q$. If a_k is terminal in B it must be preceded by two terminal a_ks and followed by one terminal a_k in $\bar{S}_q$; if a_k is non-terminal in B it must be preceded by a non-terminal a_k. Dividing $\bar{S}_q$ at the left boundary of B, we see two occurrences of each of $a_1, \ldots, a_r$ on both the left and right side of the boundary, which may take the form of one non-terminal or two terminals. Since $a_1, \ldots, a_k$ are categorized as global middle in S_q, each appears both before and after S_q, yielding an instance of $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$ in Z, a contradiction.

The aggregate potential of those symbols tagged in Step 4 is at most $(r-1)\phi \cdot m$, which are covered by the second-to-last term of Recurrence 6.1. Suppose that $a \in B_q$ is non-terminal in $\hat{S}'$ but it begets only terminal as in $\bar{S}_q$, that is, no as are tagged in Step 2. This proliferation of as causes no net increase in potential since the $a \in B_q$ carries potential at least 3, which covers the potential of the three terminal as tagged in Step 3. In general, for each molted symbol a, we will tag one non-terminal and up to three terminals in Steps 2 and 3. This will cause no net increase in potential *provided* that the a in B_q carries at least the potential of the non-terminal a in $\bar{S}_q$ plus 3. In order to avoid cumbersome statements, we will treat the non-terminal a tagged in Step 2 as the "same" $a \in B_q$. For example, if B is a block in $\bar{S}_q$ and $a \in B$ is non-terminal, to say *the* $a \in B$ *has molted four times* means that, in $\mathcal{T}_{|a}$, B has four ancestors, possibly including itself, and all strict descendants of $\mathrm{qu}_{|a}(B)$, which each have at least one sibling in $\mathcal{T}_{|a}$. This sibling corresponds to an a removed in Step 3 at some stage in the decomposition of S.

In the remainder of this section we explain why it suffices to endow each new non-terminal quill with a *constant* potential ϕ. The analysis above shows that $3 \cdot (i^\star - 1)$ suffices, which is not constant.[14]

6.2 Roosts, Eggs, and Fertility

Our analysis considers properties of blocks (and of occurrences of symbols) that depend on their context within a larger sequence.

Definition 6.2. (Roosts and Eggs) Let S be a sequence encountered in the decomposition of Z.

1. An interval I of zero or more blocks in S is a *k-roost* if there are k distinct symbols $a_1, \ldots, a_k$ such that the sequence contains

$$a_1\, a_2 \cdots a_k \;\; a_k^2 a_{k-1}^2 \cdots a_1^2 \;\; I \;\; a_1^2 a_2^2 \cdots a_k^2 \;\; a_k a_{k-1} \cdots a_1,$$

   where b^2 refers to two terminal bs or one non-terminal b. The occurrences of a_1 just to the left and right of I are called *k-left mature* and *k-right mature*. A k-mature occurrence of a symbol whose block is a k-roost is *infertile*. A k-left mature occurrence that is not infertile is k-left fertile; k-right fertile is defined analogously. (For any $l < k$, k-roosts are clearly also l-roosts, and k-mature occurrences also l-mature.)

2. An occurrence of a_1 in block B of S is a *k-egg* if the sequence contains

$$a_1\, a_2 \cdots a_k \;\; a_k^2 a_{k-1}^2 \cdots a_2^2 \;\; B \;\; a_2^2 a_3^2 \cdots a_k^2 \;\; a_k a_{k-1} \cdots a_1$$

   Note that any middle occurrence of a symbol is a 1-egg.

One may already discern from Definition 6.2 the shape of the rest of the proof. A k-roost can only exist if the sequence contains a $\mathrm{Perm}^{\mathrm{dbl}}_{k,4}$ sequence, so there cannot be r-roosts. If the proliferation of symbols *necessarily* leads to k-roosts for ever larger k, we have a cap on the proliferation of symbols. Lemma 6.3 lists some straightforward consequences of Defintion 6.2.

[14] Observe that for any $a \in \Sigma(Z)$, the height of $\mathcal{T}_{|a}$ is $i^\star + 1$ and all quills of $\mathcal{T}_{|a}$ are at distance at least 2 from $\mathrm{cr}_{|a}$. Every non-terminal quill can therefore molt up to $i^\star - 1$ times, generating up to three terminals per molting, each of which carries unit potential.

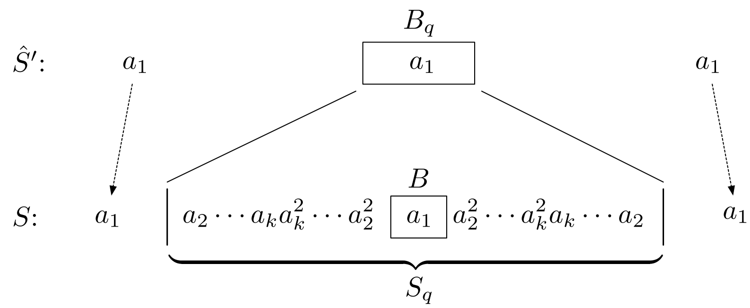

Figure 5: A k-egg is formed when a middle $a_1 \in B_q$ is dropped into a $(k-1)$-roost in $\check{S}_q$.

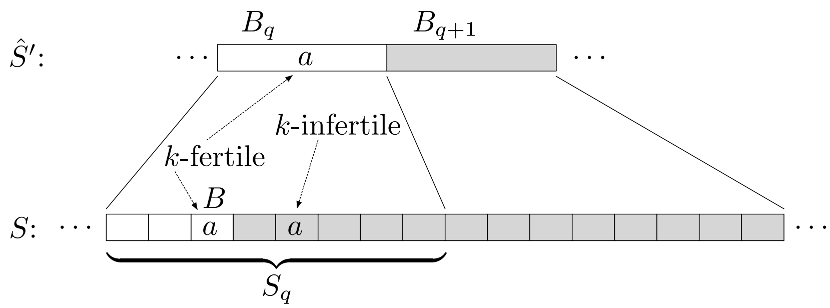

Figure 6: The shaded blocks are k-roosts. A k-left fertile occurrence of $a \in B_q$ in $\hat{S}'$ begets at most one k-left fertile occurrence in S_q, and, in this example, one k-infertile occurrence. Since B_{q+1} is a k-roost in $\hat{S}'$, all blocks in S_{q+1} are k-roosts in S whether or not they were already k-roosts in $\check{S}_{q+1}$.

Lemma 6.3. (Properties of Roosts and Eggs) *Let S be an m-block sequence encountered in the recursive decomposition of a $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free sequence Z. Define $\{S_q, \check{S}_q, \hat{S}_q\}_{1 \leqslant q \leqslant \hat{m}}$ and $\hat{S}' = B_1 \cdots B_{\hat{m}}$ as usual.*

1. *No block in S is an r-roost. All r-eggs represent at most 3 occurrences in Z.*
2. *If B_q is a k-roost in $\hat{S}'$, every block of S_q is a k-roost in S.*
3. *Let B be a block in S_q containing a global symbol a. If B is a $(k-1)$-roost in $\check{S}_q$ and the $a \in B_q$ is a middle occurrence in $\hat{S}'$ then $a \in B$ is a k-egg in S. See Figure 5.*
4. *Let B be a block in S_q containing a global symbol a. Suppose the $a \in B_q$ is k-left fertile in $\hat{S}'$ and the $a \in B$ is k-left fertile in S. All blocks following B in S_q are k-roosts in S. A symmetric statement is true of k-right fertile occurrences. See Figure 6.*

6.3 Molting and the Evolution of Potentials

Consider the status of a non-terminal symbol a as it descends, in $\mathcal{T}_{|a}$, from $\mathrm{qu}_{|a}(v)$ to some leaf v. Since $a \in \mathcal{B}(\mathrm{qu}_{|a}(v))$ is a middle symbol at that level (it is not on either wing of $\mathcal{T}_{|a}$), this a begins as a 1-egg and may become 1-fertile (left or right), then 1-infertile, then a 2-egg, 2-fertile, 2-infertile, and so on. It cannot become r-mature (fertile or infertile) for this would mean that $\mathrm{Perm}^{\mathrm{dbl}}_{r,4} \prec Z$, so there are at most $3(r-1)$ transitions. Multiple transitions may occur simultaneously. When a non-terminal first becomes a k-egg, or k-fertile, or k-infertile, its potential becomes $\phi^{\mathrm{eg}}_k, \phi^{\mathrm{fe}}_k$, or ϕ^{in}_k, where

$$\phi = \phi^{\mathrm{eg}}_1 > \phi^{\mathrm{fe}}_1 > \phi^{\mathrm{in}}_1 > \cdots > \phi^{\mathrm{eg}}_{r-1} > \phi^{\mathrm{fe}}_{r-1} > \phi^{\mathrm{in}}_{r-1} > \phi^{\mathrm{eg}}_r = 3$$

If we can show that each symbol molts $O(1)$ times between status transitions, it suffices to set the initial potential at $\phi = O(r) = O(1)$. This is clearly true of k-egg $\to$ k-mature transitions. Any k-egg a that molts three as must have molted two of them to the same side, left or right, making it k-mature. Since a non-terminal can molt up to 3 terminals in the molting event that makes it k-mature, it suffices to set $\phi^{\mathrm{eg}}_k - \phi^{\mathrm{fe}}_k = 5$. (If this a transitions directly from a k-egg to k-infertile, all the better, for $\phi^{\mathrm{in}}_k < \phi^{\mathrm{fe}}_k$.) We now analyze the k-fertile $\to$ k-infertile and k-infertile $\to$ $(k+1)$-egg transitions.

Lemma 6.4. *Fix a block index $q \leqslant [\![\hat{S}']\!]$ and let $F \subset B_q$ be those symbols newly k-left fertile, that is, they were not k-left fertile at any ancestor of B_q in their respective derivation trees. The total number of terminals molted by F-symbols before they become k-infertile is at most $2|F| + (r-1)\binom{i^\star-1}{2}$.*

Proof. Lemma 6.3(4) implies that so long as symbols in F remain k-fertile, as they travel from B_q to a block in S_q, to blocks at lower levels of the derivation tree, they will always be contained in a *single* block at that level of the tree. In other words, there is a sequence of nodes $(B_q = v_1, v_2, \ldots, v_l)$ in $\mathcal{T}$ lying on a path from $B_q = v_1$ (in $\hat{S}'$), to v_2 (in S), to a descendant leaf v_l (where $l \leqslant i^\star$) such that any symbol $a \in F$ is k-left fertile in some *prefix* of the list $\mathcal{B}(v_1), \mathcal{B}(v_2), \ldots, \mathcal{B}(v_l)$. See Figure 7. Call a symbol $a \in F$ *type* (f, g) if a molted a terminal to the right at both $\mathcal{B}(v_f)$ and $\mathcal{B}(v_g)$, for $1 < f < g \leqslant l$.[15] That is, in $\mathcal{T}_{|a}$, $\mathcal{B}(v_f)$ and $\mathcal{B}(v_g)$ have right siblings. Note that during the time in which this a is k-left fertile it can molt at most once to the left: molting two as to the left would make it k-infertile.

By the pigeonhole principle, if $(r-1)\binom{i^\star-1}{2} + 1$ symbols in F molted twice to the right then a subset $F' \subset F$ of r of them have the same type, say (f, g). However, this would imply that Z is not $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free. Since k-fertile symbols are middle symbols, every symbol in F' appears at least once before and after B_q. The occurrences of F'-symbols in $\mathcal{B}(v_g)$ are non-terminal, so they each represent at least two occurrences in Z. Finally, the F'-symbols appear twice at descendants of B_q but to the right of $\mathcal{B}(v_g)$. See Figure 7.

To sum up, we let each F-symbol molt once to the left and once to the right while k-left fertile. Some subset can molt more than once to the right, but the total number of such terminals molted by these symbols is at most $(r-1)\binom{i^\star-1}{2}$. □

A nearly symmetric analysis can be applied to right fertile symbols. The asymmetry comes from the fact that non-terminals can molt two terminals to the left but only one to the right.

Lemma 6.5. *Fix a block index $q \leqslant [\![\hat{S}']\!]$ and let $F \subset B_q$ be those symbols newly k-right fertile, that is, they were not k-left fertile at any ancestor of B_q in their respective derivation trees. The total number of terminals molted by F-symbols before they become k-infertile is at most $2|F| + (r-1)(\binom{i^\star-1}{2} + i^\star - 1)$.*

[15] Note that a symbol that molts exactly twice to the right has one type. In general, a symbol that molts h times to the right is of $\binom{h}{2}$ distinct types.

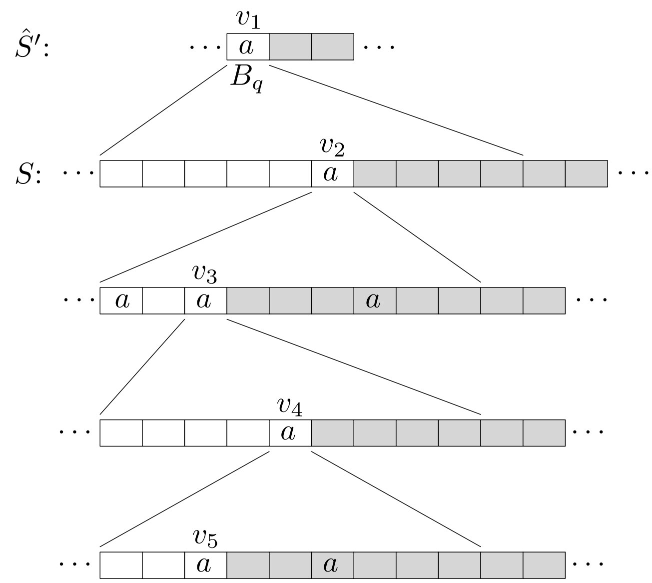

Figure 7: A newly k-left-fertile symbol $a \in B_q = \mathcal{B}(v_1)$ in $\hat{S}'$. As a progresses down $\mathcal{T}_{|a}$ it continues to be k-left fertile at $\mathcal{B}(v_2), \ldots, \mathcal{B}(v_5)$. Since it molts to the right at blocks $\mathcal{B}(v_3)$ and $\mathcal{B}(v_5)$ it has type $(3,5)$. It also molts to the left at $\mathcal{B}(v_3)$. Were it to molt twice to the left at $\mathcal{B}(v_3)$, $\mathcal{B}(v_3)$ would then become a k-roost and the $a \in \mathcal{B}(v_3)$ k-infertile.

Proof. The argument is the same as above, except that we allow types (f, f) if a symbol molts twice to the left at $\mathcal{B}(v_f)$. There are now at most $(\binom{i^\star-1}{2} + i^\star - 1)$ possible types, and we cannot see r symbols of the same type. □

According to Lemmas 6.4 and 6.5, it suffices to set $\phi_k^{\mathrm{fe}} = \phi_k^{\mathrm{in}} + 2$. The total number of molted terminals unaccounted for, over all q, all $k < r$, counting both k-left fertile and k-right fertile symbols in B_q, is $\hat{m} \cdot (r-1)^2(2\binom{i^\star-1}{2} + i^\star - 1) < \hat{m} \cdot [(r-1)(i^\star - 1)]^2$, which are covered by the last term of Recurrence 6.1.

