The 3/4 Conjecture for Fix-Free Codes With at Most Three Distinct Codeword Lengths

Spencer Congero, *Student Member, IEEE*, and Kenneth Zeger

***Abstract*—The $3/4$ conjecture was posed 25 years ago by Ahlswede, Balkenhol, and Khachatrian, and states that if a multiset of positive integers has Kraft sum at most $3/4$, then there exists a code that is both a prefix code and a suffix code with these integers as codeword lengths. We prove that the $3/4$ conjecture is true whenever the given multiset of positive integers contains at most three distinct values.**

***Index Terms*—Prefix codes, Kraft inequality, Huffman codes, unique decodability.**

I. Background on Fix-Free Codes

ONE of the most intriguing unsolved questions in information theory is the so-called "3/4 conjecture" for fix-free codes. The conjecture was posed 25 years ago by Ahlswede, Balkenhol, and Khachatrian, and states that if a multiset of positive integers has Kraft sum at most $3/4$, then there exists a code that is both a prefix code and a suffix code with these integers as codeword lengths. This conjecture is analogous to the well-known fact that if a multiset of positive integers has Kraft sum at most 1, then there exists a prefix code with these integers as codeword lengths.

In this paper, we prove that the $3/4$ conjecture is true whenever the given multiset of positive integers contains at most three distinct values.

Our proof technique is partially constructive and partially existential, the latter approach relying on a random coding argument, similar in spirit to that used in the classical channel coding theorem of Shannon [60].

For any two binary words u and v, let uv denote their concatenation. Let ϵ denote the empty word, such that $\epsilon u = u\epsilon = u$ for any binary word u. Denote the binary alphabet by $A = \{0, 1\}$. Let $A^0 = \{\epsilon\}$, and for each $n \geq 1$ let A^n denote the set of all n-bit binary words. Also, let $A^* = \cup_{n=0}^{\infty} A^n$ be the set of all finite-length binary words. For any sets $S, T \subseteq A^*$, denote their direct product by $ST = \{uv : u \in S, v \in T\}$. Note that $\varnothing T = T\varnothing = \varnothing$ vacuously. For any binary word $u \in A^*$, let $|u|$ denote its length.

Manuscript received 7 December 2021; revised 7 August 2022; accepted 14 October 2022. Date of publication 31 October 2022; date of current version 16 February 2023.

The authors are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: scongero@ucsd.edu; ken@zeger.us).

Communicated by O. Ordentlich, Associate Editor for Signal Processing and Source Coding.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2022.3218212.

Digital Object Identifier 10.1109/TIT.2022.3218212

A *code* is a finite subset of A^*, and a code's elements are called *codewords*. A word $u \in A^*$ is a *prefix* (respectively, *suffix*) of a word $v \in A^*$ if there exists $x \in A^*$ such that $v = ux$ (respectively, $v = xu$).

A *prefix code* (respectively, *suffix code*) is a code for which no codeword is a prefix (respectively, suffix) of any other codeword. A *fix-free code*[1] is a code that is both a prefix code and a suffix code.

If C is a code, then CA^* (respectively, A^*C) is the set of all words having a prefix (respectively, suffix) in C.

A *pattern* is a code described by a string in $\{0, 1, A\}^*$. The code consists of all possible words obtained by assigning either 0 or 1 to each occurrence of A in the pattern's string. For example, $11A0A^20$ is a pattern that contains $|11A0A^20| = 2^3 = 8$ strings, each of length 7, namely

$$\{1100000, 1100010, 1100100, 1100110,\\ 1110000, 1110010, 1110100, 1110110\}.$$

Also note that the length-one patterns 0 and 1 are the sets $\{0\}$ and $\{1\}$, respectively.

The *multiplicity* of an integer in a multiset is the number of occurrences of that integer in the multiset. If a multiset of positive integers has distinct integers $\lambda_1, \lambda_2, \ldots$ with corresponding multiplicities $\mu_1, \mu_2, \ldots$, then the *Kraft sum* of the multiset is the quantity

$$\sum_{n \geq 1} \mu_n 2^{-\lambda_n}$$

and the *Kraft sum* of a code C is the quantity

$$\mathcal{K}(C) = \sum_{u \in C} 2^{-|u|}.$$

Note that the Kraft sum of any pattern $U \in \{0, 1, A\}^p$ is $\mathcal{K}(U) = |U|/2^p$, where $|U|$ equals 2 raised to the number of As in U.

As an example, the multiset $\{2, 3, 3, 4, 4, 4, 4\}$ has distinct lengths 2, 3, and 4, with corresponding multiplicities $\mu_1 = 1$, $\mu_2 = 2$, $\mu_3 = 4$, and its Kraft sum is $1 \cdot 2^{-2} + 2 \cdot 2^{-3} + 4 \cdot 2^{-4} = 1$.

Variable length codes have been successfully used for transmission and storage of information for at least 75 years.

[1]Fix-free codes have also been called *"biprefix codes"* (e.g., [7], [49], [50], [51], [52], and [54]), *"bifix codes"* (e.g., [6], [8], [9], [10], [11], and [12]), *"affix codes"* (e.g., [21] and [53]), *"reversible variable length codes"* (e.g., [5], [27], [30], [34], [45], [61], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], and [82]), and *"never-self-synchronizing codes"* (e.g., [23]).

In particular, binary prefix codes have been the most commonly used variable length codes, and are widely embedded in many practical communication systems, such as speech, image, and video coding standards.

Prefix codes have been extensively studied and are well understood both in theory and practice. The existence of prefix codes with a given set of codeword lengths was characterized by Kraft [43] in 1949, and an optimal construction algorithm was given by Huffman [29] in 1952 that finds prefix codes with minimum average length with respect to a source distribution.

A fix-free code is a special type of prefix code, namely one that is also a suffix code. Fix-free codes have been studied for primarily four reasons: (1) theoretical and algebraic properties; (2) data compression; (3) error correction; (4) sufficient conditions for existence using Kraft-type inequalities.

Theoretical analyses of fix-free codes were originated in 1956 by Schützenberger [57] and in 1959 by Gilbert and Moore [23]. Various other algebraic properties were given from the 1960s to 1980s by Berstel and Perrin [7], Césari [16] and [17], Leonard [47], Perrin [49], [50], [51], and [52], Reutenauer [54], and Schützenberger [58] and [59], and more recently by Berstel, Berthe, DeFelice, Dolce, Leroy, Perrin, Reutenauer, and Rindone [8], [9], [10], [11], and [12], and Gillman and Rivest [24].

A special case of a fix-free code is a *palindromic* (or "symmetric") code which is defined as a prefix code, all of whose codewords are palindromes. Constructions of such codes were considered in [1], [3], [26], [55], [61], [63], [64], and [74].

In 1995, Takishima, Wada, and Murakami [61] studied fix-free codes for providing error correction capability by decoding both in the forward and reverse directions. Numerous other studies applying such codes to error correction appeared later (e.g., [5], [15], [22], [25], [27], [30], [34], [35], [45], [46], [48], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], and [82]). In fact, the practical application of fix-free codes was adopted into international standards for video compression, including ISO MPEG-4 [31] in 1998 and ITU-T H.263+ [32] in 2000.

In terms of data compression, prefix codes achieving a minimum possible average length with respect to a given source distribution are well known from Huffman's algorithm [29]. For fix-free codes, the situation is a bit more complicated. Some studies of this include [1], [33], [36], [37], [39], [41], [42], [55], [74], [77], [80], and [81].

In addition to the practical use of fix-free codes for error correction of variable length lossless codes, the foundational theory of fix-free codes has been a topic of great interest.

In order for any variable length code to be useful, it is generally required that it be uniquely decodable (UD), which means that there is only one way to correctly parse a concatenation of variable length codewords. Prefix codes are always UD, and it is known that for every UD code, there exists a prefix code with the same codeword lengths [18]. So there is no loss of generality in restricting one's attention from general UD variable length codes to prefix codes.

On the other hand, for the purpose of lossless data compression, one would like the average codeword length to be as short as possible, in order to reduce transmission and storage costs. This assumes each codeword is assigned to represent a particular outcome of a discrete source random variable. The desire to have codes be UD and short on average are opposing needs. That is, if a code is too short, it cannot also be UD.

The Kraft inequality makes this idea quantitatively precise. Specifically, the Kraft inequality gives an upper bound of 1 on the Kraft sum of a multiset of positive integers corresponding to the codeword lengths of a prefix code. In other words, as long as this upper bound is not violated, a prefix code exists having those positive integers as its codeword lengths. In fact, the converse to the Kraft theorem is also true, namely that the Kraft sum of the codeword lengths of any prefix code can be at most 1.

For fix-free codes, a similar trade-off exists between having short codeword lengths and being both a prefix and a suffix code. An analogous question to the prefix code case asks for the lowest possible upper bound on the Kraft sum of a multiset of positive integers that would ensure the existence of a fix-free code having those positive integers as its codeword lengths. No improved converse can exist however, since fix-free codes can indeed have Kraft sum equal to 1, such as a code consisting of all codewords of a given length.

In 1996, Ahlswede, Balkenhol, and Khachatrian [4] showed the weaker result that if the Kraft sum is at most $1/2$, then a fix-free code is guaranteed to exist with the corresponding codeword lengths. They also showed the existence of a fix-free code if the Kraft sum of the multiset of codeword lengths is at most $3/4$ and each integer in the multiset is at least twice any smaller integer in the multiset. More generally, they showed that any upper bound on the Kraft sum that ensures the existence of a fix-free code cannot be larger than $3/4$. Perhaps most interestingly, the authors of [4] conjectured that $3/4$ itself is in fact such an upper bound on the Kraft sum. This is now commonly referred to as the "$3/4$ conjecture", and is stated next.

Conjecture I.1 (Ahlswede, Balkenhol and Khachatrian [4]): If a multiset of positive integers has Kraft sum at most $3/4$, then there exists a fix-free code whose codeword lengths are the elements of the multiset.

Since the $3/4$ conjecture was made, numerous attempts to prove it have failed. However, many interesting special cases of the conjecture have been proven, which we review next.

In 1999, Harada and Kobayashi [28] proved that Conjecture I.1 holds if the multiset contains at most two distinct positive integers. Initially, they attempted to use a randomized algorithm in their proof, but demonstrated that it is not guaranteed to find the desired fix-free code. To achieve their result, they used a deterministic algorithm. They were unable to extend their methods beyond multisets containing at most two distinct positive integers and, in fact, stated the following:

> *"However, finding a fix-free code for $l_1, \ldots, l_n$ which consists of three or more different lengths seems not to be always easy."*

In 2012, Savari, Yazdi, Abedini, and Khatri [55, p. 5121] proved, among other things, one special case of Conjecture I.1 where the given multiset has three distinct values. In particular, their result is limited to the case where the smallest such value is 2 and appears exactly once, and the remaining two values have further specific restrictions. The authors also stated the following that acknowledges the Harada-Kobayashi result for multisets with two distinct values and confirms the difficulty of proving Conjecture I.1 for multisets containing three distinct values:

> *"The progress on the $3/4$ conjecture has been slow even over binary code alphabets. One of the early results [by Harada-Kobayashi] is that the $3/4$ conjecture holds for length sequences $(l_1, \ldots, l_n)$ for which $l_i \in \{\lambda_1, \lambda_2\}$ for all i. The case where $l_i \in \{\lambda_1, \lambda_2, \lambda_3\}$ is only partly understood."*

It is precisely the proof of Conjecture I.1 for at most three distinct integers that we achieve in the present paper (in our Theorem II.1).

In 2001, Ye and Yeung [75] proved that Conjecture I.1 holds when the multiset values do not exceed 7. They also proved the weaker result that a fix-free code exists when the multiset contains the integer 1 and the Kraft sum is at most $5/8$.

In 2001, also Yekhanin [76] gave a proof sketch that Conjecture I.1 holds in two different cases: (1) when the multiset values do not exceed 8; or (2) when the Kraft sum of the submultiset of is and $(i+1)$s is at least $1/2$, where i is the smallest integer in the multiset. A special case of this second result is stated as the following theorem, which we use as one component of our main result, Theorem II.1. Theorem I.2 is proved in more detail in Yekhanin's unpublished notes in [78].

Theorem I.2: Conjecture I.1 holds when the Kraft sum of the multiset of smallest length words is at least $1/2$.

In 2004, Yekhanin [77] also proved Conjecture I.1 holds when the Kraft sum is at most $5/8$.

In 2005, Kukorelly and Zeger [44] proved that Conjecture I.1 holds in two different cases: (1) when the minimum integer i in the multiset is at least 2, and no integer in the multiset, except possibly the largest one, occurs more than 2^{i-2} times; or (2) when every integer in the multiset, except possibly the largest one, occurs at most twice.

In 2007, Schnettler [56] (see also [19], [20], and [40]) gave a survey of sufficient conditions for the existence of fix-free codes and generalized to nonbinary alphabets the result described above in [44]. He also expanded the proof sketch given in [76] to a more general version of Theorem I.2, and proved several specialized cases of Conjecture I.1.

In 2008, Khosravifard and Gulliver [38] further studied and improved the algorithm used by Harada and Kobayashi [28] to establish Conjecture I.1 for two-level integer multisets. They experimentally showed that their algorithm tends to almost always find fix-free codes, when they exist, for multisets containing at most 30 integers, with two or more distinct values.

In 2013, Aghajan and Khosravifard [2] calculated the fraction of cases covered by Yekhanin's result (2) in [76].

In 2015, Bodewig [13], [14] proved several special cases of Conjecture I.1 for infinite multisets.

Today, there still remains an infinite number of unsolved cases of Conjecture I.1.

II. Summary of the Main Result

Our main result covers an infinite number of new cases not previously known in the literature, and is summarized in Theorem II.1.

Theorem II.1: Conjecture I.1 is true whenever the multiset contains at most three distinct integers.

Proof: By Lemma V.1, it suffices to prove the result when the Kraft sum is exactly $3/4$. If the multiset contains only one distinct integer λ_1, then any subset of A^{λ_1} of size $\mu_1 = 3 \cdot 2^{\lambda_1 - 2}$ will give the desired fix-free code. If the multiset contains exactly two distinct integers, then the result is known by [28] (see also our Theorem III.1).

Suppose the multiset contains exactly three distinct integers, which, in increasing order, are λ_1, λ_2, λ_3, with nonzero multiplicities μ_1, μ_2, μ_3, respectively, and such that

$$\mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} + \mu_3 2^{-\lambda_3} = 3/4.$$

Theorem I.2 implies that Conjecture I.1 holds when $\mu_1 2^{-\lambda_1} \geq 1/2$, so it suffices to assume $\mu_1 2^{-\lambda_1} \leq 1/2$, in which case the proof follows from our following three results:

- Theorem VI.2, i.e., when $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\mu_2 2^{-\lambda_2} \leq \frac{1}{4}$
- Theorem VII.2, i.e., when $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\frac{1}{4} \leq \mu_2 2^{-\lambda_2} \leq \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right)$
- Theorem VIII.1, i.e., when $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) \leq \mu_2 2^{-\lambda_2}$

These theorems are stated and proven in Sections VI, VII, VIII, respectively. If $\lambda_1 = 1$, then $\mu_1 = 1$, so Theorem I.2 applies. Thus, for each of Theorems VI.2, VII.2, and VIII.1, it suffices to assume $\lambda_1 \geq 2$. Throughout the proofs of these three theorems, we will use the following transformed quantities:

$$\begin{aligned} n &= \lambda_1 - 1 \\ l &= \lambda_2 - \lambda_1 + 1 \\ k &= \lambda_3 - \lambda_1 + 1. \end{aligned} \tag{1}$$

■

The main idea used in proving Theorems VI.2, VII.2, and VIII.1 is to build sets of codewords of the three desired lengths so that none of the shorter words is a prefix or suffix of any longer word.

The codewords of the shortest length λ_1 will be elements of the set $U_1 A^n$, where $U_1 = 0$. The set of codewords of the middle length λ_2 will be a union of at most three sets of the form $U_2 A^{l-2} U_3 A^n$, where U_2 and U_3 are two fixed bits. Since $\lambda_1 = n+1$, the conditions $U_1 \neq U_2$ and $U_2 = U_3$ ensure any word from $U_2 A^{l-2} U_3 A^n$ will not have a length-λ_1 prefix or suffix from $U_1 A^n$. Additionally, even if $U_1 = U_2$ or $U_1 = U_3$, we will still be able to choose codewords of length λ_2 as long as these words avoid having prefixes or suffixes among the codewords from $U_1 A^n$.

Once the codewords of lengths λ_1 and λ_2 are constructed with the correct multiplicities μ_1 and μ_2, and with no offending prefixes or suffixes, we then carefully construct enough codewords of length λ_3 to make the total Kraft sum equal $3/4$, while avoiding prefixes and suffixes from codewords of lengths λ_1 and λ_2.