The remaining task is to analyze the k-infertile $\to$ $(k+1)$-egg transition.

Lemma 6.6. *Let u, v, w be distinct nodes such that $a, b \in \mathcal{B}(u), a \in \mathcal{B}(v), b \in \mathcal{B}(w)$, where v is the parent of u in $\mathcal{T}_{|a}$ and w is the parent of u in $\mathcal{T}_{|b}$. If a, b were k-infertile in blocks $\mathcal{B}(v)$ and $\mathcal{B}(w)$ then at least one of a, b became a $(k+1)$-egg when it was inserted into $\mathcal{B}(u)$.*

Proof. This is a consequence of Lemma 6.3(2,3). Without loss of generality w is a strict ancestor of v, so a was inserted into $\mathcal{B}(u)$ before b was inserted into $\mathcal{B}(u)$. Since the $a \in \mathcal{B}(v)$ was k-infertile, $\mathcal{B}(v)$ was a k-roost, by definition. By Lemma 6.3(2), $\mathcal{B}(u)$ became a k-roost after a was inserted there. By Lemma 6.3(3), when b was inserted in $\mathcal{B}(u)$ it became a $(k+1)$-egg. □

Lemma 6.7. *Let $I \subset \Sigma(\hat{S}_q)$ be those non-terminals that were k-infertile, non-$(k+1)$-eggs in B_q but became $(k+1)$-eggs in S_q. The number of terminals molted by I symbols while they were k-infertile, non-$(k+1)$-eggs is at most $2|I| + (r-1)(2\binom{i^\star-2}{2} + i^\star - 2)$.*

Proof. Lemma 6.6 implies that on a path from B_q to the root of $\mathcal{T}$ we encounter nodes $v_1 = B_q, v_2, \ldots, v_l$, not necessarily adjacent, such that, for each symbol $a \in I$, the set of blocks in which

a is k-infertile and not a quill is some *prefix* of $\mathcal{B}(v_1), \ldots, \mathcal{B}(v_l)$, where $l \leqslant i^\star - 2$. Call an $a \in I$ type $(\rightarrow, f, g)$ if it molted a terminal to the right in both $\mathcal{B}(v_f)$ and $\mathcal{B}(v_g)$, where $1 \leqslant f < g \leqslant l$. Call it type $(\leftarrow, f, g)$, where $1 \leqslant f \leqslant g \leqslant l$, if it molted a terminal to the left in both $\mathcal{B}(v_f)$ and $\mathcal{B}(v_g)$, or two terminals to the left if $f = g$. There are $2\binom{l}{2} + l$ distinct types. There cannot be r symbols of one type, for this would imply that Z is not $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free. (The argument is the same as in the proof of Lemma 6.4.) Since every symbol that molts more than two terminals is of at least one type, the total number of terminals molted by I while being k-infertile, non-$(k+1)$-eggs is $2|I| + (r-1)(2\binom{i^\star-2}{2} + i^\star - 2)$. □

We set $\phi^{\mathrm{in}}_k - \phi^{\mathrm{eg}}_{k+1} = 2$, so the total number of terminals unaccounted for, over all $q < \hat{m}$ and $k < r$, is at most $\hat{m} \cdot [(r-1)(i^\star - 2)]^2$, which is covered by the last term of Recurrence 6.1. Given the constraints we have established on potentials it suffices to set $\phi = \phi^{\mathrm{eg}}_1 = 7(r-1) + 1$, since $|\phi^{\mathrm{eg}}_k - \phi^{\mathrm{fe}}_k| = 5, |\phi^{\mathrm{fe}}_k - \phi^{\mathrm{in}}_k| = |\phi^{\mathrm{in}}_k - \phi^{\mathrm{eg}}_{k+1}| = 2$, and $\phi^{\mathrm{eg}}_r = 3$.

Remark 6.8. Observe the asymmetry in the arguments of Lemmas 6.4–6.5 and Lemma 6.7. In Lemmas 6.4 and 6.5 we are tracking moltings that will happen "in the future" (below the level of S in $\mathcal{T}$) whereas in Lemma 6.7 we are accounting for moltings that have already occurred at and above the level of $\hat{S}'$ in $\mathcal{T}$.

6.4 Wrapping Up the Analysis

Since $\Lambda_{r,2}(\cdot,\cdot)$ and $\Lambda^{\mathrm{dbl}}_{r,2}(\cdot,\cdot)$ are both linear and $\hat{m} < m$, we can simplify Recurrence 6.1 to

$$\Upsilon(n,m) \leqslant \sum_{1 \leqslant q \leqslant \hat{m}} \Upsilon(\check{n}_q, m_q) + \Upsilon(\hat{n}, \hat{m}) + C[\hat{n} + (i^\star)^2 m],$$

for some constant C depending only on r. A straightforward proof by induction shows that for any $i \leqslant i^\star$ and j minimal such that $m \leqslant a_{i,j}$, $\Upsilon(n,m) \leqslant Ci(n + (i^\star)^2 jm)$. Putting it all together we have, for $\|Z\| = n^\star$ and $[\![Z]\!] = m^\star$,

$$|Z| \leqslant \Lambda^{\mathrm{dbl}}_{r,3}(n^\star, m^\star) \leqslant \Upsilon(n^\star, m^\star) \leqslant Ci^\star n^\star + C(i^\star)^3 j^\star m^\star. \tag{9}$$

Eqn. (9) leads to an upper bound of $\Lambda^{\mathrm{dbl}}_{r,3}(n,m) = O(n\alpha(n,m) + m\alpha^3(n,m))$, which, by Lemma 3.1, implies an upper bound of $\Lambda^{\mathrm{dbl}}_{r,3}(n) = O(n\alpha^3(n))$. Theorem 6.9 reduces this to $O(n\alpha(n))$, which is asymptotically tight since $\Lambda^{\mathrm{dbl}}_{r,3}(n) = \Omega(\lambda_3(n))$.

Theorem 6.9. *For any $r \geqslant 2$, $\Lambda^{\mathrm{dbl}}_{r,3}(n) = \Theta(n\alpha(n))$ and $\Lambda^{\mathrm{dbl}}_{r,3}(n,m) = \Theta(n\alpha(n,m) + m)$.*

Proof. Let S be a $\mathrm{Perm}^{\mathrm{dbl}}_{r,4}$-free sequence. To bound $|S|$ asymptotically we can assume, using Lemmas 3.1 and 3.3, that S consists of $m \leqslant 2n$ blocks. (If there are $m > 2n$ blocks, remove up to $r-1$ symbols at block boundaries to make it r-sparse. If the sequence is r-sparse, we can discard a constant fraction of occurrences to partition the sequence into $2n$ blocks.) Choose i to be minimal such that $m \leqslant a_{i,j}$, where $j = \max\{3, \lceil n/m \rceil\}$. Partition $S = S_1 \cdots S_{\hat{m}}$ into $\hat{m} = \lceil m/i^2 \rceil$ intervals, each consisting of i^2 blocks. Define $\hat{S}, \hat{S}', \check{S}_q$, etc. as usual. Applying Eqn. (9) with $i^\star = i$, we have $|\hat{S}'| \leqslant C(i\hat{n} + i^3 j\hat{m}) \leqslant C(i(\hat{n} + jm)) = O(in)$. Since each $\acute{S}_q, \grave{S}_q$, and $\bar{S}_q$ is $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$-free and $\Lambda^{\mathrm{dbl}}_{r,2}(n_q, m_q) = O(n_q + m_q)$ is linear, it follows that $|\hat{S}| = O(in + m) = O(in)$. We now apply Eqn. (9) to local symbols with $i^\star = 1$, that is, for each index $q \leqslant \hat{m}$, j is chosen to be minimal such that $m_q \leqslant a_{1,j}$. Since $a_{1,j} = 2^j$, $j = \lceil \log m_q \rceil \leqslant \lceil \log i^2 \rceil$. It follows that $|\check{S}| = \sum_q |\check{S}_q| \leqslant \sum_q C(\check{n}_q + m_q \log m_q) = O(\check{n} + m \log(i^2)) = O(n \log i)$. Since $i = \alpha(n,m) + O(1)$, $|S| = |\hat{S}| + |\check{S}| = O(n\alpha(n,m)) = O(n\alpha(n))$. □

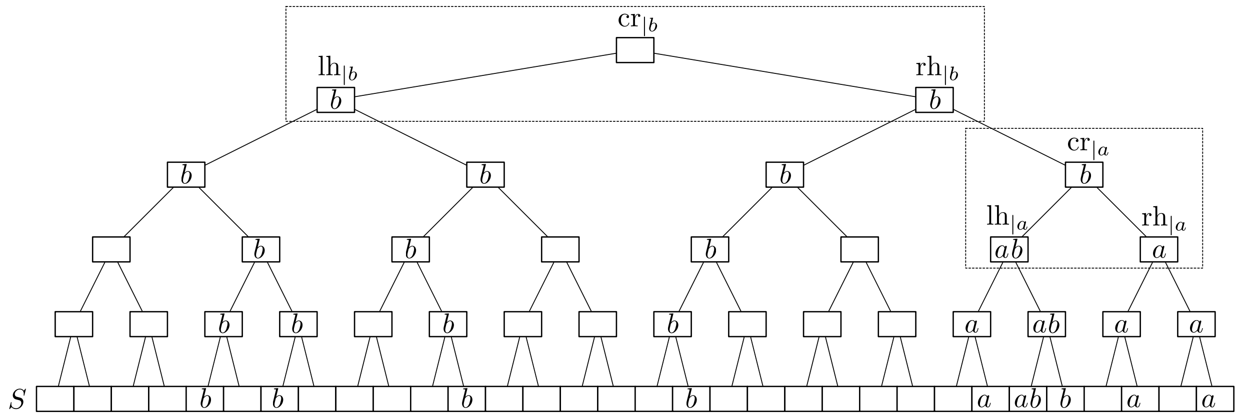

Figure 8: An example of a canonical derivation tree for S. Dashed boxes isolate the base case trees that assign $a, b \in \Sigma(S)$ their crowns and heads.

Theorem 6.9 and Lemma 1.2 immediately give us asymptotically sharp bounds on the extremal functions for certain doubled forbidden sequences.

Corollary 6.10. *(See Nivasch [16, Rem. 5.1], Pettie [20], Geneson, Prasad, and Tidor [8], and Klazar [13, p. 13].)*

$$\begin{aligned} \lambda_3^{\mathrm{dbl}}(n) &= \Theta(\Lambda_{2,3}^{\mathrm{dbl}}(n)) = \Theta(n\alpha(n)), \\ \mathrm{Ex}(\mathrm{dbl}(abcacbc), n) &= \Theta(\Lambda_{4,3}^{\mathrm{dbl}}(n)) = \Theta(n\alpha(n)), && \textit{See [20]} \\ \mathrm{Ex}(\mathrm{dbl}(abcabca), n) &= \Theta(\Lambda_{3,3}^{\mathrm{dbl}}(n)) = \Theta(n\alpha(n)), && \textit{See [16]} \end{aligned}$$

and, more generally,

$$\mathrm{Ex}(\mathrm{dbl}(1 \cdots k\, 1 \cdots k\, 1), n) = \Theta(\Lambda_{r,3}^{\mathrm{dbl}}(n)) = \Theta(n\alpha(n)),$$

where $r = (k-1)^3 + 1$.

7 Double Davenport-Schinzel Sequences

Recall from Section 5.1 that the *canonical* derivation tree $\mathcal{T}^\star(S)$ is obtained by decomposing S in the least aggressive way possible, choosing $\hat{m} = \lceil [\![S]\!]/2 \rceil$ whenever $[\![S]\!] > 2$. Figure 8 gives an example of such a tree.

The structure of the canonical derivation tree is, in many respects, simpler than general derivation trees. For example, all wing nodes in any projection tree $\mathcal{T}_{|a}$, where $a \in \Sigma(S)$, have either one or two children. Those with two children (*branching* nodes) are associated with precisely one quill and therefore one feather,[16] so counting the number of feathers is tantamount to counting branching wing nodes.

[16] Recall that a *feather* of $\mathcal{T}_{|a}$ is the rightmost descendant of a dove quill or leftmost descendant of a hawk quill.

Nesting was a concept introduced in [22] to analyze odd-order DS sequences. Here we generalize it to deal with double DS sequences.

Definition 7.1. (Nesting) Let B be a block of S containing $a, b \in \Sigma(S)$. If S contains either

$$a\, b\, b\; B\; b\, b\, a \qquad \text{or} \qquad b\, a\, a\; B\; a\, a\, b$$

then a and b are called *double-nested* in B.

Lemma 7.2 can be thought of as a generalization of [22, Lem. 4.3] to deal with double-nestedness. Whereas [22, Lem. 4.3] assumed any derivation tree, Lemma 7.2 refers to the canonical derivation tree $\mathcal{T}^\star(S')$ as this makes the proof slightly simpler. This assumption is actually without much loss of generality since any derivation tree obtained with uniform block partitions is "contained" in the canonical derivation tree, that is, its blocks are subsequences of the corresponding blocks in the canonical tree.

Lemma 7.2. *Consider a sequence S', its canonical derivation tree $\mathcal{T}^\star(S')$, and a leaf v for which $a, b \in \mathcal{B}(v)$. Let S be obtained from S' by substituting, for each leaf $u \neq v$, a sequence $S(u)$ containing at least two copies of each symbol in $\mathcal{B}(u)$. (The block $\mathcal{B}(v)$ appears verbatim in S.) If v is neither a wingtip nor feather in both $\mathcal{T}^\star_{|a}$ and $\mathcal{T}^\star_{|b}$ then, in S, a and b are double-nested in $\mathcal{B}(v)$.*

Proof. Without loss of generality we can assume that v is a dove in $\mathcal{T}^\star_{|a}$ and $\mathrm{cr}_{|b}$ is ancestral to $\mathrm{cr}_{|a}$. Because v is neither a wingtip nor feather in $\mathcal{T}^\star_{|a}$, it must be distinct from the leftmost and rightmost leaf descendants of $\mathrm{wi}_{|a}(v)$, namely $\mathrm{lt}_{|a}$ and $\mathrm{fe}_{|a}(v)$. Moreover, since v is a dove in $\mathcal{T}^\star_{|a}$ it descends from the right child of $\mathrm{wi}_{|a}(v)$, namely $\mathrm{qu}_{|a}(v)$. Partition S into four intervals

I_1 : everything preceding $\mathcal{B}(\mathrm{lt}_{|a})$.

I_2 : everything from I_1 to the beginning of $\mathcal{B}(v)$.

I_3 : everything from the end of $\mathcal{B}(v)$ to the end of $\mathcal{B}(\mathrm{fe}_{|a}(v))$.

I_4 : everything following I_3.

If b appeared in both I_1 and I_4 then $a, b \in \mathcal{B}(v)$ would clearly be double-nested in S. Therefore it suffices to consider two cases, (1) I_1 contains no bs, and (2) I_4 contains no bs. Figures 9 and 10 illustrate the two cases.

Case 1. The wingtip $\mathrm{lt}_{|b}$ must be in interval I_2, though it may be identical to $\mathrm{lt}_{|a}$. Since $\mathrm{wi}_{|a}(v)$ is ancestral to both $\mathrm{lt}_{|b}$ and v, and is a strict descendant of $\mathrm{cr}_{|b}$, it follows that v is a dove in $\mathcal{T}^\star_{|b}$ and that $\mathrm{wi}_{|b}(v)$ is a descendant of $\mathrm{wi}_{|a}(v)$. The rightmost descendant of $\mathrm{wi}_{|b}(v)$ in $\mathcal{T}_{|b}$ is $\mathrm{fe}_{|b}(v)$, which is distinct from v. Since $\mathrm{wi}_{|a}(v)$ is a descendant of $\mathrm{lh}_{|a}$, any descendant of $\mathrm{rh}_{|a}$, such as $\mathrm{rt}_{|a}$, lies to the right of $\mathrm{fe}_{|b}(v)$, in interval I_4. By the same reasoning, $\mathrm{rt}_{|b}$ lies in I_4.

Regardless of whether $\mathrm{lt}_{|a}$ and $\mathrm{lt}_{|b}$ are identical or distinct, $\mathcal{B}(v)$ is preceded, in S, by either abb or baa. In the first case $\mathrm{lt}_{|a}, \mathrm{lt}_{|b}, v, \mathrm{fe}_{|b}(v), \mathrm{rt}_{|a}$ certify that a, b are double-nested in $\mathcal{B}(v)$; see Figure 9. In the latter case $\mathrm{lt}_{|b} = \mathrm{lt}_{|a}, v, \mathrm{fe}_{|a}(v), \mathrm{rt}_{|b}$ certify that a, b are double-nested in $\mathcal{B}(v)$.

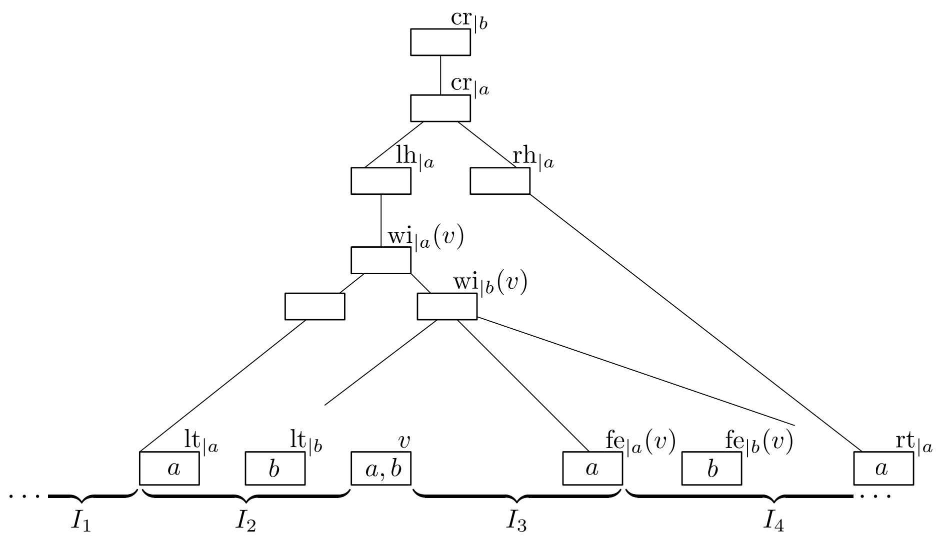

Figure 9: In Case 1 interval I_1 contains no bs. Contrary to the depiction, $\mathrm{lt}_{|a}$ and $\mathrm{lt}_{|b}$ are not necessarily distinct, nor are $\mathrm{wi}_{|a}(v)$ and $\mathrm{wi}_{|b}(v)$ or $\mathrm{cr}_{|a}$ and $\mathrm{cr}_{|b}$. In this depiction $\mathrm{qu}_{|a}(v)$, the right child of $\mathrm{wi}_{|a}(v)$, happens to be identical to $\mathrm{wi}_{|b}(v)$.

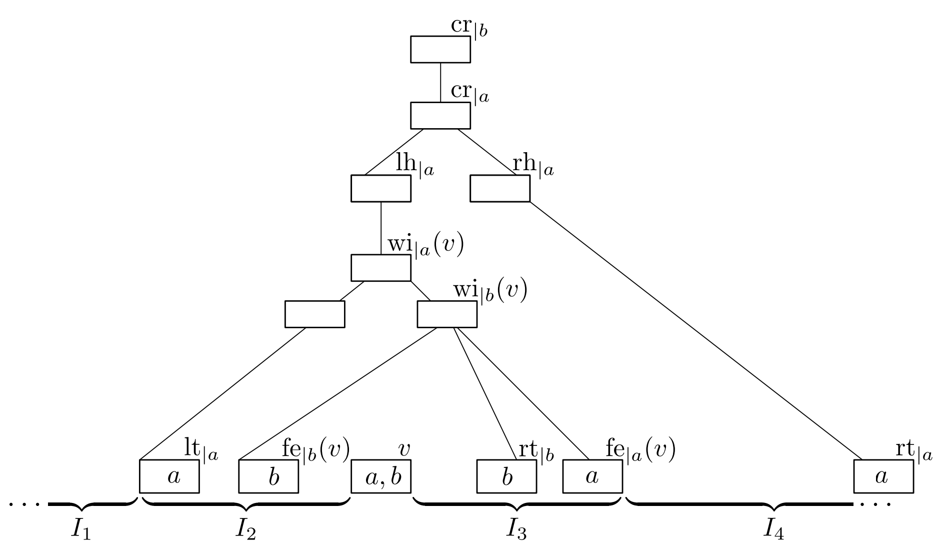

Figure 10: In Case 2 interval I_4 contains no bs. Contrary to the depiction, $\mathrm{rt}_{|b}$ and $\mathrm{fe}_{|a}(v)$ are not necessarily distinct.

Case 2. The wingtip $\mathrm{rt}_{|b}$ must lie in I_3, so v and $\mathrm{rt}_{|b}$ are both descendants of $\mathrm{qu}_{|a}(v)$, the right child of $\mathrm{wi}_{|a}(v)$. It follows that v is a hawk in $\mathcal{T}^\star_{|b}$ and that no descendants of $\mathrm{wi}_{|b}(v)$ are in interval I_1. Since $\mathrm{fe}_{|b}(v)$ is the leftmost descendant of $\mathrm{wi}_{|b}(v)$ in $\mathcal{T}^\star_{|b}$, and $\mathrm{fe}_{|b}(v) \neq v$, the distinct nodes $\mathrm{lt}_{|a}, \mathrm{fe}_{|b}(v), v, \mathrm{rt}_{|b}, \mathrm{rt}_{|a}$ certify that a, b are double-nested in $\mathcal{B}(v)$. See Figure 10. □

Recurrence 7.3 gives a significantly simpler method for bounding the number of feathers, compared to [22, Recs. 5.1 and 7.6]. Whereas [22] considered feathers in an arbitrary derivation tree, Recurrence 7.3 *only* considers the canonical derivation tree.

Recurrence 7.3. *Let S be an m-block, order-s DS sequence over an n-letter alphabet and $\mathcal{T} = \mathcal{T}^\star(S)$ be its canonical derivation tree. Define $\Phi_s(n,m)$ to be the maximum number of feathers of one type (dove or hawk) in such a sequence, where* feather *is with respect to $\mathcal{T}$. For any $s \geqslant 2$,*

$$\Phi_s(n,2) = 0$$
$$\Phi_2(n,m) < m$$

and for any uniform block partition $\{m_q\}_{1\leqslant q\leqslant \hat{m}}$ and alphabet partition $\{\hat{n}\} \cup \{\check{n}_q\}_{1\leqslant q\leqslant \hat{m}}$,

$$\Phi_s(n,m) \leqslant \sum_{q=1}^{\hat{m}} \Phi_s(\check{n}_q, m_q) + \Phi_s(\hat{n},\hat{m}) + \Phi_{s-1}(\hat{n},m) + \hat{n}$$

Proof. Suppose we only wish to bound dove feathers. If there are only two blocks then all occurrences are wingtips and feathers are not wingtips. This gives the first equality. In the most extreme case every non-wingtip is a dove feather, so $\Phi_s(n,m) \leqslant \lambda_s(n,m) - 2n$. In particular, $\Phi_2(n,m) \leqslant \lambda_2(n,m) - 2n < m$. Decompose S into $\hat{S}, \hat{S}', \acute{S}_q, \grave{S}_q, \bar{S}_q$ in the usual way with respect to the given uniform block partition. Let $\hat{\mathcal{T}} = \mathcal{T}^\star(\hat{S}')$ be the canonical derivation tree of the contracted global sequence $\hat{S}'$. It follows that $\acute{S}_q$ is an order-$(s-1)$ DS sequence. Define $\acute{\mathcal{T}}_q = \mathcal{T}^\star(\acute{S}_q)$ to be its canonical derivation tree. The branching nodes on the left wing of $\mathcal{T}_{|a}$, where $a \in \Sigma(\acute{S}_q)$, consist of (i) the branching nodes on the left wing of $\hat{\mathcal{T}}_{|a}$, (ii) the branching nodes on the left wing of $(\acute{\mathcal{T}}_q)_{|a}$, and (iii) the crown $\acute{\mathrm{cr}}_{|a}$ of $(\acute{\mathcal{T}}_q)_{|a}$, which is on the left wing of $\mathcal{T}_{|a}$ but not $(\acute{\mathcal{T}}_q)_{|a}$. Each branching node is identified with one feather in $\mathcal{T}_{|a}$. The total number of branching nodes/feathers covered by (i), summed over all $a \in \Sigma(\hat{S})$, is at most $\Phi_s(\hat{n},\hat{m})$. The total number covered by (ii), summed over all $q \leqslant \hat{m}$ and $a \in \Sigma(\hat{S}_q)$, is $\sum_q \Phi_{s-1}(\acute{n}_q, m_q) \leqslant \Phi_{s-1}(\hat{n}, m)$. The number covered by (iii) is clearly $\hat{n}$, which gives the last inequality. □

Recurrence 7.4 generalizes [16, Rec. 3.1] and [22, Recs. 3.3, 5.2, and 7.7], from DS sequences to double DS sequences. When $s = 3$ or $s \geqslant 4$ is even, Recurrence 7.4 is substantively no different than Recurrence 4.3 for $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free sequences.

Recurrence 7.4. *Let s, n, and m be the order, alphabet size, and block count parameters. Let $\{m_q\}_{1\leqslant q\leqslant \hat{m}}$ be a uniform block partition, where $\hat{m} \geqslant 2$, and $\{\hat{n}\} \cup \{\check{n}_q\}_{1\leqslant q\leqslant \hat{m}}$ be an alphabet partition. When $\hat{m} = 2$, for any $s \geqslant 3$,*

$$\lambda^{\mathrm{dbl}}_s(n,m) \leqslant \sum_{q\in\{1,2\}} \lambda^{\mathrm{dbl}}_s(\check{n}_q, m_q) + \lambda^{\mathrm{dbl}}_{s-1}(2\hat{n}, m) + 2\hat{n}.$$

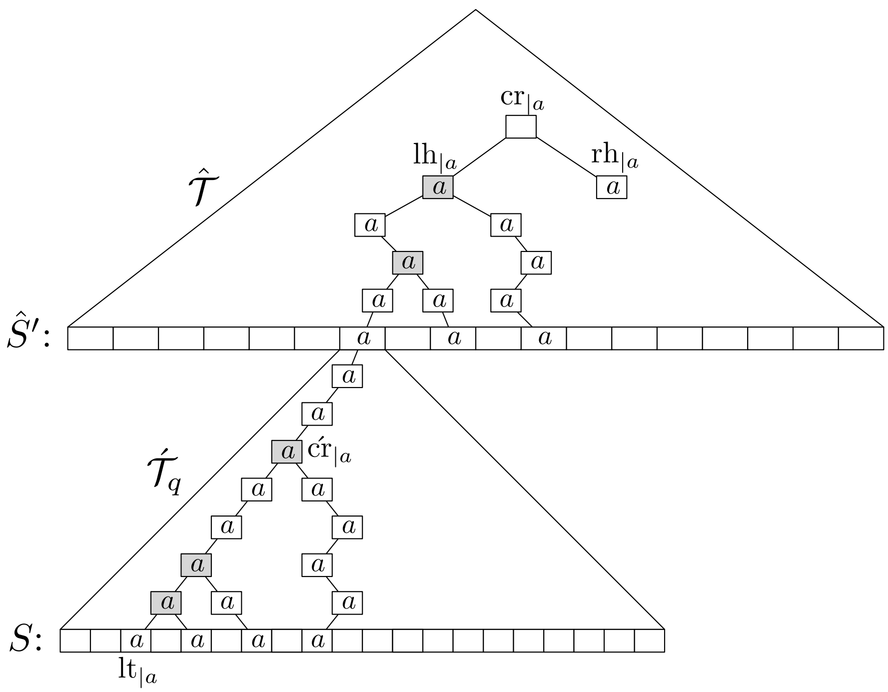

Figure 11: Counting dove feathers in $T_{|a}$ is tantamount to counting branching nodes on the left wing of $\mathcal{T}_{|a}$.

When $\hat{m} > 2$ and either $s = 3$ or $s \geqslant 4$ is even,

$$\lambda_s^{\mathrm{dbl}}(n, m) \leqslant \sum_q \lambda_s^{\mathrm{dbl}}(\check{n}_q, m_q) + \lambda_s^{\mathrm{dbl}}(\hat{n}, \hat{m}) + 2 \cdot \lambda_{s-1}^{\mathrm{dbl}}(\hat{n}, m) + \lambda_{s-2}^{\mathrm{dbl}}(\lambda_s(\hat{n}, \hat{m}), m) + 2 \cdot \lambda_s(\hat{n}, \hat{m}),$$

and when $s \geqslant 5$ is odd,

$$\lambda_s^{\mathrm{dbl}}(n, m) \leqslant \sum_{q=1}^{\hat{m}} \lambda_s^{\mathrm{dbl}}(\check{n}_q, m_q) + \lambda_s^{\mathrm{dbl}}(\hat{n}, \hat{m}) + 2 \cdot \lambda_{s-1}^{\mathrm{dbl}}(\hat{n}, m) + \lambda_{s-2}^{\mathrm{dbl}}(2 \cdot \Phi_s(\hat{n}, \hat{m}), m) + 4 \cdot \Phi_s(\hat{n}, \hat{m}) + \lambda_{s-3}^{\mathrm{dbl}}(\lambda_s(\hat{n}, \hat{m}), m) + 2 \cdot \lambda_s(\hat{n}, \hat{m})$$

Proof. First consider the case when $s \geqslant 5$ is odd. Let S be an order-s double DS sequence, decomposed into $\hat{S}$ and $\{\check{S}_q\}$ as usual. The contribution of local symbols is $\sum_q \lambda_s^{\mathrm{dbl}}(\check{n}_q, m_q)$. If a global symbol occurs exactly once in an $\hat{S}_q$ this occurrence is a *singleton*. Let $\dot{S} \prec \hat{S}$ be the subsequence of singletons and $\ddot{S} \prec \hat{S}$ be the subsequence of non-singletons. By definition $\dot{S}$ is partitioned into $\hat{m}$ blocks, so $|\dot{S}| \leqslant \lambda_s^{\mathrm{dbl}}(\hat{n}, \hat{m})$. Symbols in $\Sigma(\ddot{S}_q)$ are classified as *first*, *last*, and *middle* if they appear, in $\ddot{S}$, after $\ddot{S}_q$ but not before, before $\ddot{S}_q$ but not after, and both before and after $\ddot{S}_q$, respectively. In the worst case these three criteria are exhaustive. However, it may be that all non-singleton occurrences of a symbol appear *exclusively* in $\Sigma(\ddot{S}_q)$. In this case we call the symbol *first* if it appears after interval q in $\dot{S}$ and *last* if it is not first and appears before interval q in $\dot{S}$. Define $\acute{S}_q, \grave{S}_q, \bar{S}_q \prec \ddot{S}_q$ to be the subsequences of first, last, and middle occurrences in $\ddot{S}_q$.