During this construction, the locations in words of length λ_3 of the fixed bits U_1, U_2, and U_3 (which are fixed in words of lengths λ_1 and λ_2) play an important role in our ability to avoid prefixes and suffixes of words of lengths λ_1 and λ_2. There are three possible "overlap cases" that are separately considered, depending on how much overlap there is between the prefix and suffix of length λ_2 in a codeword of length λ_3. The three cases are illustrated in Figure 1, and correspond to whether the value $\lambda_2 - \lambda_1$ is less than, equal to, or greater than, $\lambda_3 - \lambda_2$, An equivalent condition, using the terminology from (1), is whether the value $2l - k$ is less than, equal to, or greater than 1.

In Overlap Case 1, the fixed bits U_2 and U_3 of the length-λ_2 prefixes and suffixes do not coincide, and also the length-λ_2 prefixes do not overlap the length-λ_2 suffixes in their first l positions. In this case, length-λ_3 codewords are drawn from sets of the form

$$Z_1 A^{l-2} Z_2 A^{k-2l} Z_3 A^{l-2} Z_4 A^n$$

for some subset of the 16 possible assignments of (Z_1, Z_2, Z_3, Z_4). In Overlap Case 2, the fixed bit U_3 in length-λ_2 prefixes of length-λ_3 codewords coincides with the fixed bit U_2 in length-λ_2 suffixes of such length-λ_3 codewords.

In this case, length-λ_3 codewords are drawn from sets of the form

$$Z_1 A^{l-2} Z_2 A^{l-2} Z_3 A^n,$$

for some subset of the 8 possible assignments of (Z_1, Z_2, Z_3). In Overlap Case 3, the fixed bits U_2 and U_3 of the length-λ_2 prefixes and suffixes do not coincide, but the length-λ_2 prefixes do overlap the length-λ_2 suffixes in their first l positions. It turns out that this latter property is a source of complication throughout the construction. In this case, length-λ_3 codewords are drawn from sets of the form $Z_1 A^{k-l-1} Z_3 A^{2l-k-2} Z_2 A^{k-l-1} Z_4 A^n$, for some subset of the 16 possible assignments of (Z_1, Z_2, Z_3, Z_4).

In all three cases, the codewords of length λ_3 are randomly selected from carefully designed patterns to avoid words of lengths λ_1 and λ_2 as prefixes and suffixes. The proof of our main result demonstrates that, on average, the random choices of codewords of lengths λ_1 and λ_2 leave enough remaining potential codewords of length λ_3 to satisfy the Kraft sum being $3/4$ without violating the prefix or suffix conditions. This bound on the average Kraft sum implies that there exists at least one particular choice of words of the desired lengths that forms a fix-free code and satisfies the requirements.

III. The 3/4 Conjecture With Two Distinct Lengths

In order to illustrate aspects of our random coding technique in a relatively simple example, we next prove Conjecture I.1 when the multiset of positive integers is restricted to having only two distinct values. This result was originally given by Harada and Kobayashi [28] using a different, and considerably longer, proof.

Theorem III.1: Suppose a multiset of positive integers consists of μ_1 copies of λ_1 and μ_2 copies of λ_2, such that $\lambda_1 < \lambda_2$ and $\mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} \leq 3/4$. Then there exists a fix-free code with μ_1 codewords of length λ_1 and μ_2 codewords of length λ_2.

Proof: Let F be a randomly chosen set of μ_1 distinct words of length λ_1. Each word of length $\lambda_2 - \lambda_1$ is the prefix of a unique word of length λ_2 whose length-λ_1 prefix equals its length-λ_1 suffix. So the number of such words is $2^{\lambda_2-\lambda_1}$, and their Kraft sum is $2^{\lambda_2-\lambda_1} \cdot 2^{-\lambda_2} = 2^{-\lambda_1}$. The probability that any such word does not have its common length-λ_1 prefix/suffix in F is $\frac{2^{\lambda_1}-\mu_1}{2^{\lambda_1}}$. On the other hand, any word of length λ_2 whose length-λ_1 prefix and suffix differ does not have a prefix or suffix in F with probability

$$\frac{2^{\lambda_1}-\mu_1}{2^{\lambda_1}} \cdot \frac{2^{\lambda_1}-\mu_1-1}{2^{\lambda_1}-1}$$

(using Lemma V.6), and the Kraft sum of the set of such words is $(2^{\lambda_2} - 2^{\lambda_2-\lambda_1}) \cdot 2^{-\lambda_2} = 1 - 2^{-\lambda_1}$. Therefore, the expected Kraft sum of the set of length-λ_2 words that have neither a prefix nor suffix in F is

$$\begin{aligned} & 2^{-\lambda_1} \cdot \frac{2^{\lambda_1}-\mu_1}{2^{\lambda_1}} \\ & \quad + (1 - 2^{-\lambda_1}) \cdot \frac{(2^{\lambda_1}-\mu_1)(2^{\lambda_1}-\mu_1-1)}{2^{\lambda_1}(2^{\lambda_1}-1)} \\ & = \frac{3}{4} - \mu_1 2^{-\lambda_1} + \left(\frac{1}{2} - \mu_1 2^{-\lambda_1}\right)^2 \\ & \geq \frac{3}{4} - \mu_1 2^{-\lambda_1} \\ & \geq \mu_2 2^{-\lambda_2}. \end{aligned}$$

Thus, there exists at least one particular choice of F such that there are at least μ_2 words of length λ_2 that have neither a prefix nor suffix in F, i.e., there are then enough available words of length λ_2 to create the claimed fix-free code. ■

IV. Overview of the Proof of the 3/4 Conjecture With Three Distinct Lengths

We give here a preview and high-level description of the proof of the $3/4$ conjecture with three distinct lengths (i.e., the proof of Theorem II.1). In Sections VI–VIII, detailed proofs are given, and some useful lemmas are given in Section V and proven in the appendix.

The random coding method illustrated in the proof of Theorem III.1 for two distinct lengths plays an important role in the case of three distinct lengths, but significant complications arise when trying to avoid prefixes and suffixes in the words of longest length. To avoid such prefixes and suffixes in our constructions, we assign fixed values to certain bit locations in the chosen words of all three lengths, which in turn does make the analysis based on random coding more difficult. Also, the method in the proof of Theorem III.1 of counting each length-λ_2 word whose length-λ_1 prefix is also a suffix does not work when there are bits with fixed values, as in the proofs with three distinct lengths, so we instead develop a more widely applicable result that is proven in our Lemma V.3. As a result, the proofs provided in subsequent sections are substantially longer and more complex than that of Theorem III.1.

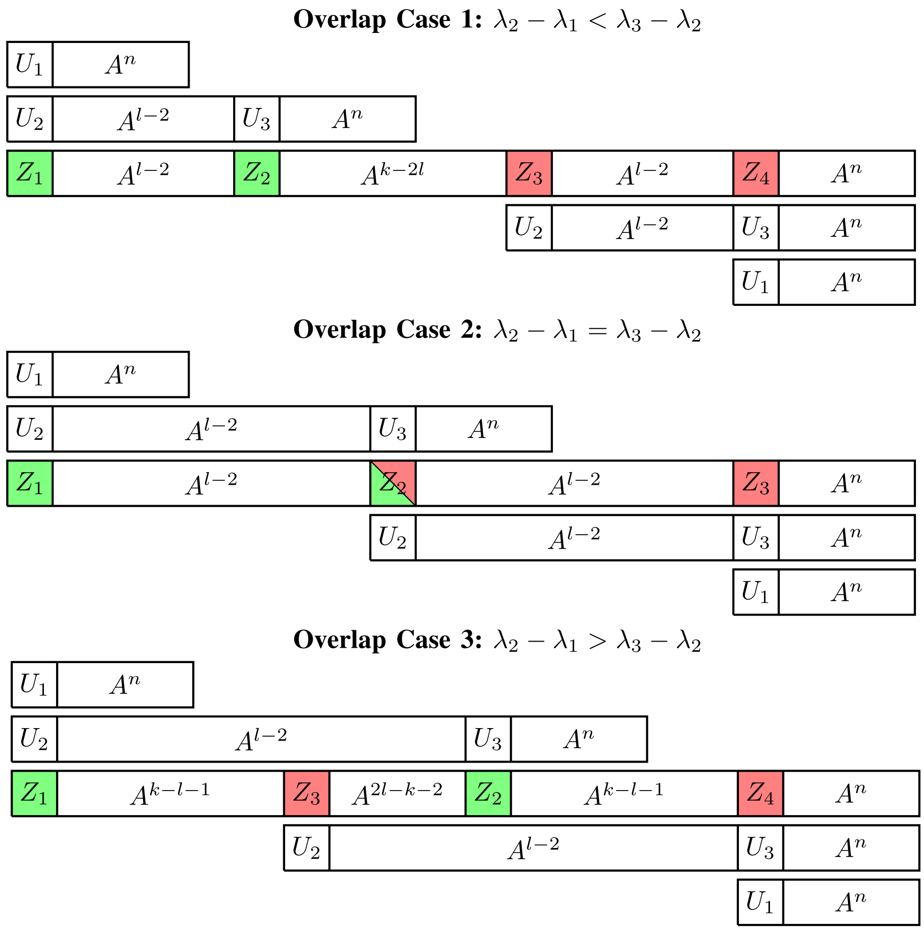

Fig. 1. Three cases of code word overlap. The three word lengths illustrated are λ_1, λ_2, and λ_3. The bit positions U_1, U_2, U_3 correspond to certain fixed bits in patterns of length λ_1 and λ_2, and the bit positions Z_1, Z_2, Z_3, Z_4 represent fixed bits in patterns of length λ_3. These fixed bit positions are used to avoid prefixes and suffixes in order to create a fix-free code.

Explicitly constructing the needed numbers of strings of lengths λ_1, λ_2, and λ_3 appears to be a difficult task, so we chose an alternative approach that randomly chooses such strings according to certain rules that maintain the prefix/suffix conditions. The construction process chooses the correct number of strings of lengths λ_1 and λ_2 and then we show that on average there remains enough strings of length λ_3 to complete the process.

The proof of Theorem II.1 is broken into three main cases, depending on the values of the Kraft sum components $\mu_1 2^{-\lambda_1}$ and $\mu_2 2^{-\lambda_2}$. The three cases are:

(1) $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\mu_2 2^{-\lambda_2} \leq \frac{1}{4}$
(2) $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\frac{1}{4} \leq \mu_2 2^{-\lambda_2} \leq \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right)$
(3) $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) \leq \mu_2 2^{-\lambda_2}$.

The third main case is broken into the following four subcases:

(a) $\lambda_2 \geq 2\lambda_1$ (i.e., $n \leq l-2$)
(b) $\lambda_2 < 2\lambda_1$ (i.e., $n > l-2$) and $\frac{1}{4} \leq \mu_1 2^{-\lambda_1} \leq \frac{1}{2}$
(c) $\lambda_2 < 2\lambda_1$ (i.e., $n > l-2$) and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\frac{1}{4} \leq \mu_2 2^{-\lambda_2} \leq \frac{1}{2}$
(d) $\lambda_2 < 2\lambda_1$ (i.e., $n > l-2$) and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\frac{1}{2} < \mu_2 2^{-\lambda_2}$.

Each of the main cases (1) and (2) and the subcases (3a)–(3d) are further divided into the three overlap cases illustrated in Figure 1, namely:

- $\lambda_2 - \lambda_1 < \lambda_3 - \lambda_2$ (i.e., $2l - k < 1$)
- $\lambda_2 - \lambda_1 = \lambda_3 - \lambda_2$ (i.e., $2l - k = 1$)
- $\lambda_2 - \lambda_1 > \lambda_3 - \lambda_2$ (i.e., $2l - k > 1$).

Within each overlap case of each main case or subcase, specific definitions are given of the three sets F_1, F_2, and F_3. These are the sets containing codewords of lengths λ_1, λ_2, and λ_3, respectively. Our construction chooses these three sets using various randomizations, and we show that in each case, on average, there are enough codewords to correctly populate the sets without violating the prefix or suffix conditions. Once this step is accomplished, we then conclude that there must be at least one (non-random) code with the correct sizes of F_1, F_2, and F_3, and without violating the prefix or suffix conditions.

- Constructing F_1:

  In all cases and subcases we define $F_1 = 0A^n - C$, where C is a set of size $2^n - \mu_1$ chosen randomly from certain subsets of $0A^n$. In other words, we construct F_1 by starting with all binary strings of length $n+1$ that start with 0, and then we delete in a random way enough of those strings to leave exactly μ_1 remaining.

The motivation behind this definition of F_1 is that when we construct larger codewords of lengths λ_2 and λ_3, they can avoid having length-λ_1 codewords as prefixes by having a 1 in their leftmost position or having a word in C as a prefix, and they can avoid having length-λ_1 codewords as suffixes by having a 1 in position $n+1$ from the right or having a word in C as a suffix.

In the main cases (1) and (2) and the subcase (3a), we choose C uniformly at random from among the 2^{n-1} length-λ_1 strings of $0A^n$. For these cases, the construction of F_1 is equivalent to simply choosing μ_1 elements at random without replacement from $0A^n$.

In subcase (3b), we choose C uniformly at random from among the 2^{n-1} length-λ_1 strings of $0A^{l-2}1\ A^{n-l+1}$. In other words, in this case F_1 is constructed by randomly deleting enough strings from $0A^n$ containing a 1 in the lth position to leave exactly μ_1 strings remaining.

In subcases (3c) and (3d), since $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ the value of μ_1 is smaller than in case (3b), so the random set C must be made larger than in (3b). So we choose C to have all 2^{n-1} strings in $0A^{l-2}1\ A^{n-l+1}$ together with $2^{n-1} - \mu_1$ strings chosen uniformly at random from $0A^{l-2}0\ A^{n-l+1}$. In other words, in these cases F_1 is constructed by deleting all strings from $0A^n$ that contain a 1 in the lth position and also randomly deleting enough strings from $0A^n$ that contain a 0 in the lth position to leave exactly μ_1 strings remaining.

- Constructing F_2:

The construction of F_2 requires that F_2 has μ_2 strings, each of length λ_2, and that none of these strings contains a prefix or suffix in F_1.

Notice that each word in $1A^{l-2}1\ A^n$ avoids both prefixes and suffixes from F_1. There are $2^{n+l-2} = \frac{1}{4}2^{\lambda_2}$ such words available for F_2, and each is of length $n+l = \lambda_2$. For main case (1), this number of codewords is sufficient since $\mu_2 \leq \frac{1}{4}2^{\lambda_2}$, but for main cases (2) and (3) more codewords are needed of length λ_2 since $\mu_2 > \frac{1}{4}2^{\lambda_2}$ in those cases. In these two cases, one way to increase the number of codewords of length λ_2 is to include in F_2 some words from $0A^{l-2}1\ A^n$ or $1A^{l-2}0\ A^n$, and then require such words to have a prefix or suffix, respectively, from C, in order to avoid prefixes or suffixes from F_1.

When constructing F_2, we make use of a new set D which is chosen uniformly at random from one of the four sets:

$$1A^{l-2}1A^n$$
$$0A^{l-2}1A^n$$
$$1A^{l-2}C$$
$$CA^{l-1}.$$

In all cases except (3d), the words in set D are avoided when constructing F_2, which allows words of length λ_3 to have prefixes or suffixes from D in the construction of F_3. In contrast, in case (3d) we add words from D when constructing F_2, and then avoid such words in constructing F_3.

Table I shows for each main case, subcase, and overlap case which of the four sets D is chosen from, how many words D contains, and the exact definition of F_2. The precise usage of these quantities will become apparent in the detailed proofs given in Sections VI–VIII. The involvement of D for constructing F_2 in each case allows sufficient codewords of length λ_2 while avoiding prefixes and suffixes from F_1.

- Constructing F_3:

Our general strategy for constructing the set F_3 is to form a union of subsets of A^{k+n}, where each subset obeys certain constraints (according to which overlap case is being considered) that prevent prefixes or suffixes from F_1 or F_2. Each subset in such a union is an intersection of two specially constructed sets $Y_{p,q}$ and $W_{r,s}$ over various binary values of p, q, r, s, specified by an index set $\mathcal{I}$. The intersection of $Y_{p,q}$ and $W_{r,s}$ produces a pattern that falls into one of three specific forms, as seen in the third column of the Table II. The patterns are designed based on the locations of the bits Z_1, Z_2, Z_3, Z_4 in Figure 1. By controlling the values of these four bits we prevent the strings in F_3 from having prefixes and suffixes in F_1 and F_2. The sets $Y_{p,q}$ regulate prefixes and the sets $W_{r,s}$ regulate suffixes.

Table II summarizes the construction of F_3 and the pattern used for each overlap case.

These constructions of the random set F_3 are used for all cases, except for main case (3d), where a slightly different construction is used.

The proofs in Sections VI–VIII calculate the average size of F_3 and show that it is at least μ_3. This ensures that there is at least one (deterministic) instance of the random sets C and D that guarantees a (deterministic) instance of the set F_3 with at least μ_3 elements, and this instance of F_3 can be pruned back to have exactly μ_3 elements.

V. Lemmas About Kraft Sums

This section gives many technical lemmas used to prove the main theorems in the sections to follow. Most of the proofs of the lemmas in this section can be found in the appendix.

For any set $S \subseteq A^*$, the *indicator function* $1_S : A^* \to \{0,1\}$ is defined by

$$1_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{else.} \end{cases}$$

Lemma V.1: If a multiset of positive integers has Kraft sum less than $3/4$, then the multiplicity of its largest value can be increased to make the Kraft sum equal to $3/4$.

Some basic facts about Kraft sums will be used throughout this paper. As some examples: (i) if S and T are disjoint codes, then $\mathcal{K}(S \cup T) = \mathcal{K}(S) + \mathcal{K}(T)$; (ii) if S and T are codes and at least one of them is fixed length, then $\mathcal{K}(ST) = \mathcal{K}(S)\mathcal{K}(T)$; (iii) $\mathcal{K}(A^n) = 1$ for all $n \geq 1$; and (iv) $\mathcal{K}(0) = \mathcal{K}(1) = 1/2$. Two consequences of these facts are given in the following lemma.

TABLE I

SETS F_2 AND D FOR ALL OVERLAP CASES AND SUBCASES USED IN THE PROOF OF THEOREM III.1

Case : Overlap	F_2	$\|D\|$	D is from
(1)	$1A^{l-2}1A^n - D$	$2^{n+l-2} - \mu_2$	$1A^{l-2}1A^n$
(2): 1,3	$(1A^{l-2}1A^n - D) \cup 1A^{l-2}C$	$\frac{1}{2}(1-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	$1A^{l-2}1A^n$
(2): 2	$1A^{l-2}1A^n \cup (1A^{l-2}C - D)$	$\frac{1}{2}(1-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	$1A^{l-2}C$
(3a): 1	$(1A^{l-2}1A^n - D) \cup 1A^{l-2}C \cup CA^{l-2-n}1A^n$	$(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	$1A^{l-2}1A^n$
(3a): 2,3	$1A^{l-2}1A^n \cup (1A^{l-2}C - D) \cup CA^{l-2-n}1A^n$	$(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	$1A^{l-2}C$
(3b): 1,3	$(1A^{l-2}1A^n - D) \cup CA^{l-1}$	$(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	$1A^{l-2}1A^n$
(3b): 2	$1A^{l-2}1A^n \cup (CA^{l-1} - D)$	$(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$	CA^{l-1}
(3c): 1,3	$(1A^{l-2}1A^n - D) \cup 0A^{l-2}1A^n$	$2^{\lambda_2-1} - \mu_2$	$1A^{l-2}1A^n$
(3c): 2	$1A^{l-2}1A^n \cup (0A^{l-2}1A^n - D)$	$2^{\lambda_2-1} - \mu_2$	$0A^{l-2}1A^n$
(3d)	$1A^{l-2}1A^n \cup 0A^{l-2}1A^n \cup D$	$\mu_2 - 2^{\lambda_2-1}$	$1A^{l-2}C$

TABLE II

SET F_3 AND CORRESPONDING PATTERN FOR ALL OVERLAP CASES USED IN THE PROOF OF THEOREM III.1

Overlap Case	F_3	Pattern
(1)	$\bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})$	$Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n$
(2)	$\bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})$	$Z_1A^{l-2}Z_2A^{l-2}Z_3A^n$
(3)	$\bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_3} \cap W_{Z_2,Z_4})$	$Z_1A^{k-l-1}Z_2A^{2l-k-2}Z_3A^{k-l-1}Z_4A^n$

Lemma V.2:

(i) If $p_1, \ldots, p_n$ are nonnegative integers and $u_1, \ldots, u_n \in A^*$, then

$$\mathcal{K}\left(u_1 A^{p_1} u_2 A^{p_2} \ldots u_n A^{p_n}\right) = 2^{-(|u_1|+\cdots+|u_n|)}.$$

(ii) If S is a code and T is a fixed-length random code, then $E[\mathcal{K}(ST)] = \mathcal{K}(S)\, E[\mathcal{K}(T)]$.

Let $m \geq l \geq 1$ be integers, and let $U \in \{0,1,A\}^m$. Define $R_l(U) \subseteq A^m$ to be the set of words $w \in U$ such that the l-bit prefix of w equals the l-bit suffix of w.

The following lemma is used in many of the proofs of our other lemmas.

Lemma V.3: Let $U = U_1 U_2 \cdots U_m \in \{0,1,A\}^m$ and let $l \leq m$ be a positive integer. Then the number of words in U whose length-l prefix and suffix are the same is

$$|R_l(U)| = \prod_{p=1}^{m-l} \left| \bigcap_{\substack{1 \leq i \leq m \\ i \equiv p \bmod (m-l)}} U_i \right|.$$

Three examples illustrating the usage of Lemma V.3 are given next.

- Let $U = 0A^2 0A^3 1$ and $l = 5$. So $m - l = 3$. Then

$$\begin{aligned} U_1 \cap U_4 \cap U_7 &= 0 \cap 0 \cap A = 0 \\ U_2 \cap U_5 \cap U_8 &= A \cap A \cap 1 = 1 \\ U_3 \cap U_6 &= A \cap A = A. \end{aligned}$$

  So $|R_l(U)| = |0| \cdot |1| \cdot |A| = 1 \cdot 1 \cdot 2 = 2$. The two words in $R_l(U)$ are 01001001 and 01101101.
- Let $U = 0A^2 0A^2 1\ A$ and $l = 5$. So $m - l = 3$. Then

$$\begin{aligned} U_1 \cap U_4 \cap U_7 &= 0 \cap 0 \cap 1 = \varnothing \\ U_2 \cap U_5 \cap U_8 &= A \cap A \cap A = A \\ U_3 \cap U_6 &= A \cap A = A. \end{aligned}$$

  So $|R_l(U)| = |\varnothing| \cdot |A| \cdot |A| = 0 \cdot 2 \cdot 2 = 0$.
- Let $U = 1A^2 1\ A\ 1\ A^2 1$ and $l = 6$. So $m - l = 3$. Then

$$\begin{aligned} U_1 \cap U_4 \cap U_7 &= 1 \cap 1 \cap A = 1 \\ U_2 \cap U_5 \cap U_8 &= A \cap A \cap A = A \\ U_3 \cap U_6 \cap U_9 &= A \cap 1 \cap 1 = 1. \end{aligned}$$

  So $|R_l(U)| = |1| \cdot |A| \cdot |1| = 1 \cdot 2 \cdot 1 = 2$. The two words in $R_l(U)$ are 101101101 and 111111111.

A *fixed point* in a pattern $X \in \{0,1,A\}^*$ is a position in the pattern's string whose value is not equal to A.

We will say that a randomly generated set of words of a given length is *of a fixed size* if the set is chosen according to some probability distribution on all sets of the same cardinality that contain words of the given length.

Lemma V.4: Let m be a positive integer. For each $i = 1, 2$ let $X_i \in \{0, 1, A\}^{m_i}$, with $m_i \leq m$, and let Y_i be a set of a fixed size drawn uniformly at random and without replacement from X_i, where the words of Y_1 and Y_2 are drawn independently of each other. Let $W_1 = A^a Y_1 A^b \cap U_1$ and $W_2 = A^c Y_2 A^d \cap U_2$ for some patterns $U_1 \subseteq A^a X_1 A^b$ and $U_2 \subseteq A^c X_2 A^d$, where $a+b = m-m_1$ and $c+d = m-m_2$. Let p denote the number of positions where U_1 and U_2 both have a fixed point, and assume that the values of U_1 and U_2 agree at each such position. Then

$$E[\mathcal{K}(W_1 \cap W_2)] = 2^p \cdot \prod_{i=1}^{2} \mathcal{K}(U_i) \frac{\mathcal{K}(Y_i)}{\mathcal{K}(X_i)}.$$

Note that in the above lemma, the cardinality of each random set Y_i is fixed and its elements are all of length m_i, so the Kraft sum of Y_i is deterministic.

Corollary V.5: Let Y be a set of a fixed size chosen uniformly at random and without replacement from a pattern $X \in \{0, 1, A\}^{n+1}$. Let $U \in \{0, 1, A\}^{n+k}$. If $U \subseteq A^a X A^b$, then

$$\mathcal{K}\left(A^a Y A^b \cap U\right) = \mathcal{K}(U) \cdot \frac{\mathcal{K}(Y)}{\mathcal{K}(X)}.$$

Lemma V.6: Let X be a set of size at least 2 and let C be a set of a fixed size chosen uniformly at random from X. For any particular element of X, the probability that the element lies in C is $|C|/|X|$. For any two particular distinct elements of X, the probability that both lie in C is $\frac{|C|(|C|-1)}{|X|(|X|-1)}$.

Lemma V.7: Let $n \geq 1$ and $p \geq 0$ be integers, let $b \in A$, and let C be a set of a fixed size chosen uniformly at random from bA^n. Then for any $U \in \{0, 1, A\}^p$,

$$E[\mathcal{K}\left(CA^{p+1} \cap bUbA^n \cap A^{p+1}C\right)] = \mathcal{K}(U)\, \mathcal{K}(C)^2.$$

The next lemma calculates the expected Kraft sum of the set all $(k+n)$-bit words that have both a prefix and suffix in a randomly chosen set of words of the form $1A^{l-2}1A^n$, where $2 \leq l < k$.

Lemma V.8: Let $n, l, k \geq 0$ be integers, with $2 \leq l < k$, and let D be a subset of a fixed size chosen uniformly at random from $1A^{l-2}1\ A^n$. Then

$$E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] = \begin{cases} \mathcal{K}(D)^2 & \text{if } 2l-k < 1 \\ 2\mathcal{K}(D)^2 & \text{if } 2l-k = 1 \\ \mathcal{K}(D)^2 & \text{if } 2l-k > 1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(D)^2 + \frac{\mathcal{K}(D)(\frac{1}{4}-\mathcal{K}(D))}{2^{n+l-2}-1} & \text{if } 2l-k > 1 \text{ and } (k-l) \mid (2l-k-1). \end{cases}$$

Corollary V.9: Let $n, l, k \geq 1$ be integers, with $2 \leq l < k$ and $n > l-2$. Let C_0 be a subset of a fixed size chosen uniformly at random from $0A^{l-2}0\ A^{n-(l-1)}$. Then

$$E[\mathcal{K}\left(C_0 A^{k-l} \cap A^{k-l} C_0\right)] = \begin{cases} \mathcal{K}(C_0)^2 & \text{if } 2l-k < 1 \\ 2\mathcal{K}(C_0)^2 & \text{if } 2l-k = 1 \\ \mathcal{K}(C_0)^2 & \text{if } 2l-k > 1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(C_0)^2 + \frac{\mathcal{K}(C_0)(\frac{1}{4}-\mathcal{K}(C_0))}{2^{n-1}-1} & \text{if } 2l-k > 1 \text{ and } (k-l) \mid (2l-k-1). \end{cases}$$

Proof: This corollary follows from Lemma V.8 by changing 1s to 0s. ■

Lemma V.10: Let $C \subseteq A^{n+1}$ be a random set of a fixed size. Let $g(C) \subseteq A^{n+k}$ be a set that is some function of C. If D is a set of a fixed size chosen uniformly at random from $1A^{l-2}C$, then

$$E[\mathcal{K}\left(DA^{k-l} \cap g(C)\right)] = \frac{\mathcal{K}(D)}{\mathcal{K}(C)/2} \cdot E[\mathcal{K}\left(1A^{l-2}CA^{k-l} \cap g(C)\right)]$$

$$E[\mathcal{K}\left(g(C) \cap A^{k-l}D\right)] = \frac{\mathcal{K}(D)}{\mathcal{K}(C)/2} \cdot E[\mathcal{K}\left(g(C) \cap A^{k-l}1A^{l-2}C\right)],$$

and if D is a set of a fixed size chosen uniformly at random from CA^{l-1}, then

$$E[\mathcal{K}\left(DA^{k-l} \cap g(C)\right)] = \frac{\mathcal{K}(D)}{\mathcal{K}(C)} \cdot E[\mathcal{K}\left(CA^{k-1} \cap g(C)\right)],$$

where $\mathcal{K}(D)/\mathcal{K}(C) = 0$ whenever $\mathcal{K}(C) = \mathcal{K}(D) = 0$.

In Theorem VIII.1(c) and Theorem VIII.1(d) in Section VIII, the set C is not chosen uniformly at random from a fixed pattern, but instead $C = C_1 \cup C_0$, where $C_1 = 0\ A^{l-2}1A^{n-(l-1)}$ and C_0 is chosen uniformly at random from $0A^{l-2}0A^{n-(l-1)}$. The following lemma calculates $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-1}C\right)]$ in this situation.

Lemma V.11: Let $n, l, k \geq 1$ be integers, with $2 \leq l < k$ and $n > l-2$. Let $C = C_1 \cup C_0$ be a set of a fixed size where $C_1 = 0A^{l-2}1A^{n-(l-1)}$ and C_0 is chosen uniformly at random from $0A^{l-2}0\ A^{n-(l-1)}$. Let D be a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. Then

$$E[\mathcal{K}\left(DA^{k-l} \cap A^{k-1}C\right)] = \begin{cases} \mathcal{K}(C)\,\mathcal{K}(D) & \text{if } 2l-k < 1 \\ 2(\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(D) & \text{if } 2l-k = 1 \\ \mathcal{K}(C)\,\mathcal{K}(D) & \text{if } 2l-k > 1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(C)\,\mathcal{K}(D) + \frac{\mathcal{K}(D)(\mathcal{K}(C)-\frac{1}{4})(\frac{1}{2}-\mathcal{K}(C))}{\mathcal{K}(C)(2^{n-1}-1)} & \text{if } 2l-k > 1 \text{ and } (k-l) \mid (2l-k-1). \end{cases}$$

Lemma V.12: Let $n \geq 1$ and $a, b \geq 0$ be integers and let C be a subset of a fixed size chosen uniformly at random from $0A^n$. Then

$$E[\mathcal{K}\left(1\ A^a C A^{a+b+2} \cap 1\ A^a 0\ A^b 0\ A^a 1\ A^n \cap A^{a+b+2} C A^{a+1}\right)] = \begin{cases} \frac{\mathcal{K}(C)^2}{4} - \frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} \\ \quad \text{if } n > a \text{ and } (b+1) \mid (a+1) \\ \frac{\mathcal{K}(C)^2}{4} \\ \quad \text{else.} \end{cases}$$

Lemma V.13: Let $n \geq 1$ and $a, b \geq 0$ be integers and let C be a subset of a fixed size chosen uniformly at random from $0A^n$. Then

$$E[\mathcal{K}\left(CA^{2a+b+3} \cap 0\ A^a 0\ A^b 0\ A^a 1\ A^n \cap 0 A^a C A^{a+b+2}\right)] = \begin{cases} \frac{\mathcal{K}(C)^2}{4} & \text{if } n \leq b \\ \frac{\mathcal{K}(C)^2}{4} & \text{if } b < n \leq a+b+1 \\ & \quad \text{and } (a+1) \nmid (b+1) \\ \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } b < n \leq a+b+1 \\ & \quad \text{and } (a+1) \mid (b+1) \\ \frac{\mathcal{K}(C)^2}{4} - \frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } n > a+b+1. \end{cases}$$

Lemma V.14: Let $n \geq 1$ and $a, b \geq 0$ be integers and let $l = a+b+3$. Let C be a subset of a fixed size of at least 1 chosen uniformly at random from $0A^n$, and let D be a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. Then

$$E[\mathcal{K}\left(DA^{a+1} \cap 1\ A^a 1\ A^b 0\ A^a 0\ A^n \cap A^{a+1} D\right)] = \begin{cases} \mathcal{K}\left(D\right)^2 \\ \quad \text{if } (a+1) \nmid (b+1) \\ \mathcal{K}\left(D\right)^2 - \frac{\mathcal{K}(D)}{|C|2^{l-2}-1} \cdot \left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}\left(D\right)\right) \\ \quad \text{if } (a+1) \mid (b+1). \end{cases}$$

Lemma V.15: Let $n \geq 1$ and $l \geq 3$ and

$$f(x,y) = \left(\frac{1}{4} - y\right)^2 - \frac{\frac{1}{2} - x}{2(2^n-1)}\left(\frac{x}{2} - y\right) - \frac{y}{x2^{n+l-1} - 1}\left(\frac{x}{2} - y\right) - \frac{1}{4(2^n-1)} x \left(\frac{1}{2} - x\right).$$

Then $f(x,y) \geq 0$ for all $x \in [\frac{1}{2^{n+1}}, \frac{1}{2} - \frac{1}{2^{n+1}}] \cup \{\frac{1}{2}\}$ and $y \in [0, \frac{x}{2} - \frac{1}{2^{n+l}}] \cup \{\frac{x}{2}\}$.