If we remove the last occurrence of each letter from $\acute{S}_q$, or the first occurrence of each letter from $\grave{S}_q$, the resulting sequence is an order-$(s-1)$ double DS sequence. The contribution of first and last non-singletons is therefore at most

$$\sum_q \left[\lambda^{\mathrm{dbl}}_{s-1}(\acute{n}_q, m_q) + \acute{n}_q + \lambda^{\mathrm{dbl}}_{s-1}(\grave{n}_q, m_q) + \grave{n}_q\right] \leqslant 2(\lambda^{\mathrm{dbl}}_{s-1}(\hat{n}, m) + \hat{n}).$$

Obtain $\ddot{S}' = B_1 \cdots B_{\hat{m}}$ from $\ddot{S}$ by contracting each interval $\ddot{S}_q$ into a single block B_q. Since occurrences in $\ddot{S}'$ each represent at least two occurrences in $\ddot{S}$, we can conclude[17] that $|\ddot{S}'| \leqslant \lambda_s(\hat{n}, \hat{m})$.

Let $\ddot{\mathcal{T}} = \mathcal{T}^\star(\ddot{S}')$ be the canonical derivation tree of $\ddot{S}'$. Define $\tilde{S}'$ to be the subsequence of $\ddot{S}'$ consisting of feathers with respect to $\ddot{\mathcal{T}}$ (both dove and hawk) and let $\tilde{S}$ be the subsequence of $\ddot{S}$ begat by symbols in $\tilde{S}'$. It follows that $|\tilde{S}'| \leqslant 2 \cdot \Phi_s(\hat{n}, \hat{m})$ since Φ_s only counts feathers of one type (dove or hawk). Define $\mathring{S}' \prec \ddot{S}'$ to be the subsequence of non-feather, non-wingtips with respect to $\ddot{\mathcal{T}}$, and define $\mathring{S} \prec \ddot{S}$ analogously. Since $\tilde{S}$ consists solely of middle symbols, removing the first and last occurrence of each letter in $\tilde{S}_q$ leaves an order-$(s-2)$ double DS sequence, hence

$$\begin{aligned}
|\tilde{S}| = \sum_q |\tilde{S}_q| &\leqslant \sum_q (\lambda^{\mathrm{dbl}}_{s-2}(\tilde{n}_q, m_q) + 2\tilde{n}_q) \\
&\leqslant \lambda^{\mathrm{dbl}}_{s-2}\left(\sum_q \tilde{n}_q, m\right) + 2\sum_q \tilde{n}_q \\
&\leqslant \lambda^{\mathrm{dbl}}_{s-2}(|\tilde{S}'|, m) + 2(|\tilde{S}'|) \\
&\leqslant \lambda^{\mathrm{dbl}}_{s-2}(2 \cdot \Phi_s(\hat{n}, \hat{m}), m) + 4 \cdot \Phi_s(\hat{n}, \hat{m})
\end{aligned}$$

We have accounted for every part of S except for $\mathring{S}$. Fix an interval q and $a, b \in \Sigma(\mathring{S}_q)$. Since $a, b \in B_q$ are neither feathers nor wingtips in $\ddot{\mathcal{T}}$, Lemma 7.2 implies that $\ddot{S}$ contains $a\,b\,b\,\ddot{S}_q\,b\,b\,a$. Suppose we remove the first and last occurrence of each letter in $\mathring{S}_q$. (These letters are underlined below.) The resulting sequence must be an order-$(s-3)$ double DS sequence, for if it contained a doubled alternating sequence with length $s-1$, which is even, we would see either

$$a\ b\ b\ \left|\ \overbrace{\underline{a}\ a\ b\ b\ \cdots a\ a\ b\ \underline{b}}^{s-1 \text{ alternations}}\ \right|\ b\ b\ a$$

or

$$a\ b\ b\ \left|\ \overbrace{\underline{b}\ b\ a\ a\ \cdots b\ b\ a\ \underline{a}}^{s-1 \text{ alternations}}\ \right|\ b\ b\ a,$$

[17] This is not quite true. As discussed in Remark 4.2, we can make this inference when bounding λ^{dbl}_s asymptotically.

contradicting the fact that S is an order-s double DS sequence. We can therefore bound $|\mathring{S}|$ by

$$\begin{aligned}
\sum_q |\mathring{S}_q| &\leqslant \sum_q (\lambda_{s-3}^{\mathrm{dbl}}(\mathring{n}_q, m_q) + 2\mathring{n}_q) \\
&\leqslant \lambda_{s-3}^{\mathrm{dbl}}\left(\sum_q \mathring{n}_q, m\right) + 2\sum_q \mathring{n}_q \\
&\leqslant \lambda_{s-3}^{\mathrm{dbl}}(|\mathring{S}'|, m) + 2|\mathring{S}'| \\
&\leqslant \lambda_{s-3}^{\mathrm{dbl}}(|\ddot{S}'| - 2\hat{n}, m) + 2(|\ddot{S}'| - 2\hat{n}) \\
&\leqslant \lambda_{s-3}^{\mathrm{dbl}}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m) + 2(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n})
\end{aligned}$$

This establishes the recurrence for odd $s \geqslant 5$. When $s = 3$ or $s \geqslant 4$ is even, we ignore the distinction between feathers and non-feathers and bound $|\bar{S}|$ by $\lambda_{s-2}^{\mathrm{dbl}}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m) + 2(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n})$. When $S = S_1S_2$ consists of $\hat{m} = 2$ intervals, no symbols are classified as middle, so it suffices to account for first, last, and local occurrences only. After discarding the last occurrence of each symbol from $\hat{S}_1$ and the first from $\hat{S}_2$, what remains are order-$(s-1)$ double DS sequences, so $|\hat{S}| \leqslant 2\hat{n} + \lambda_{s-1}^{\mathrm{dbl}}(\hat{n}, m_1) + \lambda_{s-1}^{\mathrm{dbl}}(\hat{n}, m_2) \leqslant 2\hat{n} + \lambda_{s-1}^{\mathrm{dbl}}(2\hat{n}, m)$. □

Recurrence 7.5 is similar to [22, Rec. 5.2] but presented in the style of Recurrence 7.4. The proof is essentially the same as that of Recurrence 7.4 except that we do not need to distinguish singletons from non-singletons, nor do we need to remove symbols from $\acute{S}_q, \grave{S}_q, \tilde{S}_q, \mathring{S}_q$, or $\bar{S}_q$ in order to make them double DS sequences with order $s-1$ or $s-2$ or $s-3$, as the case may be.

Recurrence 7.5. *Let s, n, and m be the order, alphabet size, and block count parameters. Let $\{m_q\}_{1 \leqslant q \leqslant \hat{m}}$ be a uniform block partition, where $\hat{m} \geqslant 2$, and $\{\hat{n}\} \cup \{\check{n}_q\}_{1 \leqslant q \leqslant \hat{m}}$ be an alphabet partition. When $\hat{m} = 2$, for any $s \geqslant 3$,*

$$\lambda_s(n, m) \leqslant \sum_{q \in \{1,2\}} \lambda_s(\check{n}_q, m_q) + \lambda_{s-1}(2\hat{n}, m).$$

When $\hat{m} > 2$ and either $s = 3$ or $s \geqslant 4$ is even,

$$\lambda_s(n, m) \leqslant \sum_q \lambda_s(\check{n}_q, m_q) + 2 \cdot \lambda_{s-1}(\hat{n}, m) + \lambda_{s-2}(\lambda_s(\hat{n}, \hat{m}) - 2\hat{n}, m)$$

and when $s \geqslant 5$ is odd,

$$\lambda_s(n, m) \leqslant \sum_{q=1}^{\hat{m}} \lambda_s(\check{n}_q, m_q) + 2 \cdot \lambda_{s-1}(\hat{n}, m) + \lambda_{s-2}(2 \cdot \Phi_s(\hat{n}, \hat{m}), m) + \lambda_{s-3}(\lambda_s(\hat{n}, \hat{m}), m)$$

Lemma 7.6 states some bounds on Φ_s, λ_s, and λ_s^{dbl} in terms of coefficients $\{\phi_{s,i}, \delta_{s,i}, \delta_{s,i}^{\mathrm{dbl}}\}$ and the ith row-inverse of Ackermann's function, for any $i \geqslant 1$. Refer to [22, Appendices B and C] for proofs of similar lemmas, and to the discussion following Lemma 4.4.

Lemma 7.6. *Fix parameters $i \geqslant 1$, $s \geqslant 3$, and $c \geqslant s-2$ and let n, m be the alphabet size and block count. Let j be minimal such that $m \leqslant (a_{i,j})^c$. Then Φ_s, λ_s, and λ_s^{dbl} are bounded by*

$$\begin{aligned}
\Phi_s(n, m) &\leqslant \phi_{s,i}\big(n + O((cj)^{s-2}m)\big) \\
\lambda_s(n, m) &\leqslant \delta_{s,i}\big(n + O((cj)^{s-2}m)\big) \\
\lambda_s^{\mathrm{dbl}}(n, m) &\leqslant \delta_{s,i}^{\mathrm{dbl}}\big(n + O((cj)^{s-2}m)\big)
\end{aligned}$$

where $\{\phi_{s,i}, \delta_{s,i}, \delta^{\rm dbl}_{s,i}\}$ *are defined as follows.*

$$\begin{aligned}
\phi_{2,i} &= 0 && \textit{all } i\\
\phi_{s,1} &= \phi_{s-1,1} + 1 && s \geqslant 3\\
\phi_{s,i} &= \phi_{s,i-1} + \phi_{s-1,i} + 1 && s \geqslant 3, i \geqslant 2\\
\delta_{1,i} &= 1 && \textit{all } i\\
\delta_{2,i} &= 2 && \textit{all } i\\
\delta^{\rm dbl}_{1,i} &= 2 && \textit{all } i\\
\delta^{\rm dbl}_{2,i} &= 5 && \textit{all } i\\
\delta_{s,1} &= 2\delta_{s-1,1} = 2^{s-1} && s \geqslant 3\\
\delta^{\rm dbl}_{s,1} &= 2(\delta^{\rm dbl}_{s-1,1} + 1) = 2^{s+1} - 2^{s-2} - 2 && s \geqslant 3\\
\delta_{s,i} &= \begin{cases} 2\delta_{s-1,i} + \delta_{s-2,i}(\delta_{s,i-1} - 2) & s = 3 \textit{ or even } s \geqslant 4\\ 2\delta_{s-1,i} + 2\delta_{s-2,i}\phi_{s,i-1} + \delta_{s-3,i}\delta_{s,i-1} & \textit{odd } s \geqslant 5 \end{cases}\\
\delta^{\rm dbl}_{s,i} &= \begin{cases} \delta^{\rm dbl}_{s,i-1} + 2\delta^{\rm dbl}_{s-1,i} + (\delta^{\rm dbl}_{s-2,i} + 2)\delta_{s,i-1} & s = 3 \textit{ or even } s \geqslant 4\\ \delta^{\rm dbl}_{s,i-1} + 2\delta^{\rm dbl}_{s-1,i} + 2(\delta^{\rm dbl}_{s-2,i} + 2)\phi_{s,i-1} + (\delta^{\rm dbl}_{s-3,i} + 2)\delta_{s,i-1} & \textit{odd } s \geqslant 5 \end{cases}
\end{aligned}$$

When applying Lemma 7.6, the tightest bounds are obtained by setting $i = \alpha(n,m) + O(1)$, which is $\alpha(n) + O(1)$ whenever $j = O(1)$. Lemma 7.7 gives closed form bounds on the coefficients $\{\delta_{s,i}, \delta^{\rm dbl}_{s,i}, \phi_{s,i}\}$, which immediately yield sharp bounds on the extremal functions $\lambda_s(n,m)$ and $\lambda^{\rm dbl}_s(n,m)$ for DS and double DS sequences partitioned into blocks.

Lemma 7.7. (Closed Form Bounds) *For all* $s \geqslant 3, i \geqslant 1$*, we have*

$$\begin{aligned}
\phi_{s,i} &= \binom{i+s-2}{s-2} - 1\\
\delta_{3,i} &= 2i + 2\\
\delta^{\rm dbl}_{3,i} &= \Theta(i^2)\\
\delta_{4,i}, \delta^{\rm dbl}_{4,i} &= \Theta(2^i)\\
\delta_{5,i}, \delta^{\rm dbl}_{5,i} &= \Theta(i2^i)\\
\delta_{s,i}, \delta^{\rm dbl}_{s,i} &\leqslant 2^{\binom{i+O(1)}{t}} && \textit{where } t = \lfloor \tfrac{s-2}{2} \rfloor.
\end{aligned}$$

Proof. The expression for $\phi_{s,i}$ holds in the base cases, when $s = 2$ or $i = 1$. By Pascal's identity it holds in general since

$$\phi_{s,i} = \phi_{s,i-1} + \phi_{s-1,i} + 1 = \binom{i+s-3}{s-2} + \binom{i+s-3}{s-3} - 1 = \binom{i+s-2}{s-2} - 1.$$

When $s \in \{3,4\}$, $\delta_{s,i}$ and $\delta^{\rm dbl}_{s,i}$ are identical to $\pi_{s,i}$ and $\pi^{\rm dbl}_{s,i}$, and therefore satisfy the same bounds from Lemma 4.5. Define C_4 such that $\delta_{4,i} \leqslant 2^{i+C_4}$. Assuming inductively that for some sufficiently large C_5, $\delta_{5,i-1} \leqslant (i-1)2^{(i-1)+C_5}$, we have

$$\begin{aligned}
\delta_{5,i} &\leqslant 2\delta_{4,i} + 2\delta_{3,i}\phi_{5,i-1} + \delta_{2,i}\delta_{5,i-1} \\
&\leqslant 2^{i+C_4+1} + 2(2i+2)\cdot\binom{i+2}{3} + 2\cdot(i-1)2^{i-1+C_5} \\
&\leqslant i2^{i+C_5}.
\end{aligned}$$

We claim that there are constants $\{C_s\}$ such that, for all $s > 5$, $\delta_{s,i} \leqslant 2^{\binom{i+C_s}{t}}$. When $s > 4$ is even,

$$\begin{aligned}
\delta_{s,i} &\leqslant 2\delta_{s-1,i} + \delta_{s-2,i}\delta_{s,i-1} \\
&\leqslant 2^{\binom{i+C_{s-1}}{t-1}+1} + 2^{\binom{i+C_{s-2}}{t-1}}2^{\binom{i-1+C_s}{t}} \\
&\leqslant 2^{\binom{i+C_s}{t}}, \;\text{ for some } C_s > C_{s-1} > C_{s-2}.
\end{aligned}$$