VI. Main Result, Part 1: $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\mu_2 2^{-\lambda_2} \leq \frac{1}{4}$

We are given three positive integers in increasing order, λ_1, λ_2, λ_3, and corresponding nonzero multiplicities μ_1, μ_2, μ_3, such that the Kraft sum $\mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} + \mu_3 2^{-\lambda_3}$ equals $3/4$. We are also given that the Kraft sums of the words of lengths λ_1 and λ_2 are upper bounded by $1/2$ and $1/4$, respectively. Our objective is to demonstrate that a fix-free code exists with the corresponding multiset of integers as codeword lengths.

We first construct length-λ_1 codewords by randomly removing a subset C of $0A^n$, whose size is chosen to leave exactly μ_1 codewords remaining. Then, length-λ_2 codewords are chosen from the words in $1A^{l-2}1A^n$, since none of them can have a prefix or suffix from the length-λ_1 words already chosen. Specifically, these words are chosen by randomly removing a subset D of $1A^{l-2}1A^n$, whose size is picked to leave exactly μ_2 words remaining after removal. The largest possible Kraft sum of the length-λ_2 words that can be achieved in this manner occurs when no words are removed, i.e., when $|D| = 0$. In this case, the expected Kraft sum of the length-λ_2 words is $\mathcal{K}\left(1A^{l-2}1A^n\right) = 1/4$, by Lemma V.2, which explains the upper bound on $\mu_2 2^{-\lambda_2}$ used in Theorem VI.2.

Finally, length-λ_3 words are constructed to avoid prefixes and suffixes in the randomly constructed sets of words of lengths λ_1 and λ_2.

It appears to be a somewhat difficult task to describe which codewords of lengths λ_1 and λ_2 to use in order to ensure the availability of the needed length-λ_3 codewords, while preserving the fix-free condition and the $3/4$ Kraft sum upper bound.

We use a probabilistic approach and remove the correct number of codewords of lengths λ_1 and λ_2 by random selection. In other words, we remove $2^{\lambda_1-1} - \mu_1$ of the original length-λ_1 codewords, uniformly at random from among the 2^{λ_1-1} original length-λ_1 codewords, and then we remove $2^{\lambda_2-2} - \mu_2$ of the original length-λ_2 codewords uniformly at random from among the 2^{λ_2-2} original length-λ_2 codewords. We prove that, on average, there are at least μ_3 codewords of length λ_3 that do not have any prefix or suffix from the resulting μ_1 codewords of length λ_1 and μ_2 codewords of length λ_2. So, there must exist at least one actual collection of μ_1 codewords of length λ_1 and μ_2 codewords of length λ_2 that result in at least μ_3 codewords of length λ_3 that have neither prefix nor suffix in the collection. This existential technique is somewhat analogous to that used in Shannon's proof of the channel coding theorem [60].

Throughout the proof of the main results, we shall use certain repeated terminology. The quantities λ_1, λ_2, and λ_3 will represent, in increasing order, the three distinct codeword lengths for the desired fix-free code. Throughout, we will use the quantities n, l, and k as defined in (1).

The proof of Theorem VI.2 is broken into three "overlap cases", namely when: (1) $2l - k < 1$; (2) $2l - k = 1$; and (3) $2l - k > 1$. These cases correspond, respectively, to when a length-λ_2 prefix and a length-λ_2 suffix of a length-λ_3 word: (1) overlap in at most n positions; (2) overlap in exactly $n+1$ positions; and (3) overlap in at least $n+2$ positions. The same three cases are also used to prove Theorems VII.2 and VIII.1. These three cases are illustrated in Figure 1.

The proof of Theorem VI.2 uses the following lemma, whose proof can be found in the appendix.

Lemma VI.1: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$. Let C be a set of a fixed size chosen uniformly at random from $0A^n$. Let D be a set of a fixed size chosen uniformly at random from $1A^{l-2}1\ A^n$. For any $b_1, b_2 \in A$, if $2l - k \neq 1$, then

(i) $\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n\right) = 1/16$

(ii) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}b_1 A^{n+k-l} \cap A^{k-l}0\ A^{l-2}1A^n\right)] = \mathcal{K}\left(C\right)/8$

(iii) $E[\mathcal{K}\left(1A^{l-2}0\ A^{n+k-l} \cap A^{k-l}b_1 A^{l-2}C\right)] = \mathcal{K}\left(C\right)/8$

(iv) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}b_2A^{l-2}C\right)]$
$= \mathcal{K}(C)^2/4$
(v) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)] = \mathcal{K}(D)/4$
(vi) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}D\right)] = \mathcal{K}(D)/4$
(vii) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}D\right)]$
$= \mathcal{K}(C)\mathcal{K}(D)/2$
(viii) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}b_1A^{l-2}C\right)]$
$= \mathcal{K}(C)\mathcal{K}(D)/2$
(ix) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)]$

$$= \begin{cases} \mathcal{K}(D)^2 & \text{if } 2l{-}k < 1 \\ \mathcal{K}(D)^2 & \text{if } 2l{-}k > 1 \\ & \text{and} \\ & (k{-}l) \nmid (2l{-}k{-}1) \\ \mathcal{K}(D)^2 + \frac{\mathcal{K}(D)(\frac{1}{4}-\mathcal{K}(D))}{2^{n+l-2}-1} & \text{if } 2l{-}k > 1 \\ & \text{and} \\ & (k{-}l) \mid (2l{-}k{-}1) \end{cases}$$

and if $2l-k=1$, then
(x) $\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n\right) = 1/8$
(xi) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n\right)]$
$= \mathcal{K}(C)/4$
(xii) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}C\right)]$
$= \mathcal{K}(C)/4$
(xiii) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}b_1A^{l-2}C\right)]$
$= \mathcal{K}(C)^2/2$
(xiv) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}1A^{n+k-l} \cap A^{k-l}D\right)]$
$= \mathcal{K}(C)\mathcal{K}(D)$
(xv) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}1A^{l-2}C\right)] = \mathcal{K}(C)\mathcal{K}(D)$
(xvi) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] = 2\mathcal{K}(D)^2$.

Theorem VI.2: Suppose a multiset of positive integers consists of μ_1 copies of λ_1, μ_2 copies of λ_2, and μ_3 copies of λ_3, such that $2 \le \lambda_1 < \lambda_2 < \lambda_3$. Then there exists a fix-free code with μ_1 codewords of length λ_1, μ_2 codewords of length λ_2, and μ_3 codewords of length λ_3, whenever the following conditions hold:

$$\mu_1 2^{-\lambda_1} \le \frac{1}{2}$$
$$\mu_2 2^{-\lambda_2} \le \frac{1}{4}$$
$$\mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} + \mu_3 2^{-\lambda_3} = \frac{3}{4}.$$

Proof: Let C be a set of size $2^n - \mu_1$ chosen uniformly at random from the 2^n length-λ_1 elements of $0A^n$, and let D be a set of size $2^{n+l-2} - \mu_2$ chosen uniformly at random from the 2^{n+l-2} length-λ_2 elements of $1A^{l-2}1\ A^n$. Define the following (random) sets:

$$F_1 = 0\ A^n - C$$
$$F_2 = 1\ A^{l-2}1\ A^n - D.$$

Then F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2. By Lemma V.2, we have $\mathcal{K}(F_1) = \mathcal{K}(0A^n) - \mathcal{K}(C) = \frac{1}{2} - \mathcal{K}(C) \le \frac{1}{2}$ and $\mathcal{K}(F_2) = \mathcal{K}\left(1A^{l-2}1A^n\right) - \mathcal{K}(D) = \frac{1}{4} - \mathcal{K}(D) \le \frac{1}{4}$.

In each of three cases, we will construct a third random set of words, F_3. The random set

$$F = F_1 \cup F_2 \cup F_3$$

on average forms the desired fix-free code. The union of non-random instances of F_1, F_2, and F_3 will then yield the asserted fix-free code. Let

$$Y_{i,j} = \begin{cases} CA^{k-1} \cap 0A^{l-2}jA^{n+k-l} & \text{if } i{=}0 \\ 1A^{l-2}0A^{n+k-l} & \text{if } i{=}1,\ j{=}0 \\ DA^{k-l} & \text{if } i{=}j{=}1 \end{cases}$$

$$W_{i,j} = \begin{cases} A^{k-l}iA^{l-2}C & \text{if } j{=}0 \\ A^{k-l}0A^{l-2}1A^n & \text{if } i{=}0,\ j{=}1 \\ A^{k-l}D & \text{if } i{=}j{=}1. \end{cases}$$

- Overlap Case 1: $2l{-}k < 1$.
  In this case, the set F_3 is built as a union of 16 disjoint subsets of A^{k+n}. The basic building block of each such subset is a pattern of the form $Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n$, where Z_1, Z_2, Z_3, Z_4 are fixed bits that ensure the 16 subsets are disjoint and are chosen to avoid prefixes or suffixes from F_1 or F_2. When these four bits do not prevent such prefixes or suffixes, the sets $Y_{i,j}$ and $W_{i,j}$ are constructed to remove offending prefixes or suffixes. These constructions can require certain subsets to have prefixes or suffixes in C and/or D. The terms in each intersection below satisfy

$$Y_{Z_1,Z_2} \cap W_{Z_3,Z_4} \subseteq Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n.$$

  Let $\mathcal{I} = A^4$ and define the set F_3, containing words of length λ_3, by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})$$

  Each of the 16 sets in the union comprising F_3 consists of words of length λ_3, and these sets, except when $(Z_1, Z_2, Z_3, Z_4) = (1, 0, 0, 1)$, are random, since they involve the random sets C or D. Thus, the Kraft sums of all but one term in the union are random variables.

  When $Z_1 = 0$ (respectively, $Z_4 = 0$), the words in the sets of the union are designed to contain prefixes (respectively, suffixes) in C in order to avoid prefixes (respectively, suffixes) in F_1, and when $Z_1 = Z_2 = 1$ (respectively, $Z_3 = Z_4 = 1$), the words in the sets of the union are designed to contain prefixes (respectively, suffixes) in D in order to avoid prefixes (respectively, suffixes) in F_2.

  It is easy to verify that none of the words of F_2 have prefixes or suffixes in F_1, none of the words of F_3 have prefixes or suffixes in F_1 or F_2, and that every two of the sets in the union forming F_3 are disjoint.

Next, we lower bound the expected Kraft sum of F_3:

$$\begin{aligned}
&E[\mathcal{K}(F_3)] \\
&= \sum_{(Z_1,Z_2,Z_3,Z_4)\in A^4} E[\mathcal{K}(Y_{Z_1,Z_2}\cap W_{Z_3,Z_4})] \\
&= \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)}{8} + \frac{\mathcal{K}(C)^2}{4} \\
&\quad + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)}{8} \\
&\quad + \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(C)}{8} \\
&\quad + \frac{1}{16} + \frac{\mathcal{K}(C)}{8} + \frac{\mathcal{K}(D)}{4} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} \\
&\quad + \frac{\mathcal{K}(D)}{4} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \mathcal{K}(D)^2 \qquad (2) \\
&= \left(\mathcal{K}(C)+\mathcal{K}(D)-\frac{1}{4}\right)^2 + \mathcal{K}(C)+\mathcal{K}(D) \\
&\geq \mathcal{K}(C)+\mathcal{K}(D) \\
&= \frac{1}{2} - \mu_1 2^{-\lambda_1} + \frac{1}{4} - \mu_2 2^{-\lambda_2} \qquad (3) \\
&= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \qquad (4)
\end{aligned}$$

where (2) follows from Lemma VI.1.

From (4), we can lower bound the expected size of the random set F_3 by

$$E[|F_3|] \geq 2^{\lambda_3}\left(\frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2}\right) = \mu_3.$$

There must exist at least one instance of the randomly constructed set F_3 that satisfies the same lower bound satisfied by the average size of F_3. Such an instance of the random set F_3 corresponds to some particular choices of the random sets C and D. Let $\widehat{F_1}$ and $\widehat{F_2}$ denote the resulting (non-random) instances of the random sets F_1 and F_2, respectively. Let $\widehat{F_3}$ denote the resulting (non-random) instance of F_3, but only after throwing away enough codewords to make the size of $\widehat{F_3}$ exactly equal to the lower bound on $E[|F_3|]$. That is,

$$|\widehat{F_3}| = \mu_3.$$

The code $\widehat{F_1}\cup\widehat{F_2}\cup\widehat{F_3}$ is fix-free, has Kraft sum equal to $3/4$, and has μ_1, μ_2, μ_3 codewords of sizes λ_1, λ_2, λ_3, respectively.

- Overlap Case 2: $2l-k=1$.
  In this case, the set F_3 is built in a similar manner as in Overlap Case 1, although here it will be a union of only 8 disjoint subsets of A^{k+n}, using patterns of the form $Z_1A^{l-2}Z_2A^{l-2}Z_3A^n$. The terms in each intersection below satisfy

$$Y_{Z_1,Z_2}\cap W_{Z_2,Z_3} \subseteq Z_1A^{l-2}Z_2A^{l-2}Z_3A^n.$$

Let $\mathcal{I}=A^3$ and define the set F_3 containing words of length λ_3, and lower bound its expected Kraft sum as follows:

$$\begin{aligned}
&F_3 = \bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}} (Y_{Z_1,Z_2}\cap W_{Z_2,Z_3}) \\
&E[\mathcal{K}(F_3)] \\
&= \sum_{(Z_1,Z_2,Z_3)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2}\cap W_{Z_2,Z_3})] \\
&= \frac{\mathcal{K}(C)^2}{2} + \frac{\mathcal{K}(C)}{4} + \frac{\mathcal{K}(C)^2}{2} + \mathcal{K}(C)\mathcal{K}(D) \\
&\quad + \frac{\mathcal{K}(C)}{4} + \frac{1}{8} + \mathcal{K}(C)\mathcal{K}(D) + 2\mathcal{K}(D)^2 \qquad (5) \\
&= \left(\mathcal{K}(C)+\mathcal{K}(D)-\frac{1}{4}\right)^2 + \left(\mathcal{K}(D)-\frac{1}{4}\right)^2 \\
&\quad + \mathcal{K}(C)+\mathcal{K}(D) \\
&\geq \mathcal{K}(C)+\mathcal{K}(D)
\end{aligned}$$

where (5) follows from Lemma VI.1. Overlap Case 2 is then finished by applying the same reasoning as used from (3) to the end of Overlap Case 1.

- Overlap Case 3: $2l-k>1$.
  This case is nearly identical to Overlap Case 1, but uses the following definition of F_3:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_3}\cap W_{Z_2,Z_4})$$

where

$$\begin{aligned}
&Y_{Z_1,Z_3}\cap W_{Z_2,Z_4} \\
&\subseteq Z_1A^{k-l-1}Z_2A^{2l-k-2}Z_3A^{k-l-1}Z_4A^n.
\end{aligned}$$

The only other difference is that the equal sign in (2) changes to $\geq$, since in Lemma VI.1(ix) when $2l-k>1$ we have

$$E[\mathcal{K}\left(DA^{k-l}\cap A^{k-l}D\right)] \geq \mathcal{K}(D)^2.$$

■

VII. Main Result, Part 2: $\mu_1 2^{-\lambda_1}\leq\frac{1}{2}$ and $\frac{1}{4}\leq\mu_2 2^{-\lambda_2}\leq\frac{1}{2}\left(1-\mu_1 2^{-\lambda_1}\right)$

In Theorem VI.2, sets of words of lengths λ_1 and λ_2 were initially constructed, and then some words were independently and uniformly removed from each set in order to bring their sizes down to μ_1 and μ_2, respectively. Then a set of length-λ_3 words was constructed that avoided prefixes and suffixes from the random sets of lengths λ_1 and λ_2. The constraint in Theorem VI.2 that $\mu_2 2^{-\lambda_2}\leq\frac{1}{4}$ allowed us to construct the set F_2 of length-λ_2 words entirely based on words from $1A^{l-2}1A^n$ (whose Kraft sum is $1/4$).

The construction for Theorem VII.2 is slightly different however, since this theorem requires $\mu_2 2^{-\lambda_2}\geq\frac{1}{4}$, but there are not enough words in $1A^{l-2}1A^n$ to get a Kraft sum larger than $1/4$ for the length-λ_2 words.

To solve this issue, in Theorem VII.2 we start with a larger set of length-λ_2 words, namely $1A^{l-2}1A^n\cup 1A^{l-2}C$, where C is a random set of words removed from $0A^n$ to leave exactly μ_1 of such words of length $n+1=\lambda_1$ remaining (just like in Theorem VI.2). Then, to construct a set of μ_2 length-λ_2 codewords, we remove a randomly selected subset D from $1A^{l-2}1A^n$ in Overlap Cases 1 and 3, and from $1A^{l-2}C$ in Overlap Case 2, where the cardinality of D ensures that there will be a total of μ_2 codewords of length λ_2 left after removal. In this manner, no length-λ_1 codewords can be prefixes or suffixes of any length-λ_2 codewords, since any word in the set $1A^{l-2}1A^n$ has a fixed bit of 1 where a length-λ_1 prefix or suffix from $0A^n$ would have a 0, and any word in the set $1A^{l-2}C$ has a fixed bit of 1 where a length-λ_1 prefix would have a 0, and has a length-λ_1 suffix from C, which, by construction, cannot be one of the μ_1 words of length λ_1 chosen for the fix-free code.

The largest Kraft sum of the length-λ_2 words that we can get with this technique is when we do not remove any codewords of length λ_2, i.e., when $|D| = 0$, in which case the expected Kraft sum of the length-λ_2 words, in all three overlap cases, is

$$\begin{aligned}
&\mathcal{K}\left(1A^{l-2}1A^n \cup 1A^{l-2}C\right) \\
&= \mathcal{K}\left(1A^{l-2}1A^n\right) + \mathcal{K}\left(1A^{l-2}C\right) \\
&= \mathcal{K}\left(1A^{l-2}1A^n\right) + \mathcal{K}\left(1A^{l-2}\right)\mathcal{K}\left(C\right) && (6)\\
&= \frac{1}{4} + \frac{1}{2}(2^n - \mu_1)2^{-(n+1)} && (7)\\
&= \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) && (8)
\end{aligned}$$

where (6) follows from Lemma V.2 and Corollary V.5; (7) follows from Lemma V.2 and the fact that $\mathcal{K}(C)$ equals the constant $(2^n - \mu_1)2^{-(n+1)}$; and (8) explains the upper bound on $\mu_2 2^{-\lambda_2}$ imposed in Theorem VII.2.

The proof of Theorem VII.2 uses the following lemma, whose proof can be found in the appendix.

Lemma VII.1: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$. Let C be a set of a fixed size chosen uniformly at random from $0A^n$. Let D_1 be a set of a fixed size chosen uniformly at random from $1A^{l-2}1\ A^n$, and let D_2 be a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. For any $b_1, b_2 \in A$, if $2l - k \neq 1$, then

(i) $E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)] = \left(\frac{1}{2} - \mathcal{K}(C)\right)/8$

(ii) $E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}C\right)] = \mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)/4$

(iii) $E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}D_1\right)] = \mathcal{K}(D_1)\left(\frac{1}{2} - \mathcal{K}(C)\right)/2$

and if $2l - k = 1$, then

(iv) $E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)] = \left(\frac{1}{2} - \mathcal{K}(C)\right)/4$

(v) $E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}1A^{n+k-l} \cap A^{k-l}D_2\right)] = \mathcal{K}(C)\mathcal{K}(D_2)$

(vi) $E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}C\right)] = \mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)/2$

(vii) $E[\mathcal{K}\left(D_2A^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)] = \mathcal{K}(D_2)/2$

(viii) $E[\mathcal{K}\left(D_2A^{k-l} \cap A^{k-l}0A^{l-2}C\right)] = \mathcal{K}(C)\mathcal{K}(D_2)$

Theorem VII.2: Suppose a multiset of positive integers consists of μ_1 copies of λ_1, μ_2 copies of λ_2, and μ_3 copies of λ_3, such that $2 \leq \lambda_1 < \lambda_2 < \lambda_3$, Then there exists a fix-free code with μ_1 codewords of length λ_1, μ_2 codewords of length λ_2, and μ_3 codewords of length λ_3, whenever the following conditions hold:

$$\begin{aligned}
\mu_1 2^{-\lambda_1} &\leq \frac{1}{2} \\
\frac{1}{4} \leq \mu_2 2^{-\lambda_2} &\leq \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) \\
\mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} + \mu_3 2^{-\lambda_3} &= \frac{3}{4}.
\end{aligned}$$

Proof: As in Part 1, let C be a set of size $2^n - \mu_1$ chosen uniformly at random from among the 2^n length-λ_1 elements of $0A^n$ and define the following (random) set:

$$F_1 = 0\ A^n - C.$$

- For Overlap Cases 1 and 3 below:
  Let D be a set of size $\frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right)2^{\lambda_2} - \mu_2$ chosen uniformly at random from among the 2^{n+l-2} length-λ_2 elements of $1A^{l-2}1A^n$, and let
  $$F_2 = (1A^{l-2}1A^n - D) \cup (1A^{l-2}C).$$
- For Overlap Case 2 below:
  Let D be a set of size $\frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right)2^{\lambda_2} - \mu_2$ chosen uniformly at random from among the $2^{l-2} \cdot |C|$ length-λ_2 elements of $1A^{l-2}C$, and let
  $$F_2 = 1A^{l-2}1A^n \cup (1A^{l-2}C - D).$$

Then $\mathcal{K}(D) = \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) - \mu_2 2^{-\lambda_2}$, so

$$0 \leq \mathcal{K}(D) \leq \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) - \frac{1}{4} = \frac{1}{4} - \mu_1 2^{-\lambda_1 - 1} \leq \frac{1}{4}.$$

This means there are enough words from which to choose D, since $\mathcal{K}\left(1A^{l-2}1A^n\right) = \frac{1}{4}$, and

$$\begin{aligned}
\mathcal{K}\left(1A^{l-2}C\right) = \frac{1}{2}\mathcal{K}(C) &= \frac{1}{2}\left(\frac{1}{2} - \mu_1 2^{-\lambda_1}\right) \\
&= \frac{1}{4} - \mu_1 2^{-\lambda_1 - 1},
\end{aligned}$$

by Lemma V.2.

Note that the words in F_2 all start with 1, and have a 0 in position l only if they have a suffix in C. These conditions guarantee that no word in F_1 is a prefix or suffix of a word in F_2. In all 3 cases,

$$\begin{aligned}
|F_1| &= |0A^n| - |C| = 2^n - (2^n - \mu_1) = \mu_1 \\
|F_2| &= |1\ A^{l-2}1\ A^n| + |1\ A^{l-2}C| - |D| \\
&= 2^{n+l-2} + 2^{l-2}(2^n - \mu_1) \\
&\quad - \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right)2^{\lambda_2} + \mu_2 \\
&= \mu_2
\end{aligned}$$

so the set F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2. The set F_1 can be viewed as being chosen uniformly at random among all subsets of $0A^n$ of size μ_1.

In each case, we will also construct a third random set F_3, consisting of μ_3 words of length λ_3. The random set

$$F = F_1 \cup F_2 \cup F_3$$

on average forms the desired fix-free code. The union of non-random instances of F_1, F_2, and F_3 will then yield the asserted fix-free code.

- Overlap Case 1: $2l-k < 1$.
  Let

$$Y_{i,j} = \begin{cases} CA^{k-1} \cap 0A^{l-2}jA^{n+k-l} & \text{if } i{=}0 \\ 1A^{l-2}(0A^n - C)A^{k-l} & \text{if } i{=}1,\ j{=}0 \\ DA^{k-l} & \text{if } i{=}j{=}1 \end{cases}$$

$$W_{i,j} = \begin{cases} A^{k-l}0A^{l-2}C & \text{if } i{=}j{=}0 \\ A^{k-l}0A^{l-2}1A^n & \text{if } i{=}0,\ j{=}1 \\ A^{k-l}D & \text{if } i{=}j{=}1. \end{cases} \tag{9}$$

Let $\mathcal{I} = A^4 - A^2 10$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})$$

where

$$Y_{Z_1,Z_2} \cap W_{Z_3,Z_4} \subseteq Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n.$$

In contrast to Overlap Cases 1 and 3 of Part 1, here F_3 is comprised of only 12 of the 16 possible sets obtained from the pattern

$$Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n,$$

namely by excluding $(Z_3, Z_4) = (1, 0)$ from the union. One can verify that no words in F_1 or F_2 can be either prefixes or suffixes of any words in F_3.

The expected Kraft sum of F_3 is then lower bounded as follows:

$$\begin{aligned} &E[\mathcal{K}(F_3)] \\ &= \sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})] \\ &= \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)}{8} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(C)^2}{4} \\ &\quad + \frac{\mathcal{K}(C)}{8} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))}{4} \\ &\quad + \frac{\frac{1}{2} - \mathcal{K}(C)}{8} + \frac{\mathcal{K}(D)(\frac{1}{2} - \mathcal{K}(C))}{2} \\ &\quad + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(D)}{4} + \mathcal{K}(D)^2 \qquad (10) \\ &= \left(\frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) - \frac{1}{4}\right)^2 + \frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) \\ &\geq \frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) \qquad (11) \\ &= \frac{1}{4} - \frac{1}{2}\mu_1 2^{-\lambda_1} + \frac{1}{2} - \frac{1}{2}\mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \qquad (12) \\ &= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \end{aligned}$$

where (10) follows from Lemma VII.1 when $(Z_1, Z_2) = (1, 0)$ and otherwise follows from Lemma VI.1; and (12) follows from the quantities $|C|$ and $|D|$ defined at the beginning of the proof of this theorem.

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 2: $2l-k = 1$.
  Let

$$Y_{i,j} = \begin{cases} CA^{k-1} \cap 0A^{l-2}jA^{n+k-l} & \text{if } i{=}0 \\ (D \cup 1A^{l-2}(0A^n - C))A^{k-l} & \text{if } i{=}1,\ j{=}0 \end{cases}$$

$$W_{i,j} = \begin{cases} A^{k-l}0A^{l-2}C & \text{if } i{=}j{=}0 \\ A^{k-l}0A^{l-2}1A^n & \text{if } i{=}0,\ j{=}1 \\ A^{k-l}D & \text{if } i{=}1,\ j{=}0. \end{cases}$$

Let $\mathcal{I} = A^3 - (11A \cup A11)$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})$$

where

$$Y_{Z_1,Z_2} \cap W_{Z_2,Z_3} \subseteq Z_1A^{l-2}Z_2A^{l-2}Z_3A^n.$$

In this case, F_3 is comprised of only 5 of the 8 possible sets obtained from the pattern

$$Z_1A^{l-2}Z_2A^{l-2}Z_3A^n,$$

namely by excluding (Z_1, Z_2, Z_3) from being $(1, 1, 0)$, $(0, 1, 1)$, or $(1, 1, 1)$ in the union.

The expected Kraft sum of F_3 can be lower bounded as follows:

$$\begin{aligned} &E[\mathcal{K}(F_3)] \\ &= \sum_{(Z_1,Z_2,Z_3)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})] \\ &= \frac{\mathcal{K}(C)^2}{2} + \frac{\mathcal{K}(C)}{4} + \mathcal{K}(C)\mathcal{K}(D) \\ &\quad + \left(\frac{\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))}{2} + \mathcal{K}(C)\mathcal{K}(D)\right) \\ &\quad + \left(\frac{\frac{1}{2} - \mathcal{K}(C)}{4} + \frac{\mathcal{K}(D)}{2}\right) \qquad (13) \\ &= \frac{(1 - 2\mathcal{K}(C))(1 - 4\mathcal{K}(D))}{8} \\ &\quad + \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) \\ &\geq \frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) \qquad (14) \end{aligned}$$

where (13) follows from Lemma VI.1 when $Z_1 = Z_2 = 0$ and otherwise follows from Lemma VII.1; and (14) follows since $\mathcal{K}(D) \leq 1/4$ and $\mathcal{K}(C) \leq 1/2$.

Overlap Case 2 is then finished by applying the same reasoning as used from (11) to the end of Overlap Case 1.

- Overlap Case 3: $2l-k > 1$.
  We use the same sets $Y_{i,j}$ and $W_{i,j}$ as defined in (9) for Overlap Case 1. Let $\mathcal{I} = A^4 - A1A0$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_3} \cap W_{Z_2,Z_4})$$

where

$$\begin{aligned} &Y_{Z_1,Z_3} \cap W_{Z_2,Z_4} \\ &\quad \subseteq Z_1A^{k-l-1}Z_2A^{2l-k-2}Z_3A^{k-l-1}Z_4A^n. \end{aligned}$$

Here F_3 is comprised of 12 of the 16 possible sets obtained by excluding $(Z_2, Z_4) = (1, 0)$.

The expected Kraft sum of F_3 is then lower bounded as follows:

$$\begin{aligned} &E[\mathcal{K}(F_3)] \\ &= \sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_3} \cap W_{Z_2,Z_4})] \\ &\geq \left(\frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) - \frac{1}{4}\right)^2 + \frac{\mathcal{K}(C)}{2} + \mathcal{K}(D) \quad (15) \end{aligned}$$

where the $\geq$ in (15) follows from Lemma VI.1(ix), and the remainder of the proof is the same as for Overlap Case 1 starting at (11).

■

VIII. Main Result, Part 3: $\mu_1 2^{-\lambda_1} \leq \frac{1}{2}$ and $\frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) \leq \mu_2 2^{-\lambda_2}$

The methods for constructing codes in this section are similar in spirit to the methods from the past two sections, but contain some significant changes as well. As before, sets of words of lengths λ_1 and λ_2 are initially constructed, but in this section sometimes we will then remove words at random from these sets, and sometimes we will add words at random to these sets. Specifically, in the proofs of Theorem VIII.1, parts (a), (b), and (c), some words are removed, and in part (d) some words are added, but in all cases the resulting words of lengths λ_1 and λ_2 have cardinalities μ_1 and μ_2, respectively. Then, just as in previous sections, we will show that there are enough words of length λ_3 available to produce the desired fix-free code.

There are additional complications in this section that result in more cases to consider than in the previous sections. Before, we removed words of length λ_2 from $1A^{l-2}1A^n$ or $1A^{l-2}C$ only, but in this section we will need to remove words of length λ_2 from $CA^{l-1} \cap 0A^{l-2}1A^n$ as well. However, if $n > l-2$ and C happens to be a subset of $0A^{l-2}0A^{n-l+1}$, then $CA^{l-1} \cap 0A^{l-2}1A^n = \varnothing$, leaving us no length-$\lambda_2$ words to remove from this set. To remedy this, we split the proof into separate lemmas, where we first consider the case when $n \leq l-2$ (in which we proceed in a similar fashion as the previous sections), and then consider when $n > l-2$. This latter case requires us to take more care in choosing C, and so we break this case into three separate lemmas.

Additionally, it turns out that as n grows, other complications can arise depending on the values of the lengths λ_2 and λ_3. As can be seen in Lemma VIII.2, particularly in cases (ix)–(xii), the expected Kraft sums of certain sets may depend on specific divisibility conditions involving the codeword lengths. These conditions are a result of the ways in which randomly chosen codewords may overlap each other as factors in codewords of a larger length. Fortunately, these complications are present in Overlap Case 3 of Theorem VIII.1(a) only, and we use Lemma V.15 to prove our desired result even in this case.

Theorem VIII.1: Suppose a multiset of positive integers consists of μ_1 copies of λ_1, μ_2 copies of λ_2, and μ_3 copies of λ_3, such that $2 \leq \lambda_1 < \lambda_2 < \lambda_3$. Then there exists a fix-free code with μ_1 codewords of length λ_1, μ_2 codewords of length λ_2, and μ_3 codewords of length λ_3, whenever the following conditions hold:

$$\begin{aligned} \mu_1 2^{-\lambda_1} &\leq \frac{1}{2} \\ \frac{1}{2}\left(1 - \mu_1 2^{-\lambda_1}\right) &\leq \mu_2 2^{-\lambda_2} \\ \mu_1 2^{-\lambda_1} + \mu_2 2^{-\lambda_2} + \mu_3 2^{-\lambda_3} &= \frac{3}{4}. \end{aligned}$$

Theorem VIII.1 follows immediately from the following four cases, which depend on the values of λ_1, λ_2, μ_1, and μ_2:

(a) $\lambda_2 \geq 2\lambda_1$
(b) $\lambda_2 < 2\lambda_1$ and $\frac{1}{4} \leq \mu_1 2^{-\lambda_1} \leq \frac{1}{2}$
(c) $\lambda_2 < 2\lambda_1$ and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\frac{1}{4} \leq \mu_2 2^{-\lambda_2} \leq \frac{1}{2}$
(d) $\lambda_2 < 2\lambda_1$ and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\frac{1}{2} < \mu_2 2^{-\lambda_2}$.

A. Proof of Theorem VIII.1(a)

The proof of Theorem VIII.1(a) uses the following lemma, whose proof can be found in the appendix.