When $s > 5$ is odd, whether $s-2 = 5$ or not, $\delta_{s-2,i} \leqslant i2^{\binom{i+C_{s-2}}{t-1}}$ by the inductive hypothesis, so

$$\begin{aligned}
\delta_{s,i} &\leqslant 2\delta_{s-1,i} + 2\delta_{s-2,i}\phi_{s,i-1} + \delta_{s-3,i}\delta_{s,i-1} \\
&\leqslant 2^{\binom{i+C_{s-1}}{t}+1} + i2^{\binom{i+C_{s-2}}{t-1}+1}\cdot\binom{i+s-3}{s-2} + 2^{\binom{i+C_{s-3}}{t-1}}2^{\binom{i-1+C_s}{t}} \\
&\leqslant 2^{\binom{i+C_{s-1}}{t}+1} + i2^{\binom{i+C_{s-2}}{t-1}+1}\cdot\binom{i+s-3}{s-2} + 2^{-(C_s-C_{s-3})}2^{\binom{i+C_s}{t-1}+\binom{i-1+C_s}{t}} && (10) \\
&\leqslant 2^{\binom{i+C_s}{t}}. && (11)
\end{aligned}$$

Inequality (10) follows since $t-1 \geqslant 1$ and Inequality (11) follows since, for C_s sufficiently large, $2^{\binom{i+C_s}{t}}$ dominates both $\mathrm{poly}(i)\cdot 2^{\binom{i+C_{s-2}}{t-1}}$ and $2^{\binom{i+C_{s-1}}{t}+1}$. It is straightforward to show the same bounds hold on $\delta_{s,i}^{\mathrm{dbl}}$, for $s \geqslant 4$, with respect to different constants $\{D_s\}$. That is, $\delta_{s,i}^{\mathrm{dbl}} \leqslant 2^{\binom{i+D_s}{t}}$ when $s \neq 5$ and $\delta_{5,i}^{\mathrm{dbl}} \leqslant i2^{i+D_5}$. □

Choosing $i = \alpha(n,m) + O(1)$, Lemmas 7.6 and 7.7 imply that

$$\begin{aligned}
\lambda_3(n,m) &= O((n+m)\alpha(n,m)) \\
\lambda_3^{\mathrm{dbl}}(n,m) &= O((n+m)\alpha^2(n,m)) \\
\lambda_4(n,m), \lambda_4^{\mathrm{dbl}}(n,m) &= O((n+m)2^{\alpha(n,m)}) \\
\lambda_5(n,m), \lambda_5^{\mathrm{dbl}}(n,m) &= O((n+m)\alpha(n,m)2^{\alpha(n,m)}) \\
\lambda_s(n,m), \lambda_s^{\mathrm{dbl}}(n,m) &= O((n+m)2^{\alpha^t(n,m)/t! + O(\alpha^{t-1}(n,m))})
\end{aligned}$$

When $m = O(n)$ these bounds are all sharp, with the exception of λ_3^{dbl}, which was already handled in Section 6. Using the best transformations from 2-sparse to blocked sequences from Lemma 3.1, we obtain all the bounds on λ_s and λ_s^{dbl} claimed in Theorem 1.3, except at $s = 5$, where we only get $\lambda_5(n) = O(\alpha(\alpha(n)))\cdot\lambda_5(n,3n)$ and $\lambda_5^{\mathrm{dbl}}(n) = O(\alpha(\alpha(n)))\cdot\lambda_5^{\mathrm{dbl}}(n,3n)$. Refer to [22, §7.3] for an ad hoc method to eliminate this $\alpha(\alpha(n))$ factor.

8 Generalized Constructions of Nonlinear Sequences

Recall from Section 2.1 that the difference between postshuffling and preshuffling is in how blocks of one sequence are merged with copies of another. In $U_{\mathrm{sub}} \circledcirc U_{\mathrm{bot}}$ symbols from U_{sub} are inserted at the *end* of blocks in copies of U_{bot} whereas in $U_{\text{sub}} \otimes U_{\text{bot}}$ they are inserted at the beginning of blocks. It is not immediately clear why these two shuffling strategies should yield sequences with different properties. Consider the projection of symbols $R = \{a, \ldots, z\}$ in a common block B of U_{top}, where all symbols in R are middle occurrences in B. If U_{top} was constructed via a series of composition and *post*shuffling operations, the projection of U_{top} onto R, ignoring repetitions, is $ab\cdots z(zy\cdots a)zy\cdots a$, whereas if *pre*shuffling were used the projection onto R would be $ab\cdots z(ab\cdots z)zy\cdots a$. In a subsequent composition event $U_{\text{sub}} = U_{\text{top}} \circ U_{\text{mid}}$, the canonical ordering of R in $U_{\text{mid}}(B)$ is identical to their ordering in U_{top}, in the case of preshuffling, or the reversal of that ordering in the case of postshuffling.

In this section we explore the complexity of sequences avoiding "zig-zagging" patterns, which can be viewed as one natural generalization of Davenport-Schinzel sequences. Recall the definitions of N_k, M_k, and Z_k.

$$\begin{aligned} N_k &= 12\cdots(k+1)k\cdots 12\cdots(k+1) \\ M_k &= 12\cdots(k+1)k\cdots 12\cdots(k+1)k\cdots 1 \\ Z_k &= 12\cdots(k+1)k\cdots 12\cdots(k+1)k\cdots 12\cdots(k+1) \end{aligned}$$

Note that $N_1 = abab, M_1 = ababa$, and $Z_1 = ababab$ generalize order-2, -3, and -4 Davenport-Schinzel sequences. Klazar and Valtr [14] and Pettie [20] proved that $\mathrm{Ex}(N_k, n) = \Theta(\lambda_2(n)) = \Theta(n)$ and that for any $k \geqslant 1$, $\mathrm{Ex}(\{M_k, ababab\}, n) = \Theta(\lambda_3(n)) = \Theta(n\alpha(n))$. (That is, avoiding *both* M_k and *ababab* are equivalent to just avoiding M_1.) One might guess that zig-zagging patterns, in general, mimic the behavior of the corresponding order-s DS sequences.

We prove two results that, taken together, are rather surprising. Theorems 8.5 and 8.6 state the following in a more precise fashion.

(1) For all t, there exists a k such that $\mathrm{Ex}(M_k, n) = \Omega(n\alpha^t(n))$.

(2) For all t, there exists a k such that $\mathrm{Ex}(Z_k, n) = \Omega(n2^{(1+o(1))\alpha^t(n)/t!})$.

Overview. We define two classes of non-linear sequences. Class I sequences have lengths $\Theta(n\alpha^t(n))$ and Class II sequences have length $n2^{(1+o(1))\alpha^t(n)/t!}$, for any $t \geqslant 1$. Both Class I and Class II sequences are parameterized by a binary *pattern* $\pi = \pi_1\pi_2\cdots\pi_{|\pi|} \in \{\diagup, \diagdown\}^*$. The diagonals in π have the following interpretation. Consider any set $\{a_1, \ldots, a_l\}$ of symbols in a sequence T_π of type π. A *maximally* intertwined configuration is one in which each pair of symbols in $\{a_1, \ldots, a_l\}$ alternate the maximum number of times. In T_π all maximally intertwined configurations will take the form $A^{\pi_1}A^{\pi_2}\cdots A^{\pi_{|\pi|}}$, where $A^{\diagup} = a_1\cdots a_l$ and $A^{\diagdown} = a_l\cdots a_1$. Class I and II sequences are defined in Sections 8.1 and 8.2 and their forbidden sequences analyzed in Section 8.3.

8.1 Class I Sequences

The sequence $T_\pi(i,j)$ consists of a mixture of live and dead blocks. It is parameterized by a pattern π, which always begins with $\diagup$. The base cases for T_π are given below. (Recall that live blocks are indicated with parentheses and dead blocks with angular brackets.)

$$
\begin{aligned}
T_{\diagup\diagdown}(i,j) &= (12\cdots j)\ \langle j\cdots 21\rangle && \text{one live block, one dead, for any } i\\
T_{\diagup\diagup}(i,j) &= (12\cdots j)\ \langle 12\cdots j\rangle && \text{one live block, one dead, for any } i\\
T_\pi(1,j) &= \begin{cases} (12\cdots j)\ \langle j\cdots 21\rangle & \text{if } \pi_{|\pi|} = \diagdown \text{ and } |\pi|>2\\ (12\cdots j)\ \langle 12\cdots j\rangle & \text{if } \pi_{|\pi|} = \diagup \text{ and } |\pi|>2 \end{cases}\\
T_\pi(i,0) &= (\,)^2 && \text{two empty live blocks, any } \pi
\end{aligned}
$$

Note that $T_\pi(1,j)$ is identical to either $T_{\diagup\diagdown}(\cdot,j)$ or $T_{\diagup\diagup}(\cdot,j)$, depending on the last character of π. For the inductive case, when $i>1, j>0$, and $|\pi|>2$,

$$
T_\pi(i,j) = \begin{cases} T_{\mathrm{sub}} \olessthan T_{\mathrm{bot}} = (T_{\mathrm{top}} \circ T_{\mathrm{mid}}) \olessthan T_{\mathrm{bot}} & \text{if } \pi_{|\pi|} = \diagdown\\[2ex] T_{\mathrm{sub}} \ogreaterthan T_{\mathrm{bot}} = (T_{\mathrm{top}} \circ T_{\mathrm{mid}}) \ogreaterthan T_{\mathrm{bot}} & \text{if } \pi_{|\pi|} = \diagup \end{cases}
$$

$$
\begin{aligned}
\text{where}\quad T_{\mathrm{bot}} &= T_\pi(i,j-1)\\
T_{\mathrm{mid}} &= T_{\pi^-}(i, \llparenthesis T_{\mathrm{bot}} \rrparenthesis) && \pi^- = \pi_1\cdots\pi_{|\pi|-1}\\
T_{\mathrm{top}} &= T_\pi(i-1, \|T_{\mathrm{mid}}\|)
\end{aligned}
$$

The following facts can easily be proved about $T_\pi(i,j)$ by induction.

1. The first occurrence of every symbol appears in a live block and live blocks consist solely of first occurrences.
2. All live blocks have length exactly j. The length of dead blocks varies, as does the number of dead blocks between consecutive live blocks.
3. Each symbol occurs with the same multiplicity, $\nu_{\pi,i}$, defined below. Hence $|T| = \nu_{\pi,i}\|T\| = \nu_{\pi,i}\cdot j\cdot \llparenthesis T \rrparenthesis$.

The construction of T_π gives us an inductive expression for the multiplicity $\nu_{\pi,i}$ of symbols in $T_\pi(i,j)$.

$$
\begin{aligned}
\nu_{\pi,i} &= 2 && \text{for } |\pi|=2 \text{ and all } i\\
\nu_{\pi,1} &= 2 && \text{for all } \pi\\
\nu_{\pi,i} &= \nu_{\pi,i-1} + \nu_{\pi^-,i} - 1 && \text{where } \pi^- = \pi_1\cdots\pi_{|\pi|-1}
\end{aligned}
$$

A short proof by induction shows that $\nu_{\pi,i}$ has the closed form

$$
\nu_{\pi,i} = \binom{i+|\pi|-3}{|\pi|-2} + 1 \qquad \text{for all } i\geqslant 1, |\pi|\geqslant 2
$$

It can be shown that $i = \alpha(n,m) + O(1)$, where $n = \|T_\pi(i,j)\|$ and $m = \llbracket T_\pi(i,j)\rrbracket)$, from which it follows that $T_\pi(i,j)$ has length $\Theta(n\alpha^{|\pi|-2}(n,m))$, and length $\Theta(n\alpha^{|\pi|-2}(n))$ if $j=O(1)$. Theorem 8.1 summarizes two results from [21, 19, 9] using the T_π notation.

Theorem 8.1. *([9, 21, 19])*

1. $ababa, abcaccbc \not\prec T_{\diagup\diagdown\diagup}$.
2. $abaaba, abcacbc \not\prec T_{\diagup\diagup\diagdown}$.

As a consequence both $\mathrm{Ex}(ababa, n)$ *and* $\mathrm{Ex}(abcacbc, n)$ *are* $\Omega(n\alpha(n))$, *which is asymptotically tight.*

8.2 Class II Sequences

Class II Sequences consist *solely* of live blocks. They are parameterized by binary patterns, which are restricted to being even-length palindromes, starting with $\diagup$ and ending with $\diagdown$. If $\pi = \pi_1 \cdots \pi_{|\pi|}$, its *flip* $\mathrm{flip}(\pi)$ is obtained by flipping the direction of each diagonal and its *truncation* π^- is obtained by trimming π_1 and $\pi_{|\pi|}$. For example, if $\pi = \diagup\diagdown\diagdown\diagdown\diagup\diagup\diagup\diagdown$, $\mathrm{flip}(\pi^-) = \diagup\diagup\diagup\diagdown\diagdown\diagdown$.

The base cases for U_π are given below. The sequence $U_\pi(i,j)$ has the property that each block has length j and each symbol has multiplicity $\mu_{\pi,i}$, which will be defined below.

$$
\begin{array}{lr}
U_{\diagup\diagdown}(i,j) = (12\cdots j)\,(j\cdots 21) & \text{two blocks, for any } i \\
U_\pi(1,j) = (12\cdots j)\,(j\cdots 21) & \text{two blocks, for any } \pi \\
U_\pi(0,j) = (12\cdots j) & \text{one block, for any } \pi \\
U_\pi(i,1) = (1)^{\mu_{\pi,i}} & \mu_{\pi,i} \text{ identical blocks}
\end{array}
$$

For the inductive case, when $i > 1, j > 0$, and $|\pi| > 2$, we have

$$
U_\pi(i,j) = \begin{cases} U_{\mathrm{sub}} \olessthan U_{\mathrm{bot}} = (U_{\mathrm{top}} \circ T_{\mathrm{mid}}) \olessthan U_{\mathrm{bot}} & \text{if } \pi_2\pi_{|\pi|-1} = \diagup\diagdown \\ U_{\mathrm{sub}} \ogreaterthan U_{\mathrm{bot}} = (U_{\mathrm{top}} \circ T_{\mathrm{mid}}) \ogreaterthan U_{\mathrm{bot}} & \text{if } \pi_2\pi_{|\pi|-1} = \diagdown\diagup \end{cases}
$$

$$
\text{where} \quad U_{\mathrm{bot}} = U_\pi(i, j-1)
$$

$$
U_{\mathrm{mid}} = \begin{cases} U_{\pi^-}(i, [\![T_{\mathrm{bot}}]\!]) & \text{if } \pi_2\pi_{|\pi|-1} = \diagup\diagdown \\ U_{\mathrm{flip}(\pi^-)}(i, [\![T_{\mathrm{bot}}]\!]) & \text{if } \pi_2\pi_{|\pi|-1} = \diagdown\diagup \end{cases}
$$

$$
U_{\mathrm{top}} = U_\pi(i-1, \|T_{\mathrm{mid}}\|)
$$

The construction of U_π is a strict generalization of the U_s sequences defined in Section 2, for even s. Note that when $\pi = (\diagup\diagdown)^{s/2}$, only postshuffling is used, since $\mathrm{flip}(\pi^-) = (\diagup\diagdown)^{s/2-1}$. The multiplicity $\mu_{\pi,i}$ of symbols in $U_\pi(i,j)$ is not affected by which shuffling operation is used, so the analysis from Section 2 still holds: $\mu_{\pi,i} = 2^{\binom{i+t-1}{t}} \geqslant 2^{i^t/t!}$, where $t = (|\pi|-2)/2$, and $i = \alpha(\|U_\pi(i,j)\|, [\![U_\pi(i,j)]\!]) + O(1)$.