Lemma VIII.2: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$ and $n \leq l-2$. Let C be a set of a fixed size chosen uniformly at random from $0A^n$. Let D_1 be a set of a fixed size chosen uniformly at random from $1A^{l-2}1\ A^n$, and let D_2 be a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. For any $b \in A$, if $2l - k < 1$, then

- (i)
$$\begin{aligned} &E[\mathcal{K}\left(CA^{l-2-n}0A^{n+k-l}\right. \\ &\qquad \left.\cap A^{k-l}(0A^n - C)A^{l-2-n}1\ A^n\right)] \\ &= \mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)/4 \end{aligned}$$
- (ii)
$$\begin{aligned} &E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l}\right. \\ &\qquad \left.\cap A^{k-l}(0A^n - C)A^{l-2-n}1\ A^n\right)] \\ &= \left(\frac{1}{2} - \mathcal{K}(C)\right)^2/4 \end{aligned}$$
- (iii) $E[\mathcal{K}\left(D_1A^{k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1\ A^n\right)]$ $= \mathcal{K}(D_1)\left(\frac{1}{2} - \mathcal{K}(C)\right)/2.$

If $2l - k = 1$, then

- (iv)
$$\begin{aligned} &E[\mathcal{K}\left(CA^{l-2-n}0A^{n+k-l}\right. \\ &\qquad \left.\cap A^{k-l}(0A^n - C)A^{l-2-n}1\ A^n\right)] \\ &= \mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)/2 \end{aligned}$$
- (v)
$$\begin{aligned} &E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l}\right. \\ &\qquad \left.\cap A^{k-l}(0A^n - C)A^{l-2-n}1\ A^n\right)] \\ &= \left(\frac{1}{2} - \mathcal{K}(C)\right)/4. \end{aligned}$$

If $2l-k>1$, then

- (vi) $E[\mathcal{K}\left(CA^{l-2-n}0\ A^{n+k-l}\cap A^{k-l}D_2\right)]$ $=\mathcal{K}\left(C\right)\mathcal{K}\left(D_2\right)/2$
- (vii) $E[\mathcal{K}\left(1A^{l-2}(0A^n-C)A^{k-l}\cap A^{k-l}D_2\right)]$ $=(\frac{1}{2}-\mathcal{K}\left(C\right))\mathcal{K}\left(D_2\right)/2$
- (viii) $E[\mathcal{K}\left(D_2A^{k-l}\cap A^{k-l}0A^{l-2}C\right)]$ $=\mathcal{K}\left(C\right)\mathcal{K}\left(D_2\right)/2$
- (ix)
$$\begin{aligned}&E[\mathcal{K}\left(CA^{l-2-n}0\ A^{n+k-l}\right.\\&\qquad\left.\cap A^{k-l}(0A^n-C)A^{l-2-n}1A^n\right)]\\&=\frac{\mathcal{K}\left(C\right)\left(\frac{1}{2}-\mathcal{K}\left(C\right)\right)}{4}\\&\quad-\begin{cases}\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } n>2l-k-2\\ & \quad\text{and } (k-l)\mid(2l-k-1)\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$
- (x)
$$\begin{aligned}&E[\mathcal{K}\left(1A^{l-2}(0A^n-C)A^{k-l}\right.\\&\qquad\left.\cap A^{k-l}(0A^n-C)A^{l-2-n}1A^n\right)]\\&=\frac{(\frac{1}{2}-\mathcal{K}\left(C\right))^2}{4}\\&\quad-\begin{cases}\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } n>k-l-1\\ & \quad\text{and } (2l-k-1)\mid(k-l)\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$
- (xi)
$$\begin{aligned}&E[\mathcal{K}\left(D_2A^{k-l}\cap A^{k-l}(0A^n-C)A^{l-2-n}1A^n\right)]\\&=\frac{(\frac{1}{2}-\mathcal{K}\left(C\right))\mathcal{K}\left(D_2\right)}{2}\\&\quad+\begin{cases}\frac{\mathcal{K}(D_2)(\frac{1}{2}-\mathcal{K}(C))}{2(2^n-1)} & \text{if } n>k-l-1\\ & \quad\text{and } (2l-k-1)\mid(k-l)\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$
- (xii)
$$\begin{aligned}&E[\mathcal{K}\left(D_2A^{k-l}\cap A^{k-l}D_2\right)]\\&=\mathcal{K}\left(D_2\right)^2\\&\quad-\begin{cases}\frac{\mathcal{K}(D_2)}{|C|\cdot 2^{l-2}-1}\left(\frac{\mathcal{K}(C)}{2}-\mathcal{K}\left(D_2\right)\right)\\ \qquad\text{if } (k-l)\mid(2l-k-1)\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Proof of Theorem VIII.1(a): Here we assume $\lambda_2\geq 2\lambda_1$ (or equivalently $n\leq l-2$ by (1)).

Let C be a set of size $2^n-\mu_1$ chosen uniformly at random from among the 2^n length-λ_1 elements of $0A^n$. Note that $0\leq\mu_1\leq 2^n$ since $0\leq\mu_1 2^{-\lambda_1}\leq\frac{1}{2}$. Also,

$$\mathcal{K}\left(C\right)=\left(2^n-\mu_1\right)2^{-\lambda_1}=\frac{1}{2}-\mu_1 2^{-\lambda_1}.$$

Define the following (random) set:

$$F_1=0A^n-C.$$

- For Overlap Case 1 below:
Let D be a set of size $(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2}-\mu_2$ chosen uniformly at random from among the 2^{n+l-2} length-λ_2 elements of $1A^{l-2}1A^n$, and let
$$F_2=(1A^{l-2}1A^n-D)\cup 1A^{l-2}C\cup CA^{l-2-n}1A^n.$$
- For Overlap Cases 2 and 3 below:
Let D be a set of size $(\frac{3}{4}-\mu_1 2^{-\lambda_1})2^{\lambda_2}-\mu_2$ chosen uniformly at random from among the $2^{l-2}\cdot|C|$ length-λ_2 elements of $1A^{l-2}C$, and let
$$F_2=1A^{l-2}1A^n\cup(1A^{l-2}C-D)\cup CA^{l-2-n}1A^n.$$

Then

$$\begin{aligned}\mathcal{K}\left(D\right)&=|D|\cdot 2^{-\lambda_2}\\&=\frac{3}{4}-\mu_1 2^{-\lambda_1}-\mu_2 2^{-\lambda_2}\\&\leq\frac{3}{4}-\mu_1 2^{-\lambda_1}-\frac{1}{2}+\frac{\mu_1 2^{-\lambda_1}}{2}\\&=\frac{1}{4}-\frac{\mu_1 2^{-\lambda_1}}{2}.\end{aligned}\tag{16}$$

This means there are enough words from which to choose D, since

$$\frac{1}{4}-\frac{\mu_1 2^{-\lambda_1}}{2}\leq\frac{1}{4}=\mathcal{K}\left(1A^{l-2}1A^n\right)$$

and

$$\frac{1}{4}-\frac{\mu_1 2^{-\lambda_1}}{2}=\frac{\mathcal{K}\left(C\right)}{2}=\mathcal{K}\left(1A^{l-2}C\right).$$

Note that the $(n+1)$-bit prefixes and suffixes of words in F_2 either start with 1 or else lie in C, whereas all words in F_1 start with 0 and cannot lie in C. So no word in F_1 can be a prefix or a suffix of a word in F_2.

Also in all three cases,

$$\begin{aligned}|F_1|&=|0A^n|-|C|=2^n-(2^n-\mu_1)=\mu_1\\|F_2|&=|1A^{l-2}1A^n|+|1A^{l-2}C|+|CA^{l-2-n}1A^n|-|D|\\&=2^{n+l-2}+2^{l-2}(2^n-\mu_1)+2^{l-2}(2^n-\mu_1)\\&\quad-\left(\frac{3}{4}-\mu_1 2^{-\lambda_1}\right)2^{\lambda_2}+\mu_2\\&=\mu_2\end{aligned}$$

so the set F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2.

In each case, we construct a third random set F_3 consisting of μ_3 words, each of length λ_3. The random set

$$F=F_1\cup F_2\cup F_3$$

on average forms the desired fix-free code. The union of non-random instances of F_1, F_2, and F_3 will then yield the asserted fix-free code.

- Overlap Case 1: $2l-k<1$. Let

$$Y_{i,j}=\begin{cases}CA^{l-2-n}0A^{n+k-l} & \text{if } i=j=0\\ 1A^{l-2}(0A^n-C)A^{k-l} & \text{if } i=1,\ j=0\\ DA^{k-l} & \text{if } i=j=1\end{cases}$$

$$W_{i,j}=\begin{cases}A^{k-l}0\ A^{l-2}C & \text{if } i=j=0\\ A^{k-l}(0A^n-C)A^{l-2-n}1A^n & \text{if } i=0,\ j=1\\ A^{k-l}D & \text{if } i=j=1.\end{cases}$$

Let $\mathcal{I}=A^4-(01A^2\cup A^2 10)$ and define the set F_3 by:

$$F_3=\bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}}(Y_{Z_1,Z_2}\cap W_{Z_3,Z_4})$$

where

$$Y_{Z_1,Z_2}\cap W_{Z_3,Z_4}\subseteq Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n.$$

Here F_3 is comprised of only 9 of the 16 possible sets obtained from the pattern $Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n$, namely by excluding patterns with $(Z_1,Z_2)=(0,1)$ or $(Z_3,Z_4)=(1,0)$. One can verify that no words in F_1 or F_2 are either prefixes or suffixes of any words in F_3.

The expected Kraft sum of F_3 is then lower bounded as follows:

$$\begin{aligned}
&E[\mathcal{K}(F_3)]\\
&=\sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}}E[\mathcal{K}(Y_{Z_1,Z_2}\cap W_{Z_3,Z_4})]\\
&=\frac{\mathcal{K}(C)^2}{4}+\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4}+\frac{\mathcal{K}(C)\mathcal{K}(D)}{2}\\
&\quad+\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4}+\frac{(\frac{1}{2}-\mathcal{K}(C))^2}{4}\\
&\quad+\frac{\mathcal{K}(D)(\frac{1}{2}-\mathcal{K}(C))}{2}+\frac{\mathcal{K}(C)\mathcal{K}(D)}{2}\\
&\quad+\frac{\mathcal{K}(D)(\frac{1}{2}-\mathcal{K}(C))}{2}+\mathcal{K}(D)^2 &(17)\\
&=\left(\mathcal{K}(D)-\frac{1}{4}\right)^2+\mathcal{K}(D)\\
&\geq\mathcal{K}(D)\\
&=\frac{3}{4}-\mu_1 2^{-\lambda_1}-\mu_2 2^{-\lambda_2} &(18)
\end{aligned}$$

where (17) follows from Lemma VI.1 when both $Z_1=Z_2$ and $Z_3=Z_4$, from Lemma VII.1 when both $(Z_1,Z_2)=(1,0)$ and $Z_3=Z_4$, from Lemma VIII.2 when $(Z_3,Z_4)=(0,1)$; and (18) follows from (16).

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 2: $2l-k=1$.

Let

$$Y_{i,j}=\begin{cases}CA^{l-2-n}0A^{n+k-l} & \text{if } i=j=0\\ (D\cup 1A^{l-2}(0A^n-C))A^{k-l} & \text{if } i=1,\ j=0\end{cases}$$

$$W_{i,j}=\begin{cases}A^{k-l}0A^{l-2}C & \text{if } i=j=0\\ A^{k-l}(0A^n-C)A^{l-2-n}1A^n & \text{if } i=0,\ j=1.\end{cases}$$

Let $\mathcal{I}=A^3-A1A$ and define the set F_3 by:

$$F_3=\bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}}(Y_{Z_1,Z_2}\cap W_{Z_2,Z_3})$$

where the terms in the union satisfy

$$Y_{Z_1,0}\cap W_{0,Z_3}\subseteq Z_1A^{l-2}0\ A^{l-2}Z_3A^n.$$

In this case, F_3 is comprised of 4 of the 8 possible sets obtained from the pattern $Z_1A^{l-2}Z_2A^{l-2}Z_3A^n$, namely excluding (Z_1,Z_2,Z_3) being $(0,1,0)$, $(1,1,0)$, $(0,1,1)$, or $(1,1,1)$. Therefore, these conditions are equivalent to $Z_2\neq 1$.

The expected Kraft sum of F_3 can be lower bounded as follows:

$$\begin{aligned}
&E[\mathcal{K}(F_3)]\\
&=\sum_{(Z_1,Z_2,Z_3)\in\mathcal{I}}E[\mathcal{K}(Y_{Z_1,Z_2}\cap W_{Z_2,Z_3})]\\
&=\frac{\mathcal{K}(C)^2}{2}+\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{2}\\
&\quad+\left(\mathcal{K}(C)\mathcal{K}(D)+\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{2}\right)\\
&\quad+\left(0+\frac{\frac{1}{2}-\mathcal{K}(C)}{4}\right) &(19)\\
&=\frac{\mathcal{K}(C)}{2}\left(\frac{1}{2}-\mathcal{K}(C)\right)+\mathcal{K}(C)\mathcal{K}(D)+\frac{1}{8}\\
&\geq\left(\frac{\mathcal{K}(C)}{2}-\mathcal{K}(D)\right)\left(\frac{1}{2}-\mathcal{K}(C)\right)+\mathcal{K}(D) &(20)\\
&\geq\mathcal{K}(D) &(21)\\
&=\frac{3}{4}-\mu_1 2^{-\lambda_1}-\mu_2 2^{-\lambda_2} &(22)
\end{aligned}$$

where (19) follows from Lemma VI.1 when $Z_1=Z_2=Z_3=0$, from Lemma VII.1 when $(Z_1,Z_2,Z_3)=(1,0,0)$, and otherwise from Lemma VIII.2 and the fact that $D\cap A^{k-l}(0A^n-C)=\varnothing$; (20) follows since $\frac{1}{8}\geq\frac{\mathcal{K}(D)}{2}$; (21) follows since $\mathcal{K}(C)\leq\frac{1}{2}$ and $\mathcal{K}(D)\leq\frac{\mathcal{K}(C)}{2}$; and (22) follows from (16).

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 3: $2l-k>1$.

Let

$$Y_{i,j}=\begin{cases}CA^{l-2-n}0A^{n+k-l} & \text{if } i=j=0\\ (D\cup 1A^{l-2}(0A^n-C))A^{k-l} & \text{if } i=1,\ j=0\end{cases}$$

$$W_{i,j}=\begin{cases}A^{k-l}0\ A^{l-2}C & \text{if } i=j=0\\ A^{k-l}(0A^n-C)A^{l-2-n}1A^n & \text{if } i=0,\ j=1\\ A^{k-l}D & \text{if } i=1,\ j=0.\end{cases}$$

Let $\mathcal{I}=A^4-(A^2 1A\cup A1A1)$ and define the set F_3 by:

$$F_3=\bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}}(Y_{Z_1,Z_3}\cap W_{Z_2,Z_4}).$$

where

$$Y_{Z_1,Z_3} \cap W_{Z_2,Z_4} \subseteq Z_1 A^{k-l-1} Z_2 A^{2l-k-2} Z_3 A^{k-l-1} Z_4 A^n.$$

Here F_3 is comprised of only 6 of the 16 possible sets, namely by excluding patterns with $(Z_1, Z_3) \in \{(0,1),(1,1)\}$ or $(Z_2, Z_4) = (1,1)$ from the union.

Let 1_a be the indicator function for the condition $(k-l) \mid (2l-k-1)$, let 1_b be the indicator function for the condition $(2l-k-1) \mid (k-l)$, let 1_c be the indicator function for the condition $n > k-l-1$, and let 1_d be the indicator function for the condition $n > 2l-k-2$.

Regarding 1_a and 1_b, note that $k-l$ is the length of any word in $Z_1 A^{k-l-1}$, and $2l-k-1$ is the length of any word in $Z_2 A^{2l-k-2}$. If $1_c = 1$, then any word from $0A^n$ that is a prefix of a term in the union above must extend at least to the bit Z_2, which would cause, for example, overlap in prefixes of $Y_{0,0}$ that lie in C and subwords of $W_{0,1}$ that lie in $0A^n - C$. Also, if $1_d = 1$, then any word from $0A^n$ that is a subword in a term in the union above that starts at the Z_2 position must extend at least to the bit Z_3, which would cause overlap in subwords of $Y_{1,0}$ that lie in $0A^n - C$ and subwords of $W_{0,1}$ that lie in $0A^n - C$. It turns out that there are four such complications that arise in this overlap case, which are considered in cases (ix)–(xii) of Lemma VIII.2.

If $1_a = 1_b = 1_c = 1_d = 0$, then the expected Kraft sum of F_3 is

$$\begin{aligned}
&E[\mathcal{K}(F_3)] \\
&= \sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_3} \cap W_{Z_2,Z_4})] \qquad (23)\\
&= \frac{\mathcal{K}(C)^2}{4} + \frac{\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))}{4} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} \\
&\quad + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \frac{\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))}{4} \\
&\quad + \frac{\mathcal{K}(D)(\frac{1}{2} - \mathcal{K}(C))}{2} + \frac{(\frac{1}{2} - \mathcal{K}(C))^2}{4} \\
&\quad + \mathcal{K}(D)^2 + \frac{\mathcal{K}(D)(\frac{1}{2} - \mathcal{K}(C))}{2} \qquad (24)\\
&= \left(\frac{1}{4} + \mathcal{K}(D)\right)^2
\end{aligned}$$

where (24) follows from Lemma VI.1 when $Z_1 = Z_2 = Z_3 = Z_4 = 0$, from Lemma VII.1 for part of the case when $(Z_1, Z_2, Z_3, Z_4) = (1,0,0,0)$, and otherwise from Lemma VIII.2 using the fact that $1_a = 1_b = 1_c = 1_d = 0$.