8.3 Analysis of T_π and U_π

Lemmas 8.2 and 8.3 isolate some properties of T_π useful in the analysis of M-shaped sequences and comb-shaped sequences.

Lemma 8.2. *Let* $T_{\mathrm{sh}} = T_\pi(i,j)$, *where* i *and* j *are arbitrary. Let* $\chi = \pi_{|\pi|}$ *and* $\chi' = \pi_{|\pi|-1}$ *be the last and second to last characters of* π, *and let* $T_{\mathrm{top}}, T_{\mathrm{mid}}, T_{\mathrm{sub}}$, *and* T_{bot} *be the sequences arising in the formation of* T_{sh}.

1. *If $abba \prec T_{\text{sh}}$ or $baba \prec T_{\text{sh}}$ then it cannot be that $b \in \Sigma(T_{\text{sub}})$ while $a \in \Sigma(T^*_{\text{bot}})$.*

2. *If $a < b$ share a live block in one of $T_{\text{top}}, T_{\text{bot}}$, or T_{sh}, then this sequence's projection onto $\{a, b\}$ has the form $(ab)a^*b^*$ if $\chi = \diagup$ and $(ab)b^*a^*$ if $\chi = \diagdown$.*

3. *If $a_1 < \cdots < a_l$ share a live block in T_{sub}, then its projection onto $\{a_1, \ldots, a_l\}$ has the form $(a_1 \ldots a_l)A^{\chi'}A^{\chi}$ where $A^{\diagup} = a_1^* \ldots a_l^*$ and $A^{\diagdown} = a_l^* \cdots a_1^*$.*

Lemma 8.3. *Whereas $ababa \nprec T_{\diagup\diagdown\diagup}$, $abaaba \nprec T_{\pi}$, for any $\pi \in \{\diagup\diagup\diagdown, \diagup\diagdown\diagdown, \diagup\diagup\diagup\}$.*

Proof. Lemma 8.2(1) implies that $ababa$ cannot be introduced by a shuffling event, but must first appear in $T_{\text{sub}} = T_{\text{top}} \circ T_{\text{mid}}$ from a composition event. Moreover, $abaaba$ could not arise in T_{sub} from an occurrence of $ababa$ in T_{top} since, in such an occurrence, the middle a would necessarily be in a dead block and could therefore not beget multiple as in T_{sub}. It must be that a and b share a common live block in T_{top}, so its projection onto $\{a, b\}$ is contained in $(ab)a^*b^*$, if $\pi_3 = \diagup$, and $(ab)b^*a^*$ if $\pi_3 = \diagdown$. Since T_{mid} is either $T_{\diagup\diagup}$ or $T_{\diagup\diagdown}$, the projection of T_{sub} onto $\{a, b\}$ is one of

$$(ab)\ \langle ba\rangle\ a^*b^* \quad \text{or} \quad (ab)\ \langle ba\rangle\ b^*a^* \quad \text{or} \quad (ab)\ \langle ab\rangle\ a^*b^* \quad \text{or} \quad (ab)\ \langle ab\rangle\ b^*a^*.$$

The first is $ababa$-free while the remaining are $abaaba$-free. □

In Theorem 8.5 we prove that $\text{Ex}(M_{2^k}, n) = \Omega(n\alpha^{k+1}(n))$ by induction. Lemma 8.4 handles the base case for M_2.

Lemma 8.4. *$M_2 = abcbabcba \nprec T_{\pi}$, for any of the length-4 patterns $\pi \in \diagup\{\diagup, \diagdown\}^2\diagup$.*

Proof. Since M_2 contains $ababa$, any instance of M_2 must first arise in $T_{\text{sub}} = T_{\text{top}} \circ T_{\text{mid}}$ from a composition event, not in $T_{\text{sh}} = T_{\text{sub}} \circledS T_{\text{mid}}$ from a shuffling event. Here T_{mid} is defined by any of the four patterns $\pi^- \in \diagup\{\diagup, \diagdown\}^2$. It must be that a, b, c share a live block in T_{top}. If only b and c shared a live block then the projection of T_{top} onto $\{a, b, c\}$ would have the form $a^*(bc \text{ or } cb)a^*b^*c^*b^*a^*$, violating Lemma 8.2 since neither (bc) and (cb) can be followed by bcb. If only a and b shared a live block the projection onto $\{a, b, c\}$ would have the form $a^*b^*c^*(ba \text{ or } ab)c^*b^*a^*$, which violates the property that live blocks contain only first occurrences.

We have deduced that a, b, and c share a live block B in T_{top}, but they do not necessarily appear in that order. To form a copy of M_2, some prefix must arise from substituting the type π^- sequence $T_{\text{mid}}(B)$ for B; the remaining suffix must follow a, b, and c's live block in T_{top}. The split between prefix and suffix can be (i) $abcbab\,|\,cba$, or (ii) $abcbabc\,|\,ba$, or (iii) $abcbabcb\,|\,a$. In cases (i) and (ii), b must precede a in B, meaning $b < a$ in the canonical ordering of $T_{\text{mid}}(B)$. As a consequence, any occurrence of the prefix $abcbab$ (or $abcbabc$) in T_{mid} implies an occurrence of $babbab \prec T_{\text{mid}}$, contradicting Lemma 8.3. In case (iii) the prefix contains $bcbbcb$, also contradicting Lemma 8.3. □

Theorem 8.5. *For any $k \geqslant 1$, $M_{2^k} \nprec T_{\pi}$, where $\pi \in \diagup\{\diagup, \diagdown\}^2\diagup^k$. As a consequence,* $\text{Ex}(M_{2^k}, n) = \Omega(n\alpha^{k+1}(n))$.

Proof. The proof is by induction on k; the base case is covered by Lemma 8.4. For succinctness let $K = 2^k$. As in the proof of Lemma 8.4 we can restrict our attention to the case where M_K, say over the alphabet $a_1, \ldots, a_{K+1}$, arises in T_{sub} after a composition event. Moreover, we can assume $a_1, \ldots, a_{K+1}$ appear in a common live block B, so the projection of T_{top} onto $\{a_1, \ldots, a_{K+1}\}$ is $(a_1 \cdots a_{K+1})a_1^* \cdots a_{K+1}^*$. If substituting $T_{\text{mid}}(B)$ for B creates an instance of M_K, some prefix must come from $T_{\mathrm{mid}}(B)$ and the remaining suffix from the sequence $a_1^* \cdots a_{K+1}^*$ following B. There are two cases: either the suffix contains a strict majority of the $K+1$ symbols or a strict minority. In the former case we have $a_{K/2+1} < \cdots < a_{K+1}$ according to the canonical ordering of $T_{\mathrm{mid}}(B)$, so any instance of the N-shaped pattern $a_{K+1}a_K \cdots a_{K/2+1}a_{K/2+2} \cdots a_{K+1}a_K \cdots a_{K/2+1}$ in $T_{\mathrm{mid}}(B)$ implies that it also contains

$$M_{K/2} = a_{K/2+1}\; a_{K/2+2}\; \cdots\; a_{K+1}\; a_K\; \cdots\; a_{K/2+1}\; a_{K/2+2}\; \cdots\; a_{K+1}\; a_K\; \cdots\; a_{K/2+1},$$

which contradicts the hypothesis that T_{mid} is $M_{K/2}$-free. If, on the other hand, the suffix of M_K following B contains a strict minority of $\{a_1, \ldots, a_{K+1}\}$, then $T_{\mathrm{mid}}(B)$ must contain an instance of $M_{K/2}$ on the alphabet $a_1, \ldots, a_{K/2+1}$, also contradicting the inductive hypothesis. $\square$

We now turn to the analysis of the forbidden sequences of U_π.

Theorem 8.6. *For any $k \geqslant 0$, $Z_{3^k} \nprec U_\pi$, where $\pi = \diagup^{k+1}\diagdown\diagup\diagdown^{k+1}$. As a consequence,* $\mathrm{Ex}(Z_{3^k}, n) > n \cdot 2^{(1+o(1))\alpha^{k+1}(n)/(k+1)!}$.

Proof. The proof is by induction on k. For succinctness we let $K = 3^k$. In the base case $k = 0$, $Z_K = ababab$, and $U_\pi = U_{\diagup\diagdown\diagup\diagdown}$ is $ababab$-free, by Lemma 2.4. In the general case $k \geqslant 1$ and $\pi = \diagup^{k+1}\diagdown\diagup\diagdown^{k+1}$, so $U_\pi = U_{\mathrm{sub}} \circledS U_{\mathrm{bot}} = (U_{\mathrm{top}} \circ U_{\mathrm{mid}}) \circledS U_{\mathrm{bot}}$ is formed by composing U_{top} with U_{mid}, a type π^- sequence, then *pre*shuffling it with U_{bot}. We can assume that any occurrence of Z_K arises from the composition event $U_{\mathrm{sub}} = U_{\mathrm{top}} \circ U_{\mathrm{mid}}$ since $ababab \prec Z_K$ cannot be introduced by shuffling. Write Z_K as

$$a_1\; a_2\; \cdots\; a_{K+1}\; a_K\; \cdots\; a_1\; a_2\; \cdots\; a_{K+1}\; a_K\; \cdots\; a_1\; a_2\; \cdots\; a_{K+1}.$$

It is easy to verify that if Z_K occurs in U_{sub}, it must be that $\{a_1, \ldots, a_{K+1}\}$ share a *single* block B in U_{top}. (Note, however, that their canonical orderings in U_{top} and $U_{\mathrm{mid}}(B)$ are *not* necessarily $a_1 < \cdots < a_{K+1}$.) Some prefix of Z_K appears before B in U_{top}, some suffix of Z_K after B in U_{top}, and the remaining middle portion appears in $U_{\mathrm{mid}}(B)$. Suppose $a_1 \cdots a_l$ is the prefix and $a_{l'}a_{l'+1} \cdots a_{K+1}$ the suffix, for some indices l, l'. It follows that $a_1 < a_2 < \cdots < a_l$ and $a_{K+1} < a_K < \cdots < a_{l'}$ according to the canonical ordering of $U_{\mathrm{mid}}(B)$, which implies $l \leqslant l'$.[18] At least one of the following must be true

(i) the prefix contains at least $K/3+1$ symbols and is disjoint from the suffix, that is, $l \geqslant K/3+1$ and $l < l'$.

(ii) the suffix contains at least $K/3+1$ symbols and we are not in case (i), that is, $l' \leqslant 2K/3+1$.

(iii) there are at least $K/3+1$ symbols in neither the prefix nor suffix, that is, $l \leqslant K/3$ and $l' \geqslant 2K/3+2$.

Case (iii) is the simplest. To form a copy of Z_K in U_{sub}, we would need $U_{\mathrm{mid}}(B)$ to contain a copy of $Z_{K/3}$ on the alphabet $\{a_{K/3+1}, \ldots, a_{2K/3+1}\}$, contradicting the inductive hypothesis. In Case (i), $U_{\mathrm{mid}}(B)$ must contain $a_{K/3+1} \cdots a_1 \cdots a_{K/3+1} \cdots a_1 \cdots a_{K/3+1}$, which, by the canonical ordering $a_1 < \cdots < a_{K/3+1}$, implies $U_{\mathrm{mid}}(B)$ also contains a copy of $Z_{K/3}$, a contradiction. Case (ii) is symmetric to Case (i). $\square$

[18]Since preshuffling is used, the canonical ordering of *middle* symbols in B is the same in U_{top} and $U_{\mathrm{mid}}(B)$, though the same is not true of symbols making their first appearance in B.

8.4 Comb-shaped Sequences

The results of [9, 14, 19, 20] show that *ababa* and *abcacbc* are the only minimally non-linear 2-sparse forbidden sequences over a three-letter alphabet, both with extremal function $\Theta(n\alpha(n))$. Just as *ababa* can be generalized to M-shaped sequences, $C_1 = abcacbc$ can be generalized to the one-sided *comb-shaped* sequences $\{C_k\}_{k \geqslant 1}$, where

$$C_k = 1\;\;2\;\;3\;\;\cdots\;\;(k+2)\;\;1\;\;(k+2)\;\;2\;\;(k+2)\;\;3\;\;(k+2)\;\;\cdots\;\;(k+1)\;\;(k+2).$$

Our parameterized sequences let us obtain non-trivial lower bounds on comb-shaped sequences.

Theorem 8.7. *For all $k \geqslant 1$, $C_k \nprec T_\pi$, where $\pi = {\diagup}{\diagup}{\diagdown}^k$. Consequently,* $\mathrm{Ex}(C_k, n) = \Omega(n\alpha^k(n))$.

Proof. The proof is by induction on k. Theorem 8.1 (see [19]) takes care of the base case $C_1 = abcacbc$. We will focus on $C_2 = abcdadbdcd$, then note why the argument works for any k. Define $T_{\mathrm{top}}, T_{\mathrm{sub}}, T_{\mathrm{bot}}$, and T_{mid} as usual, where T_{mid} is now a type ${\diagup}{\diagup}{\diagdown}$ sequence. We first argue that $\{a, b, c, d\} \subseteq \Sigma(T_{\mathrm{top}})$. One may check that the only case that does not immediately violate Lemma 8.2(1) is that $a \in \Sigma(T^*_{\mathrm{bot}})$ while $b, c, d \in \Sigma(T_{\mathrm{top}})$. This means that $(bcd)dbdcd \prec T_{\mathrm{sub}}$, where the live block (bcd) was shuffled into a's copy of T_{bot}. However, Lemma 8.2(3) implies that the projection of T_{sub} onto $\{b, c, d\}$ has the form $(bcd)d^*c^*b^*d^*c^*b^*$, which does not contain $(bcd)dbdcd$.

One can see that a, b, c, and d must share a live block B in T_{top}. If the first two as in $C_2 \prec T_{\mathrm{sub}}$ arose from the composition that created T_{sub} then b, c, and d must have been in a's live block. If not then C_2 would have already appeared in T_{top}. Thus, some prefix of C_2 arose from substituting $T_{\mathrm{mid}}(B)$ for B and the remaining suffix followed B in T_{top}. Lemma 8.2(2) implies that the suffix cannot be dcd for otherwise $(cd)cd \prec T_{\mathrm{top}}$ or $(dc)dc \prec T_{\mathrm{top}}$. This implies that $abdadbd = C_1 \prec T_{\mathrm{mid}}(B)$ (a type ${\diagup}{\diagup}{\diagdown}$ sequence), which contradicts Theorem 8.1.