Of the 6 terms in the summation of (23), the 3 terms corresponding to (Z_1, Z_2, Z_3, Z_4) equaling $(0,0,0,0)$, $(0,1,0,0)$, and $(1,0,0,0)$, remain the same even when it's not the case that $1_a = 1_b = 1_c = 1_d = 0$. The values of the remaining 3 terms of the summation, i.e., when (Z_1, Z_2, Z_3, Z_4) is $(0,0,0,1)$, $(1,0,0,1)$, or $(1,1,0,1)$, are obtained from Lemma VIII.2 using

$$\begin{aligned}
&E[\mathcal{K}\left(CA^{l-2-n}0A^{n+k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1A^n\right)] \\
&= \frac{\mathcal{K}(C)\mathcal{K}(0A^n - C)}{4} \\
&\quad - 1_a 1_d \cdot \frac{\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))}{4(2^n - 1)} \qquad (25)\\
&E[\mathcal{K}\left(1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1A^n\right)] \\
&= \frac{\mathcal{K}(0A^n - C)^2}{4} - 1_b 1_c \cdot \frac{\frac{1}{2} - \mathcal{K}(C)}{4(2^n - 1)}\mathcal{K}(C) \qquad (26)\\
&E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1A^n\right)] \\
&= \frac{\mathcal{K}(D)\mathcal{K}(0A^n - C)}{2} \\
&\quad + 1_b 1_c \cdot \frac{\frac{1}{2} - \mathcal{K}(C)}{2(2^n - 1)} \cdot \mathcal{K}(D) \qquad (27)\\
&E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\
&= \mathcal{K}(D)^2 \\
&\quad - 1_a \cdot \frac{\mathcal{K}(D)}{|C| \cdot 2^{l-2} - 1}\left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D)\right). \qquad (28)
\end{aligned}$$

The first expressions on the right hand sides of (25)–(28) correspond to those in the calculations used to obtain (24), i.e., when $1_a = 1_b = 1_c = 1_d = 0$. Therefore, in general, the expected Kraft sum of F_3 is given by

$$\begin{aligned}
&E[\mathcal{K}(F_3)] \\
&= \left(\frac{1}{4} + \mathcal{K}(D)\right)^2 \\
&\quad - 1_b 1_c \cdot \frac{(\frac{1}{2} - \mathcal{K}(C))(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D))}{2(2^n - 1)} \\
&\quad - 1_a 1_d \cdot \frac{\mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)}{4(2^n - 1)} \\
&\quad - 1_a \cdot \frac{\mathcal{K}(D)\left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D)\right)}{|C| \cdot 2^{l-2} - 1}. \qquad (29)
\end{aligned}$$

We will show that for any binary values of 1_a, 1_b, 1_c, and 1_d, we have $E[\mathcal{K}(F_3)] \geq \mathcal{K}(D)$. Since $\mathcal{K}(C) \leq 1/2$ and $\mathcal{K}(D) \leq \mathcal{K}(C)/2$, the quantities multiplying $1_b 1_c$, $1_a 1_d$, and 1_a are all non-positive. Thus, it suffices to show that with $1_a = 1_b = 1_c = 1_d = 1$, the quantity in (29) minus $\mathcal{K}(D)$ is non-negative for any

$$\mathcal{K}(C) \in \{0\} \cup \left[\frac{1}{2^{n+1}}, \frac{1}{2} - \frac{1}{2^{n+1}}\right] \cup \left\{\frac{1}{2}\right\}$$

and

$$\mathcal{K}(D) \in \left[0, \frac{\mathcal{K}(C)}{2} - \frac{1}{2^{n+l}}\right] \cup \left\{\frac{\mathcal{K}(C)}{2}\right\}.$$

These ranges for $\mathcal{K}(C)$ and $\mathcal{K}(D)$ are sufficient to finish the proof, since $|C|$ and $|D|$ are integers, and so it is not possible that

$$\mathcal{K}(C) \in \left(0, \frac{1}{2^{n+1}}\right) \cup \left(\frac{1}{2} - \frac{1}{2^{n+1}}, \frac{1}{2}\right)$$

or

$$\mathcal{K}(D) \in \left(\frac{\mathcal{K}(C)}{2} - \frac{1}{2^{n+l}}, \frac{\mathcal{K}(C)}{2}\right).$$

Since $2l-k>1$ and $k \geq 3$, we have $l \geq 3$. If $\mathcal{K}(C)=0$, then the original multiset of lengths contains only two distinct values, and this case is covered by Theorem III.1. So suppose $\mathcal{K}(C) \geq 1/2^{n+1}$. Then

$$\begin{aligned}
&E[\mathcal{K}(F_3)] - \mathcal{K}(D)\\
&= \left(\frac{1}{4} - \mathcal{K}(D)\right)^2\\
&\quad - 1_b 1_c \cdot \frac{(\frac{1}{2} - \mathcal{K}(C))\left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D)\right)}{2(2^n-1)}\\
&\quad - 1_a \cdot \frac{\mathcal{K}(D)\left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D)\right)}{|C| \cdot 2^{l-2} - 1}\\
&\quad - 1_a 1_d \cdot \frac{\mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)}{4(2^n-1)} \qquad (30)\\
&\geq 0 \qquad (31)
\end{aligned}$$

where (31) follows by first setting $1_a = 1_b = 1_c = 1_d = 1$ to minimize (30), and then applying Lemma V.15 by setting $x = \mathcal{K}(C)$ and $y = \mathcal{K}(D)$. The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1. ■

B. Proof of Theorem VIII.1(b)

The proof of Theorem VIII.1(b) uses the following lemma, whose proof can be found in the appendix.

Lemma VIII.3: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$ and $n \geq l-1$. Let C be a set of a fixed size chosen uniformly at random from $0A^{l-2}1A^{n-(l-1)}$. Let $G = 0A^{l-2}1A^{n-(l-1)} - C$. Let D_1 be a set of a fixed size chosen uniformly at random from $1A^{l-2}1\ A^n$, and let D_2 be a set of a fixed size chosen uniformly at random from CA^{l-1}. For any $b \in A$, if $2l-k \neq 1$, then

- (i) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}bA^{l-2}C\right)]$
  $= \mathcal{K}(C)/8$
- (ii) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}GA^{l-1}\right)]$
  $= (\frac{1}{4} - \mathcal{K}(C))/4$
- (iii) $E[\mathcal{K}\left(D_1 A^{k-l} \cap A^{k-l}bA^{l-2}C\right)]$
  $= \mathcal{K}(C)\mathcal{K}(D_1)/2$
- (iv) $E[\mathcal{K}\left(D_1 A^{k-l} \cap A^{k-l}GA^{l-1}\right)]$
  $= \mathcal{K}(D_1)(\frac{1}{4} - \mathcal{K}(C))$.

If $2l-k=1$, then

- (v) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}0\ A^{l-2}C\right)]$
  $= \mathcal{K}(C)/4$
- (vi) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}GA^{l-1}\right)]$
  $= (\frac{1}{4} - \mathcal{K}(C))/2$
- (vii) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}D_2\right)]$
  $= \mathcal{K}(D_2)/2$
- (viii) $E[\mathcal{K}\left(D_2A^{k-l} \cap A^{k-l}1A^{l-2}C\right)]$
  $= \mathcal{K}(C)\mathcal{K}(D_2)$.

Proof of Theorem VIII.1(b): Here we assume $\lambda_2 < 2\lambda_1$ (or equivalently $n > l-2$ by (1)) and $\frac{1}{4} \leq \mu_1 2^{-\lambda_1} \leq \frac{1}{2}$.

Let C be a set of size $2^n - \mu_1$ chosen uniformly at random from among the 2^{n-1} length-λ_1 elements of $0A^{l-2}1\ A^{n-l+1}$. Since $\frac{1}{4} \leq \mu_1 2^{-\lambda_1} \leq \frac{1}{2}$, we have $2^{n-1} \leq \mu_1 \leq 2^n$, which implies $0 \leq |C| \leq 2^{n-1}$, so there are enough words from which to choose C.

Define the following (random) set:

$$F_1 = 0A^n - C.$$

- For Overlap Cases 1 and 3 below:
  Let D be a set of size $(\frac{3}{4} - \mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$ chosen uniformly at random from among the 2^{n+l-2} length-λ_2 elements of $1A^{l-2}1A^n$, and let
  $$F_2 = (1A^{l-2}1A^n - D) \cup CA^{l-1}.$$
- For Overlap Case 2 below:
  Let D be a set of size $(\frac{3}{4} - \mu_1 2^{-\lambda_1})2^{\lambda_2} - \mu_2$ chosen uniformly at random from among the $2^{l-1} \cdot |C|$ length-λ_2 elements of CA^{l-1}, and let
  $$F_2 = 1A^{l-2}1A^n \cup (CA^{l-1} - D).$$

Then

$$\mathcal{K}(D) = \left(\frac{3}{4} - \mu_1 2^{-\lambda_1}\right) - \mu_2 2^{-\lambda_2}, \qquad (32)$$

so

$$0 \leq \mathcal{K}(D) \leq \frac{3}{4} - \mu_1 2^{-\lambda_1} - \frac{1}{2} + \frac{\mu_1 2^{-\lambda_1}}{2} = \frac{1}{4} - \frac{\mu_1 2^{-\lambda_1}}{2}.$$

This means there are enough words from which to choose D, since

$$\frac{1}{4} - \frac{\mu_1 2^{-\lambda_1}}{2} \leq \frac{1}{4} - \frac{1}{8} = \frac{1}{8} < \frac{1}{4} = \mathcal{K}\left(1A^{l-2}1A^n\right)$$

and

$$\begin{aligned}
\frac{1}{4} - \frac{\mu_1 2^{-\lambda_1}}{2} &= \frac{1}{2}\left(\frac{1}{2} - \mu_1 2^{-\lambda_1}\right)\\
&\leq \frac{1}{2} - \mu_1 2^{-\lambda_1}\\
&= \mathcal{K}(C) = \mathcal{K}\left(CA^{l-1}\right).
\end{aligned}$$

Note that the words in F_2 all have first bit equal to 1 or prefix in C, and a 1 in the $(n+1)$th position from the right. These conditions guarantee that no word in F_1 is a prefix or suffix of a word in F_2.

Also, in all 3 cases,

$$\begin{aligned}
|F_1| &= |0A^n| - |C| = 2^n - (2^n - \mu_1) = \mu_1\\
|F_2| &= |1A^{l-2}1\ A^n| + |CA^{l-1}| - |D|\\
&= 2^{n+l-2} + 2^{l-1}(2^n - \mu_1)\\
&\quad - \left(\frac{3}{4} - \mu_1 2^{-\lambda_1}\right) 2^{\lambda_2} + \mu_2\\
&= \mu_2
\end{aligned}$$

so the set F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2.

In each case, we will also construct a third random set F_3, consisting of μ_3 words, each of length λ_3. The random set

$$F = F_1 \cup F_2 \cup F_3$$

on average meets the requirements of the desired fix-free code. The union of at least one non-random instance for each of F_1, F_2, and F_3 will then yield the asserted fix-free code.

- Overlap Case 1: $2l - k < 1$.
Let $G = 0A^{l-2}1A^{n-l+1} - C$, so that $E[\mathcal{K}(G)] = \frac{1}{4} - \mathcal{K}(C)$. Let

$$Y_{i,j} = \begin{cases} 1A^{l-2}0A^{n+k-l} & \text{if } i{=}1,\ j{=}0 \\ DA^{k-l} & \text{if } i{=}j{=}1 \end{cases}$$

$$W_{i,j} = \begin{cases} A^{k-l}iA^{l-2}C & \text{if } j{=}0 \\ A^{k-l}GA^{l-1} & \text{if } i{=}0,\ j{=}1 \\ A^{k-l}D & \text{if } i{=}j{=}1. \end{cases}$$

Let $\mathcal{I} = A^4 - 0A^3$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{1,Z_2} \cap W_{Z_3,Z_4})$$

where

$$Y_{1,Z_2} \cap W_{Z_3,Z_4} \subseteq 1\ A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n.$$

Here F_3 is comprised of only 8 of the 16 possible sets obtained from the pattern

$$Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n,$$

namely by excluding patterns with $Z_1 = 0$ from the union. One can verify that no words in F_1 or F_2 can be either prefixes or suffixes of any words in F_3.

The expected Kraft sum of F_3 is then lower bounded as follows:

$$\begin{aligned} &E[\mathcal{K}(F_3)] \\ &= \sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})] \\ &= \frac{\mathcal{K}(C)}{8} + \frac{\frac{1}{4} - \mathcal{K}(C)}{4} + \frac{\mathcal{K}(C)}{8} \\ &\quad + \frac{\mathcal{K}(D)}{4} + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} \\ &\quad + \mathcal{K}(D)\left(\frac{1}{4} - \mathcal{K}(C)\right) + \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} \\ &\quad + \mathcal{K}(D)^2 \qquad (33) \\ &= \mathcal{K}(D) + \left(\mathcal{K}(D) - \frac{1}{4}\right)^2 \\ &\geq \mathcal{K}(D) \\ &= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \qquad (34) \end{aligned}$$

where (33) follows from Lemma VI.1 when $Z_1 = Z_3 = Z_4 = 1$, and otherwise from Lemma VIII.3; and (34) follows from (32).

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 2: $2l - k = 1$.
Let $G = 0A^{l-2}1A^{n-l+1} - C$, so that $E[\mathcal{K}(G)] = \frac{1}{4} - \mathcal{K}(C)$. Let

$$Y_{i,j} = \begin{cases} DA^{k-l} & \text{if } i{=}0,\ j{=}1 \\ 1A^{l-2}0A^{n+k-l} & \text{if } i{=}1,\ j{=}0 \end{cases}$$

$$W_{i,j} = \begin{cases} A^{k-l}iA^{l-2}C & \text{if } j{=}0 \\ A^{k-l}(D \cup GA^{l-1}) & \text{if } i{=}0,\ j{=}1. \end{cases}$$

Let $\mathcal{I} = \{(0,1,0),(1,0,0),(1,0,1)\}$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})$$

where $Y_{Z_1,Z_2} \cap W_{Z_2,Z_3} \subseteq Z_1A^{l-2}Z_2A^{l-2}Z_3A^n$. In this case, F_3 is comprised of 3 of the 8 possible sets obtained from the pattern $Z_1A^{l-2}Z_2A^{l-2}Z_3A^n$.

The expected Kraft sum of F_3 can be lower bounded as follows:

$$\begin{aligned} &E[\mathcal{K}(F_3)] \\ &= \sum_{(Z_1,Z_2,Z_3)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})] \\ &= \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(C)}{4} + \frac{\mathcal{K}(D)}{2} + \frac{\frac{1}{4} - \mathcal{K}(C)}{2} \qquad (35) \\ &= \frac{(1 - 2\mathcal{K}(C))(1 - 4\mathcal{K}(D))}{8} + \mathcal{K}(D) \\ &\geq \mathcal{K}(D) \qquad (36) \\ &= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \qquad (37) \end{aligned}$$

where (35) follows from Lemma VIII.3; (36) follows since $\mathcal{K}(C) \leq \frac{1}{2}$ and $\mathcal{K}(D) \leq \frac{1}{4}$; and (37) follows from (32).

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 3: $2l - k > 1$.
This case follows from the same reasoning as in Overlap Case 1, except using $\geq$ in (33), since in this case (i.e., when $2l - k > 1$) Lemma VI.1 shows

$$\begin{aligned} E[\mathcal{K}(Y_{1,1} \cap W_{1,1})] &= E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\ &\geq \mathcal{K}(D)^2. \end{aligned}$$

■

C. Proof of Theorem VIII.1(c)

The proof of Theorem VIII.1(c) uses the following lemma, whose proof can be found in the appendix.