For $k > 2$ write $C_k = a_1a_2 \cdots a_{k+1}ba_1ba_2b \cdots ba_{k+1}b$. The same argument from above shows that $\{a_1, \ldots, a_{k+1}, b\}$ are contained in a single block B of T_{top}. For C_k to arise in T_{sub} a prefix of it must come from $T_{\mathrm{mid}}(B)$ and a suffix from the part of T_{top} following B. By Lemma 8.2(2) the suffix cannot be $ba_{k+1}b$, which means the prefix in $T_{\mathrm{mid}}(B)$ must contain $a_1 \cdots a_kba_1ba_2b \cdots ba_kb = C_{k-1}$, contradicting the inductive hypothesis. □

9 Conclusions

In Theorem 1.3 we established sharp bounds on the functions $\Lambda_{r,s}$ and $\Lambda^{\mathrm{dbl}}_{r,s}$, for all values of r and s, and showed, perhaps surprisingly, that these extremal functions are essentially the same. Moreover, they match λ_s and λ^{dbl}_s only when $s \leqslant 3$, or $s \geqslant 4$ is even, or $r = 2$. However, Theorem 1.3 is *not* the last word on $\Lambda^{\mathrm{dbl}}_{r,s}$. In Cibulka and Kynčl's [3] application of $\Lambda^{\mathrm{dbl}}_{r,s}(n, m)$, s is a fixed parameter whereas r is variable and cannot be bounded as a function of s. Cibulka and Kynčl require upper bounds on $\Lambda^{\mathrm{dbl}}_{r,s}(n, m)$ that are linear in r whereas the leading constant in our bounds matches that of $\Lambda^{\mathrm{dbl}}_{r,2}(n, m)$, currently known to be at most $O(6^r)$. See Lemma 3.3. In other words, we now have two incomparable upper bounds on $\Lambda^{\mathrm{dbl}}_{r,2}(n, m)$ when r is not treated as a constant, namely $O((n + rm)\alpha(n, m))$ [3], which is optimal as a function of r, and $O(6^r(n + m))$, which is optimal for fixed r. Whether $\Lambda^{\mathrm{dbl}}_{r,2}(n, m) = O(n + rm)$ or not is an intriguing open question.

We have shown that *doubling* various forbidden patterns (alternating sequences and catenated permutations) has no significant effect on their extremal functions. It is an open problem whether

$\text{Ex}(\text{dbl}(\sigma), n)$ is asymptotically equivalent to $\text{Ex}(\sigma, n)$ for *every* σ. We conjecture the answer is no when σ can be a *set* of forbidden sequences, though it seems plausible the answer is yes for any single forbidden sequence.

Conjecture 9.1. *In general, it is not true that* $\text{Ex}(\text{dbl}(\sigma), n) = \Theta(\text{Ex}(\sigma, n))$. *In particular, whereas* $\text{Ex}(\text{dbl}(\{ababa, abcacbc\}), n) = \Theta(n\alpha(n))$*, we conjecture* $\text{Ex}(\{ababa, abcacbc\}, n) = O(n)$.

The main open problem in the realm of generalized Davenport-Schinzel sequences is to characterize linear forbidden sequences, or equivalently, to enumerate all minimally non-linear forbidden sequences. The number of minimally non-linear sequences (with respect to the partial order $\prec$) is almost certainly infinite [19], but whether there are infinitely many *genuinely* different non-linear sequences is open. Refer to [19] for a discussion of how "genuinely" might be formally defined.

Conjecture 9.2. *(Informal) Every nonlinear sequence* σ *(having* $\text{Ex}(\sigma, n) = \omega(n)$*) contains ababa, abcacbc, or some sequence morally equivalent to abcacbc.*

Our lower bounds on $\text{Ex}(M_k, n)$ are weak, as a function of k, and we have provided no non-trivial upper bounds. It may be possible to generalize the proof of Theorem 6.9 to show $\text{Ex}(M_k, n) = O(n\,\text{poly}(\alpha(n)))$, where the degree of the polynomial depends on k.

References

[1] R. Adamec, M. Klazar, and P. Valtr. Generalized Davenport-Schinzel sequences with linear upper bound. *Discrete Math.*, 108(1-3):219–229, 1992.

[2] P. Agarwal, M. Sharir, and P. Shor. Sharp upper and lower bounds on the length of general Davenport-Schinzel sequences. *J. Combin. Theory Ser. A*, 52:228–274, 1989.

[3] J. Cibulka and J. Kynčl. Tight bounds on the maximum size of a set of permutations with bounded VC-dimension. *J. Combin. Theory Ser. A*, 119(7):1461–1478, 2012.

[4] H. Davenport and A. Schinzel. A combinatorial problem connected with differential equations. *American J. Mathematics*, 87:684–694, 1965.

[5] H. Davenport and A. Schinzel. A note on sequences and subsequences. *Elemente Der Mathematik*, 20:63–64, 1965.

[6] J. Fox, J. Pach, and A. Suk. The number of edges in k-quasi-planar graphs. *CoRR*, abs/1112.2361, 2011. To appear in SIAM J. Discrete Mathematics.

[7] Z. Füredi and P. Hajnal. Davenport-Schinzel theory of matrices. *Discrete Mathematics*, 103(3):233–251, 1992.

[8] J. Geneson, R. Prasad, and J. Tidor. Bounding sequence extremal functions with formations. *CoRR*, abs/1308.3810, 2013.

[9] S. Hart and M. Sharir. Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes. *Combinatorica*, 6(2):151–177, 1986.

[10] M. Klazar. A general upper bound in extremal theory of sequences. *Comment. Math. Univ. Carolin.*, 33(4):737–746, 1992.

[11] M. Klazar. Extremal functions for sequences. *Discrete Mathematics*, 150(1–3):195–203, 1996.

[12] M. Klazar. On the maximum lengths of Davenport-Schinzel sequences. In *Contemporary Trends in Discrete Mathematics, Štiřín Castle 1997 (Czech Republic), American Mathematical Society, Providence RI*, pages 169–178, 1999.

[13] M. Klazar. Generalized Davenport-Schinzel sequences: results, problems, and applications. *Integers*, 2:A11, 2002.

[14] M. Klazar and P. Valtr. Generalized Davenport-Schinzel sequences. *Combinatorica*, 14(4):463–476, 1994.

[15] P. Komjáth. A simplified construction of nonlinear Davenport-Schinzel sequences. *J. Combin. Theory Ser. A*, 49(2):262–267, 1988.

[16] G. Nivasch. Improved bounds and new techniques for Davenport-Schinzel sequences and their generalizations. *J. ACM*, 57(3), 2010.

[17] S. Pettie. Splay trees, Davenport-Schinzel sequences, and the deque conjecture. In *Proceedings 19th ACM-SIAM Symposium on Discrete Algorithms*, pages 1115–1124, 2008.

[18] S. Pettie. Degrees of nonlinearity in forbidden 0-1 matrix problems. *Discrete Mathematics*, 311:2396–2410, 2011.

[19] S. Pettie. Generalized Davenport-Schinzel sequences and their 0-1 matrix counterparts. *J. Comb. Theory Ser. A*, 118(6):1863–1895, 2011.

[20] S. Pettie. On the structure and composition of forbidden sequences, with geometric applications. In *Proceedings 27th Annual Symposium on Computational Geometry*, pages 370–379, 2011.

[21] S. Pettie. Origins of nonlinearity in Davenport-Schinzel sequences. *SIAM J. Discrete Mathematics*, 25(1):211–233, 2011.

[22] S. Pettie. Sharp bounds on Davenport-Schinzel sequences of every order. Technical Report arXiv:1204.1086v2, 2013. An extended abstract appears in the Proceedings of the 2013 Symposium on Computational Geometry, pp. 319–328.

[23] M. Sharir. Almost linear upper bounds on the length of general Davenport-Schinzel sequences. *Combinatorica*, 7(1):131–143, 1987.

[24] A. Suk. k-quasi-planar graphs. In *Graph Drawing*, pages 266–277, 2011.

[25] R. Sundar. On the deque conjecture for the splay algorithm. *Combinatorica*, 12(1):95–124, 1992.

[26] E. Szemerédi. On a problem of Davenport and Schinzel. *Acta Arith.*, 25:213–224, 1973/74.

[27] G. Tardos. On 0-1 matrices and small excluded submatrices. *J. Combin. Theory Ser. A*, 111(2):266–288, 2005.

[28] R. E. Tarjan. Efficiency of a good but not linear set merging algorithm. *J. ACM*, 22(2):215–225, 1975.

[29] P. Valtr. Graph drawings with no k pairwise crossing edges. In *Proceedings 5th Int'l Symposium on Graph Drawing*, pages 205–218, 1997.

[30] A. Wiernik and M. Sharir. Planar realizations of nonlinear Davenport-Schinzel sequences by segments. *Discrete Comput. Geom.*, 3(1):15–47, 1988.

A Proofs

A.1 Proof of Lemma 1.2

Recall that $\mathrm{dbl}(\mathrm{Perm}_{r,s+1}) = \{\mathrm{dbl}(\sigma) \mid \sigma \in \mathrm{Perm}_{r,s+1}\}$ whereas sequences in $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$ are formed by taking the concatenation of $s+1$ sequences, the first and last being a permutation of $\{1,\ldots,r\}$ and all the rest containing two occurrences of $\{1,\ldots,r\}$. For example, $abc\,abaccb\,bca \in \mathrm{Perm}^{\mathrm{dbl}}_{3,3}$ whereas $abbcc\,ccbbaa\,bbcca \in \mathrm{dbl}(\mathrm{Perm}_{3,3})$. We restate Lemma 1.2.

Lemma 1.2 *The following bounds hold for any $r \geqslant 2, s \geqslant 1$.*

$$\mathrm{Ex}(\mathrm{dbl}(\mathrm{Perm}_{r,s+1}), n, m) \leqslant r \cdot \Lambda^{\mathrm{dbl}}_{r,s}(n,m) + 2rn$$
$$\mathrm{Ex}(\mathrm{dbl}(\mathrm{Perm}_{r,s+1}), n) = O(\Lambda^{\mathrm{dbl}}_{r,s}(n)).$$

Proof. Let S be a $\mathrm{dbl}(\mathrm{Perm}_{r,s+1})$-free sequence over an n-letter alphabet. Obtain S' from S by discarding the first occurrence and last r occurrences of each letter, then retaining every rth occurrence of each letter, discarding the rest. Clearly S' has the property that each b is preceded and followed by at least r bs in S, and between two bs in S' there are at least $r-1$ bs in S. It follows that $|S'| \geqslant (|S| - 2rn)/r$. Suppose $|S'|$ contained some sequence $\sigma'_1 \cdots \sigma'_{s+1} \in \mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$. (Recall that σ'_1 and σ'_{s+1} contain one copy of $\{1,\ldots,r\}$ whereas $\sigma'_2,\ldots,\sigma'_s$ contain two copies of $\{1,\ldots,r\}$.) This implies that S contains a sequence $\sigma_1 \cdots \sigma_{s+1}$ where each σ_k contains $r+1$ copies of $\{1,\ldots,r\}$. We claim each σ_k contains a doubled permutation of $\{1,\ldots,r\}$, which implies that S is not $\mathrm{dbl}(\mathrm{Perm}_{r,s+1})$-free, a contradiction. Find the symbol b in σ_k whose second occurrence is earliest, that is, we can write $\sigma_k = \sigma'_k\, b\, \sigma''_k\, b\, \sigma'''_k$, where $\sigma'_k \sigma''_k$ contains at most one copy of each symbol. Since σ'''_k contains at least r copies of the $r-1$ symbols in $\{1,\ldots,r\}\backslash\{b\}$ we can continue to find a doubled permutation of $\{1,\ldots,r\}\backslash\{b\}$ by induction. If S is an m-block sequence then S' is too, giving the first bound. When S is merely r-sparse we can only bound S' by $\Lambda^{\mathrm{dbl}}_{r,s}(n)$ if it, too, is r-sparse. This is done as follows.

Greedily partition $S = S_1 S_2 \ldots S_m$ into maximal sequences $\{S_q\}$ over alphabets of size exactly $2r^2$, with $\|S_m\|$ perhaps smaller. Since each S_q has length at most $\mathrm{Ex}(\mathrm{dbl}(\mathrm{Perm}_{r,s+1}), 2r^2) = O(1)$, it follows that $m = \Omega(|S|)$. Obtain T be replacing each S_q with a block consisting of its alphabet $\Sigma(S_q)$. If $|T| \leqslant 2r^2 n$ there is nothing to prove since $|S| = \Theta(|T|) = O(n) = O(\Lambda^{\mathrm{dbl}}_{r,s}(n))$, so assume otherwise. Obtain T' from T by discarding the first occurrence and last r occurrences of each letter, then retaining every rth occurrence of each letter. It follows that $|T'| \geqslant (|T| - 2rn)/r \geqslant |T|\frac{r-1}{r^2}$, that is, the average length of blocks in T' is $2(r-1)$. Let T'' be an r-sparse subsequence of T' obtained by scanning T' from left to right, removing a symbol if it is identical to one of the preceding $r-1$ symbols. At most $r-1$ letters from each block of T' can be removed in this process. The average block length of T'' is at least $2(r-1)-(r-1) \geqslant 1$, hence $|T''| \geqslant m = \Omega(|S|)$. Since T'' is $\mathrm{Perm}^{\mathrm{dbl}}_{r,s+1}$-free, we have $|S| = O(\Lambda^{\mathrm{dbl}}_{r,s}(n))$. □

A.2 Proof of Lemma 3.1

There is no theorem to the effect that $\mathrm{Ex}(\sigma, n) = O(\mathrm{Ex}(\sigma, n, O(n)))$. Lemma 3.1 restates the best known reductions from r-sparse to blocked sequences. Some ad hoc reductions are known to be superior, for example, those for order-5 DS sequences [22].

Lemma 3.1 *(Cf. Sharir [23], Füredi and Hajnal [7], and Pettie [22].) Define $\gamma_s, \gamma_s^{\mathrm{dbl}}, \gamma_{r,s}, \gamma_{r,s}^{\mathrm{dbl}} : \mathbb{N} \to \mathbb{N}$ to be non-decreasing functions bounding the leading factors of $\lambda_s(n), \lambda_s^{\mathrm{dbl}}(n), \Lambda_{r,s}(n)$, and $\Lambda^{\mathrm{dbl}}_{r,s}(n)$, e.g., $\Lambda^{\mathrm{dbl}}_{r,s} \leqslant \gamma^{\mathrm{dbl}}_{r,s}(n) \cdot n$. The following bounds hold.*

$$\begin{aligned}
\lambda_s(n) &\leqslant \gamma_{s-2}(n) \cdot \lambda_s(n, 2n) & \lambda_s^{\mathrm{dbl}}(n) &\leqslant (\gamma_{s-2}^{\mathrm{dbl}}(n) + 4) \cdot \lambda_s^{\mathrm{dbl}}(n, 2n)\\
\lambda_s(n) &\leqslant \gamma_{s-2}(\gamma_s(n)) \cdot \lambda_s(n, 3n) & \lambda_s^{\mathrm{dbl}}(n) &\leqslant (\gamma_{s-2}^{\mathrm{dbl}}(\gamma_s^{\mathrm{dbl}}(n)) + 4) \cdot \lambda_s^{\mathrm{dbl}}(n, 3n)\\
\Lambda_{r,s}(n) &\leqslant \gamma_{r,s-2}(n) \cdot \Lambda_{r,s}(n, 2n) + 2n & \Lambda_{r,s}^{\mathrm{dbl}}(n) &\leqslant (\gamma_{r,s-2}^{\mathrm{dbl}}(n) + O(1)) \cdot \Lambda_s^{\mathrm{dbl}}(n, 2n))\\
\Lambda_{r,s}(n) &\leqslant \gamma_{r,s-2}(\gamma_{r,s}(n)) \cdot \Lambda_{r,s}(n, 3n) + 2n & \Lambda_{r,s}^{\mathrm{dbl}}(n) &\leqslant (\gamma_{r,s-2}^{\mathrm{dbl}}(\gamma_{r,s}^{\mathrm{dbl}}(n)) + O(1)) \cdot \Lambda_s^{\mathrm{dbl}}(n, 3n))
\end{aligned}$$

where the $O(1)$ term in the last two inequalities depends on r and s.

Proof. All the bounds are obtained from the following sequence manipulations, which were first used by Hart and Sharir [9] and Sharir [23]. Let S be an r-sparse sequence avoiding some set σ of subsequences over an r-letter alphabet, so $|S| \leqslant \mathrm{Ex}(\sigma, n)$. Greedily parse S into m intervals $S_1 S_2 \cdots S_m$ by choosing S_1 to be the maximum-length prefix satisfying some property $\mathcal{P}$, S_2 to be the maximum-length prefix of the remaining sequence satisfying $\mathcal{P}$, and so on. Form $S' = \Sigma(S_1)\Sigma(S_2)\cdots\Sigma(S_m)$ by replacing each interval S_i with a single block $\Sigma(S_i)$ containing its alphabet, listed in order of first appearance. Since S' is a subsequence of S, $|S'| \leqslant \mathrm{Ex}(\sigma, n, m)$. To bound $|S|$ we only need to determine upper bounds on m and the *shrinkage* factor $|S|/|S'|$.

Bounds on λ_s. If we parse S into maximal order-$(s-2)$ sequences then each S_i must contain either the first or last occurrence of some symbol, hence $m \leqslant 2n$. The shrinkage factor is $|S_i|/\|S_i\| \leqslant \gamma_{s-2}(\|S_i\|) \leqslant \gamma_{s-2}(n)$, which gives the first inequality. Now consider parsing S into m maximal sequences that are both order-$(s-2)$ DS sequences *and* have length at most $\gamma_s(n)$. It follows that $m \leqslant 3n$: at most n sequences were terminated because they reached length $\gamma_s(n)$ (by definition of γ_s) and the remaining sequences number at most $2n$ since each must contain the first or last occurrence of some letter.

Bounds on λ_s^{dbl}. Let σ_{s+2} be the alternating sequence with length $s+2$. Order-s double DS sequences are $\mathrm{dbl}(\sigma_{s+2})$-free. Obtain σ'_{s+2} by doubling each letter of σ_{s+2}, including the first and last. It is easy to show that $\mathrm{Ex}(\sigma'_{s+2}, n) \leqslant \lambda_s^{\mathrm{dbl}}(n) + 4n$ so we can take $\gamma_s^{\mathrm{dbl}}(n) + 4$ to be the leading factor in this extremal function. Consider parsing an order-s double DS sequence S. If we parse S into maximal σ_s'-free sequences then each subsequence must contain the first or last occurrence of some symbol, so $m \leqslant 2n$ and the shrinkage factor is at most $\gamma_{s-2}^{\rm dbl}(n) + 4$. If, further, we truncate any subsequence in the parsing at length $\gamma_s^{\rm dbl}(n)$, then $m \leqslant 3n$ and the shrinkage factor is at most $\gamma_{s-2}^{\rm dbl}(\gamma_s^{\rm dbl}(n)) + 4$.

Bounds on $\Lambda_{r,s}$ and $\Lambda_{r,s}^{\rm dbl}$. The argument is the same, except that during the parsing step, we discard any symbol that triggers the termination of a subsequence. For example, if S is a $\mathrm{Perm}_{r,s+1}$-free sequence we parse it into $S_1 a_1 S_2 a_2 \cdots a_{m-1} S_m a_m$, where the $\{S_i\}$ are maximal $\mathrm{Perm}_{r,s-1}$-free sequences and $\{a_i\}$ the single letters following them, where a_m might not be present. Since $S_i a_i$ contains some element of $\mathrm{Perm}_{r,s-1}$, $S_i a_i$ must contain the first or last occurrence of some letter, hence $m \leqslant 2n$. We form S' by contracting each S_i to a single block, discarding a_i, so the shrinkage factor is at most $\gamma_{r,s-2}(n)$. It follows that $|S| \leqslant \gamma_{r,s-2}(n) \cdot \Lambda_{r,s}(n, 2n) + 2n$. The procedure for $\Lambda_{r,s}^{\rm dbl}$ is a straightforward combination of the procedures described above, for $\Lambda_{r,s}$ and $\lambda_s^{\rm dbl}$. □

A.3 Proof of Lemma 3.2

We restate the lemma.

Lemma 3.2 *The extremal functions for order-s (double) Davenport-Schinzel sequences and* $\mathrm{Perm}_{2,s+1}$*-free (*$\mathrm{Perm}_{2,s+1}^{\rm dbl}$*-free) sequences are equivalent up to constant factors. In particular,*

$$\begin{aligned}
\lambda_s(n) \leqslant{}& \Lambda_{2,s}(n) && < 3 \cdot \lambda_s(n) + 2n \\
\lambda_s(n,m) \leqslant{}& \Lambda_{2,s}(n,m) && < 2 \cdot \lambda_s(n,m) + n \\
\lambda_s^{\rm dbl}(n) \leqslant{}& \Lambda_{2,s}^{\rm dbl}(n) && < 5 \cdot \lambda_s^{\rm dbl}(n) + 4n \\
\lambda_s^{\rm dbl}(n,m) \leqslant{}& \Lambda_{2,s}^{\rm dbl}(n,m) && < 3 \cdot \lambda_s^{\rm dbl}(n,m) + 2n
\end{aligned}$$

Proof. Order-s DS sequences are $\mathrm{Perm}_{2,s+1}$-free, which gives the 1st and 3rd inequalities. Let S be a 2-sparse $\mathrm{Perm}_{2,s+1}$-free sequence. Form $S' \prec S$ by filtering S as follows.

(i) Discard the 1st occurrence of each letter.

(ii) Discard up to n additional occurrences to restore 2-sparseness.

(iii) Discard every even occurrence of each letter.

(iv) Discard additional occurrences to restore 2-sparseness.

We claim S' has length at least $(|S| - 2n)/3$. The number of letters removed in steps (i) and (ii) is at most $2n$. The number removed in step (iii) is at most $(|S| - 2n)/2$ and the number removed in step (iv) is at most 1/3 of that of step (iii). This is because between any two even occurrences of some letter a, there must be another a and, due to 2-sparseness, at least two other letters. Thus, each letter removed in step (iv) corresponds to at least three removed in step (iii). Suppose that S' were not an order-s DS sequence, that it contained an alternating subsequence $a \cdots b \cdots a \cdots b \cdots$ with length $s + 2$. Together with the first occurrence of b and the missing odd occurrences of a and b from S, we can form a $\mathrm{Perm}_{2,s+1}$ subsequence in S, a contradiction. This gives the 2nd inequality. If S is composed of m blocks then we only need to form S' using steps (i) and (iii). The

4th inequality follows. The 5th and 7th inequalities follow since order-s double DS sequences are $\Lambda^{\mathrm{dbl}}_{2,s+1}$-free. To obtain the 6th inequality, let S be a 2-sparse $\Lambda^{\mathrm{dbl}}_{2,s+1}$-free sequence and S' be derived as follows.

(i) Discard the first and last occurrence of each letter.

(ii) Discard up to $2n$ additional occurrences to restore 2-sparseness.

(iii) Retain every third occurrence of each letter; discard all others.

(iv) Discard additional occurrences to restore 2-sparseness.

By the same argument as above, the number of letters discarded in step (iii) is at most $2(|S|-4n)/3$ and the number discarded in step (iv) at most 1/5th that of step (iii), hence $|S'| \geqslant (|S|-4n)/5$. Suppose S' contained a doubled alternating sequence $abbaabb\cdots$ having $s+2$ runs of as and bs. This implies that S contains $\underline{a}ab\underline{bb}ba\underline{aa}ab\underline{bb}b\cdots$, where the underlined letters appear in S but not S', and therefore that S contains an instance of $\mathrm{Perm}^{\mathrm{dbl}}_{2,s+1}$. The 6th inequality follows. The 8th follows from the same argument, omitting steps (ii) and (iv) in the construction of S'. $\square$

A.4 Proof of Lemma 3.3

Some of the results cited in Lemma 3.3 refer to (or implicitly use) results on forbidden 0-1 matrices. See Füredi and Hajnal [7] and Pettie [19, 18, 20] for more details on the connection between matrices and sequences.

Lemma 3.3 *At orders $s = 1$ and $s = 2$, the extremal functions $\lambda_s, \lambda^{\mathrm{dbl}}_s, \Lambda_{r,s}$, and $\Lambda^{\mathrm{dbl}}_{r,s}$ obey the following.*

$\lambda_1(n) = n$	$\lambda_1(n,m) = n+m-1$	
$\lambda_2(n) = 2n-1$	$\lambda_2(n,m) = 2n+m-2$	*(Davenport-Schinzel [4])*
$\lambda^{\mathrm{dbl}}_1(n) = 3n-2$	$\lambda^{\mathrm{dbl}}_1(n,m) = 2n+m-2$	*(Dav.-Sch. [5],Klazar [13])*
$\lambda^{\mathrm{dbl}}_2(n) < 8n$	$\lambda^{\mathrm{dbl}}_2(n,m) < 5n+m$	*(Klazar [11], Füredi-Hajnal [7])*
$\Lambda_{r,1}(n) = \Lambda^{\mathrm{dbl}}_{r,1}(n) < rn$	$\Lambda_{r,1}(n,m) = \Lambda^{\mathrm{dbl}}_{r,1}(n,m) < n+(r-1)m$	*(Klazar [10])*
$\Lambda_{r,2}(n) < 2rn$	$\Lambda_{r,2}(n,m) < 2n+(r-1)m$	*(Klazar [10])*
$\Lambda^{\mathrm{dbl}}_{r,2}(n) < 6^r rn$	$\Lambda^{\mathrm{dbl}}_{r,2}(n,m) < 2\cdot 6^{r-1}(n+m/3)$	*(Pettie [20], cf. [14])*

Proof. Davenport and Schinzel [4] noted the bounds on $\lambda_1(n)$ and $\lambda_2(n)$; their extension to blocked sequences is trivial. In an overlooked note Davenport and Schinzel [4] observed without proof that $\lambda^{\mathrm{dbl}}_1(n) = 3n-2$, which was formally proved by Klazar [13]. Its extension to blocked sequences is also trivial. Adamec, Klazar, and Valtr [1] proved that $\lambda^{\mathrm{dbl}}_2(n) = O(n)$ and Klazar [11] bounded the leading constant between 7 and 8. A blocked sequence S can be represented as a 0-1 incidence matrix A_S whose rows correspond to symbols and columns to blocks, where $A_S(i,j) = 1$ if and only if symbol i appears in block j. A forbidden sequence becomes a forbidden 0-1 pattern. The bound on $\lambda^{\mathrm{dbl}}_2(n,m)$ follows from Füredi and Hajnal's [7] analysis of a certain 0-1 pattern. The bounds on $\Lambda_{r,1}$ and $\Lambda_{r,2}$ were noted by Klazar [10] and Nivasch [16]. They are straightforward to prove.

Since the N-shaped sequence $12\cdots r\, r(r-1)\cdots 1\, 12\cdots r$ over r letters is contained in $\mathrm{Perm}_{r,3}$, the linear upper bound on $\mathrm{Ex}(\mathrm{dbl}(12\cdots r\, r(r-1)\cdots 1\, 12\cdots r), n)$ due to Klazar and Valtr [14] (see also [20]) immediately extend to $\Lambda^{\mathrm{dbl}}_{r,2}(n)$. With some care the leading constants of $\Lambda^{\mathrm{dbl}}_{r,2}(n)$ and $\Lambda^{\mathrm{dbl}}_{r,2}(n,m)$ can be made reasonably small using the 0-1 matrix representation of (forbidden) sequences from [20]. Consider an m-block, $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$-free sequence S. Without loss of generality assume the alphabet $\Sigma(S) = \{1,\ldots,n\}$ is ordered according to their first appearance in S. Let A_S be an $n \times m$ 0-1 matrix where $A_S(i,j) = 1$ if and only if symbol i appears in block j. By virtue of being $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$-free, A_S does not contain P as a submatrix,[19] where P is defined below. Following convention [27, 18] we use bullets for 1s and blanks for 0s.

$$P = \left(\begin{array}{c|ccccccccc|ccccc}
 & & & & & & & & \bullet & \bullet & \bullet & & & & \\
 & & & & & & \bullet & \bullet & & & & \bullet & & & \\
 & & & & & \iddots & & & & & & & \ddots & & \\
 & & & \bullet & \bullet & & & & & & & & & \bullet & \\
\bullet & \bullet & \bullet & & & & & & & & & & & & \bullet
\end{array}\right) \updownarrow r$$

The vertical bars are not part of the pattern; they mark the boundaries of the three components of a $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$ sequence. The results of [20] imply $\Lambda^{\mathrm{dbl}}_{r,2}(n,m) \leqslant \mathrm{Ex}(P,n,m) \leqslant 2 \cdot 6^{r-1}(n + m/3)$, where $\mathrm{Ex}(P,n,m)$ is the maximum number of 1s in P-free $n \times m$ matrix. To get a bound on $\Lambda^{\mathrm{dbl}}_{r,2}(n)$ we will show how to convert an r-sparse, $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$-free sequence S into a blocked one. Greedily partition $S = S_1 a_1 S_2 a_2 \cdots S_m$ into maximal $\mathrm{Perm}_{r,3}$-free sequences $S_1,\ldots,S_m$, separated by single symbols $a_1,\ldots,a_m$. That is, S_1 is $\mathrm{Perm}_{r,3}$-free but $S_1 a_1$ is not; S_2 is $\mathrm{Perm}_{r,3}$-free but $S_2 a_2$ is not, and so on. Each interval S_k must contain the last occurrence of some symbol, hence $m \leqslant n$. If this were not the case then S necessarily contains a $\mathrm{Perm}_{r,4}$ pattern, each of which is also a $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$ pattern, contradicting the $\mathrm{Perm}^{\mathrm{dbl}}_{r,3}$-freeness of S. Obtain S' by discarding $a_1,\ldots,a_m$ and contracting each S_k to a single block containing its alphabet $\Sigma(S_k)$. Since $|S_k| \leqslant \Lambda_{r,2}(\|S_k\|) < 2r\|S_k\|$, we have $|S| \leqslant 2r|S'| + n$. Being an n-block sequence, $|S'| \leqslant \Lambda^{\mathrm{dbl}}_{r,2}(n,n) < 2 \cdot 6^{r-1}(4n/3)$, so $|S| < 6^r rn$. $\square$

[19] In this context a submatrix is obtained by deleting rows and columns from A_S, and possibly flipping some 1s to 0s.