Lemma VIII.4: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$ and $n \geq l - 1$. Let C_0 be a set of a fixed size chosen uniformly at random from $0A^{l-2}0A^{n-l+1}$, and let $C = 0A^{l-2}1A^{n-l+1} \cup C_0$. For $i \in \{0, 1\}$, let D_i be a set of a fixed size chosen uniformly at random from $iA^{l-2}1\ A^n$. For any $b \in A$, if $2l - k \neq 1$, then

- (i) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}bA^{l-2}C\right)]$
  $= \mathcal{K}(C)/8$
- (ii) $E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}bA^{l-2}C\right)]$
  $= \mathcal{K}(C)(\mathcal{K}(C) - \frac{1}{4})/2$
- (iii) $E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}D_1\right)]$
  $= (\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(D_1)$
- (iv) $E[\mathcal{K}\left(D_1A^{k-l} \cap A^{k-l}bA^{l-2}C\right)]$
  $= \mathcal{K}(C)\,\mathcal{K}(D_1)/2.$

If $2l - k = 1$, then

- (v) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}0\ A^{l-2}C\right)]$
  $= \mathcal{K}(C)/4$
- (vi) $E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}C\right)]$
  $= \mathcal{K}(C)(\mathcal{K}(C) - \frac{1}{4})$
- (vii) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}D_0\right)]$
  $= \mathcal{K}(D_0)/2$
- (viii) $E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}D_0\right)]$
  $= 2(\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(D_0)$
- (ix) $E[\mathcal{K}\left(D_0A^{k-1} \cap A^{k-l}1A^{l-2}C\right)]$
  $= \mathcal{K}(C)\,\mathcal{K}(D_0).$

Proof of Theorem VIII.1(c): Here we assume $\lambda_2 < 2\lambda_1$ (or equivalently $n > l - 2$ by (1)) and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\frac{1}{4} \le \mu_2 2^{-\lambda_2} \le \frac{1}{2}$.

Let $C = C_1 \cup C_0$ be a set of size $2^n - \mu_1$, where $C_1 = 0\ A^{l-2}1\ A^{n-l+1}$ and C_0 is chosen uniformly at random from among the 2^{n-1} length-λ_1 elements of $0A^{l-2}0\ A^{n-l+1}$. Note that

$$|C_0| = 2^n - \mu_1 - 2^{n-1} = 2^{n-1} - \mu_1$$

and

$$\mathcal{K}(C_0) = \mathcal{K}(C) - \mathcal{K}(C_1) = \mathcal{K}(C) - \frac{1}{4}.$$

Since $0 \le \mu_1 2^{-\lambda_1} < \frac{1}{4}$, we have $0 \le \mu_1 \le 2^{n-1}$, which shows $0 \le |C_0| \le 2^{n-1} = |0A^{l-2}0\ A^{n-l+1}|$, and so there are enough words from which to choose C_0. Define the following (random) set:

$$F_1 = 0A^n - C.$$

- For Overlap Cases 1 and 3 below:
  Let D be a set of size $2^{\lambda_2 - 1} - \mu_2$ chosen uniformly at random from among the 2^{n+l-2} length-λ_2 elements of $1A^{l-2}1A^n$, and let
  $$F_2 = (1A^{l-2}1A^n - D) \cup 0A^{l-2}1A^n.$$
- For Overlap Case 2 below:
  Let D be a set of size $2^{\lambda_2 - 1} - \mu_2$ chosen uniformly at random from among the 2^{n+l-2} length-λ_2 elements of $0A^{l-2}1A^n$, and let
  $$F_2 = 1A^{l-2}1A^n \cup (0A^{l-2}1A^n - D).$$

Then $\mathcal{K}(D) = \frac{1}{2} - \mu_2 2^{-\lambda_2}$, so

$$0 \le \mathcal{K}(D) \le \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

This means there are enough words from which to choose D, since

$$\frac{1}{4} = \mathcal{K}\left(1A^{l-2}1A^n\right) = \mathcal{K}\left(0A^{l-2}1A^n\right).$$

Note that none of the words in F_2 start with a word from $0A^{l-2}0A^{n-l+1}$ or end with a word from $0A^n$, and so no word in $F_1 \subseteq 0A^{l-2}0A^{n-l+1}$ is a prefix or suffix of any word in F_2.

Also, in all 3 cases,

$$\begin{aligned}|F_1| &= |0A^n| - |C| = 2^n - (2^n - \mu_1) = \mu_1\\ |F_2| &= |1A^{l-2}1A^n| + |0A^{l-2}1A^n| - |D|\\ &= 2^{n+l-2} + 2^{n+l-2} - 2^{\lambda_2 - 1} + \mu_2\\ &= \mu_2\end{aligned}$$

so the set F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2.

In each case, we will also construct a third random set F_3, consisting of μ_3 words, each of length λ_3. The random set

$$F = F_1 \cup F_2 \cup F_3$$

on average forms the desired fix-free code. The union of non-random instances of F_1, F_2, and F_3 will then yield the asserted fix-free code.

- Overlap Case 1: $2l - k < 1$.
  Let
  $$Y_{i,j} = \begin{cases} C_0A^{k-1} & \text{if } i{=}j{=}0\\ 1A^{l-2}0A^{n+k-l} & \text{if } i{=}1,\ j{=}0\\ DA^{k-l} & \text{if } i{=}j{=}1\end{cases}$$
  $$W_{i,j} = \begin{cases} A^{k-l}iA^{l-2}C & \text{if } j{=}0\\ A^{k-l}D & \text{if } i{=}j{=}1.\end{cases}$$
  Let $\mathcal{I} = A^4 - (01A^2 \cup A^2 01)$ and define the set F_3 by:
  $$F_3 = \bigcup_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})$$
  where
  $$Y_{Z_1,Z_2} \cap W_{Z_3,Z_4} \subseteq Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n.$$
  Here F_3 is comprised of only 9 of the 16 possible sets obtained from the pattern
  $$Z_1A^{l-2}Z_2A^{k-2l}Z_3A^{l-2}Z_4A^n,$$
  namely by excluding patterns with $(Z_1, Z_2) = (0, 1)$ or $(Z_3, Z_4) = (0, 1)$ from the union. One can verify that no words in F_1 or F_2 can be either prefixes or suffixes of any words in F_3.
  The expected Kraft sum of F_3 is then lower bounded as follows:
  $$\begin{aligned}&E[\mathcal{K}(F_3)]\\ &= \sum_{(Z_1,Z_2,Z_3,Z_4)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_3,Z_4})]\\ &= \frac{\mathcal{K}(C)(\mathcal{K}(C) - \frac{1}{4})}{2} + \frac{\mathcal{K}(C)(\mathcal{K}(C) - \frac{1}{4})}{2}\\ &\quad + \mathcal{K}(D)\left(\mathcal{K}(C) - \frac{1}{4}\right) + \frac{\mathcal{K}(C)}{8}\\ &\quad + \frac{\mathcal{K}(C)}{8} + \frac{\mathcal{K}(D)}{4} + \frac{\mathcal{K}(C)\,\mathcal{K}(D)}{2}\end{aligned}$$

$$+ \frac{\mathcal{K}(C)\mathcal{K}(D)}{2} + \mathcal{K}(D)^2 \quad (38)$$
$$= \mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{4} + \left(\mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{2}\right)^2$$
$$\geq \mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{4}$$
$$= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \quad (39)$$

where (38) follows from Lemma VI.1 when $Z_1 = Z_3 = Z_4 = 1$, and otherwise from Lemma VIII.4; and (39) follows from the quantities $|C|$ and $|D|$ stated earlier in the proof.

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 2: $2l{-}k = 1$.
  Let

$$Y_{i,j} = \begin{cases} C_0 A^{k-1} & \text{if } i{=}j{=}0 \\ 1A^{l-2}0A^{n+k-l} & \text{if } i{=}1,\ j{=}0 \\ DA^{k-l} & \text{if } i{=}0,\ j{=}1 \end{cases}$$
$$W_{i,j} = \begin{cases} A^{k-l} i A^{l-2} C & \text{if } j{=}0 \\ A^{k-l} D & \text{if } i{=}0,\ j{=}1. \end{cases}$$

Let $\mathcal{I} = A^3 - (11A \cup A11)$ and define the set F_3 by:

$$F_3 = \bigcup_{(Z_1,Z_2,Z_3)\in\mathcal{I}} (Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})$$

where

$$Y_{Z_1,Z_2} \cap W_{Z_2,Z_3} \subseteq Z_1 A^{l-2} Z_2 A^{l-2} Z_3 A^n.$$

In this case, F_3 is comprised of only 5 of the 8 possible sets obtained from the pattern $Z_1 A^{l-2} Z_2 A^{l-2} Z_3 A^n$, namely excluding (Z_1, Z_2, Z_3) being $(1,1,0)$, $(0,1,1)$, or $(1,1,1)$.

The expected Kraft sum of F_3 can be lower bounded as follows:

$$E[\mathcal{K}(F_3)]$$
$$= \sum_{(Z_1,Z_2,Z_3)\in\mathcal{I}} E[\mathcal{K}(Y_{Z_1,Z_2} \cap W_{Z_2,Z_3})]$$
$$= \mathcal{K}(C)\left(\mathcal{K}(C) - \frac{1}{4}\right) + 2\mathcal{K}(D)\left(\mathcal{K}(C) - \frac{1}{4}\right)$$
$$+ \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(C)}{4} + \frac{\mathcal{K}(D)}{2} \quad (40)$$
$$= \mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{4} + \left(\mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{2}\right)^2$$
$$+ \mathcal{K}(D)(\mathcal{K}(C) - \mathcal{K}(D))$$
$$\geq \mathcal{K}(C) + \mathcal{K}(D) - \frac{1}{4} \quad (41)$$
$$= \frac{3}{4} - \mu_1 2^{-\lambda_1} - \mu_2 2^{-\lambda_2} \quad (42)$$

where (40) follows from Lemma VIII.4; (41) follows since $\mathcal{K}(D) \leq \frac{1}{4} \leq \mathcal{K}(C)$; and (42) follows from the quantities $|C|$ and $|D|$ stated earlier in the proof.

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 3: $2l{-}k > 1$.
  This case follows from the same reasoning as in Overlap Case 1, except using $\geq$ in (38), since in this case (i.e., when $2l - k > 1$) Lemma VI.1 shows

$$E[\mathcal{K}(Y_{1,1} \cap W_{1,1})] = E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)]$$
$$\geq \mathcal{K}(D)^2.$$

■

D. Proof of Theorem VIII.1(d)

The proof of Theorem VIII.1(d) uses the following lemma, whose proof can be found in the appendix.

Lemma VIII.5: Let $n, l, k \geq 1$ be integers such that $2 \leq l < k$ and $n \geq l - 1$. Let C_0 be a set of a fixed size chosen uniformly at random from $0A^{l-2}0A^{n-l+1}$, and let $C = 0A^{l-2}1A^{n-l+1} \cup C_0$. Let D be a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. Then

- (i) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-1}C\right)] = \mathcal{K}(C)/4$
- (ii) $E[\mathcal{K}\left(C_0 A^{k-1} \cap A^{k-1}C\right)]$
  $= \mathcal{K}(C)(\mathcal{K}(C) - \frac{1}{4})$
- (iii) $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-1}C\right)]$

$$= \begin{cases} \mathcal{K}(C)\mathcal{K}(D) \\ \quad \text{if } 2l-k<1 \\ 2(\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(D) \\ \quad \text{if } 2l-k=1 \\ \mathcal{K}(C)\mathcal{K}(D) \\ \quad \text{if } 2l-k>1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(D)(\mathcal{K}(C)-\frac{1}{4})(\frac{1}{2}-\mathcal{K}(C))}{\mathcal{K}(C)(2^{n-1}-1)} \\ \quad \text{if } 2l-k>1 \text{ and } (k-l) \mid (2l-k-1). \end{cases}$$

If $2l - k < 1$, then

- (iv) $E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}D\right)] = \mathcal{K}(D)/4$
- (v) $E[\mathcal{K}\left(C_0 A^{k-1} \cap A^{k-l}D\right)]$
  $= (\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(D)$.

Proof of Theorem VIII.1(d): Here we assume $\lambda_2 < 2\lambda_1$ (or equivalently $n > l - 2$ by (1)) and $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ and $\mu_2 2^{-\lambda_2} > \frac{1}{2}$.

Let $C = 0\ A^{l-2}1\ A^{n-l+1} \cup C_0$ be a set of size $2^n - \mu_1$, where C_0 is chosen uniformly at random from among the 2^{n-1} length-λ_1 elements of $0A^{l-2}0\ A^{n-l+1}$. Note that

$$|C_0| = 2^n - \mu_1 - 2^{n-1} = 2^{n-1} - \mu_1,$$

and thus $|C_0| \leq 2^{n-1}$, so there are enough words from which to choose C_0. Let D be a set of size $\mu_2 - 2^{\lambda_2 - 1}$ chosen uniformly at random from among the $2^{l-2} \cdot |C|$ length-λ_2 elements of $1A^{l-2}C$. Define the following (random) sets:

$$F_1 = 0A^n - C$$
$$F_2 = 1A^{l-2}1A^n \cup 0A^{l-2}1A^n \cup D.$$

Then $\mathcal{K}(D) = \mu_2 2^{-\lambda_2} - \frac{1}{2}$, so

$$\begin{aligned}0 \le \mathcal{K}(D) &\le \left(\frac{3}{4} - \mu_1 2^{-\lambda_1}\right) - \frac{1}{2}\\ &= \frac{1}{4} - \mu_1 2^{-\lambda_1}\\ &\le \frac{1}{2}\left(\frac{1}{2} - \mu_1 2^{-\lambda_1}\right)\\ &= \mathcal{K}\left(1A^{l-2}C\right)\end{aligned}$$

which means there are enough words from which to choose D. None of the words in F_2 start with a word from $0A^{l-2}0\ A^{n-l+1}$ or end with a word from F_1, and so no word in $F_1 \subseteq 0A^{l-2}0A^{n-l+1}$ is a prefix or suffix of any word in F_2. Also in all 3 cases,

$$\begin{aligned}|F_1| &= |0A^n| - |C| = 2^n - (2^n - \mu_1) = \mu_1\\ |F_2| &= |1A^{l-2}1A^n| + |0A^{l-2}1A^n| + |D|\\ &= 2^{n+l-2} + 2^{n+l-2} + \mu_2 - 2^{\lambda_2 - 1}\\ &= \mu_2\end{aligned}$$

so the set F_1 contains μ_1 words, each of length λ_1, and F_2 contains μ_2 words, each of length λ_2.

In each case, we will also construct a third random set F_3, consisting of μ_3 words, each of length λ_3. The random set

$$F = F_1 \cup F_2 \cup F_3$$

on average forms the desired fix-free code. The union of non-random instances of F_1, F_2, and F_3 will then yield the asserted fix-free code.

- Overlap Case 1: $2l - k < 1$.
  Define the following sets:

$$\begin{aligned}F_{3,1} &= 1\ A^{l-2}0\ A^{n+k-l} \cap A^{k-1}C\\ F_{3,2} &= C_0A^{k-1} \cap A^{k-1}C\\ F_{3,3} &= DA^{k-l} \cap A^{k-1}C\\ F_{3,4} &= 1\ A^{l-2}0\ A^{n+k-l} \cap A^{k-l}D\\ F_{3,5} &= C_0A^{k-1} \cap A^{k-l}D\\ F_3 &= (F_{3,1} \cup F_{3,2}) - (F_{3,3} \cup F_{3,4} \cup F_{3,5}).\end{aligned}$$

  Each set $F_{3,p}$ consists of words of length λ_3, and these sets are random, since they involve the random sets C or D. It is easy to verify that none of the words of F_1 (respectively, F_2) are prefixes or suffixes of any words in F_2 or F_3 (respectively, F_3), and that $F_{3,1}$ and $F_{3,2}$ are disjoint. Note that $F_{3,3} \cup F_{3,4} \cup F_{3,5}$ is the set of all words of $F_{3,1} \cup F_{3,2}$ that have some word of D as a prefix or suffix. We have

$$\begin{aligned}&E[\mathcal{K}(F_{3,1} \cup F_{3,2})]\\ &= E[\mathcal{K}\left(1A^{l-2}0\ A^{n+k-l} \cap A^{k-1}C\right)]\\ &\quad + E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-1}C\right)] \qquad (43)\\ &= \frac{\mathcal{K}(C)}{4} + \mathcal{K}(C_0)\mathcal{K}(C) \qquad (44)\\ &= \mathcal{K}(C)^2\\ &E[\mathcal{K}(F_{3,3} \cup F_{3,4} \cup F_{3,5})]\\ &\le E[\mathcal{K}\left(DA^{k-l} \cap A^{k-1}C\right)]\\ &\quad + E[\mathcal{K}\left(1A^{l-2}0\ A^{n+k-l} \cap A^{k-l}D\right)]\\ &\quad + E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}D\right)]\\ &= \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(D)}{4} + \mathcal{K}(C_0)\mathcal{K}(D) \qquad (45)\\ &= 2\mathcal{K}(C)\mathcal{K}(D)\\ &E[\mathcal{K}(F_3)]\\ &= E[\mathcal{K}(F_{3,1} \cup F_{3,2})] - E[\mathcal{K}(F_{3,3} \cup F_{3,4} \cup F_{3,5})] \qquad (46)\\ &\ge \mathcal{K}(C)^2 - 2\mathcal{K}(C)\mathcal{K}(D).\end{aligned}$$

$$\begin{aligned}&E[\mathcal{K}(F_1 \cup F_2 \cup F_3)]\\ &= E[\mathcal{K}(F_1)] + E[\mathcal{K}(F_2)] + E[\mathcal{K}(F_3)]\\ &= E[\mathcal{K}(0A^n - C)] + E[\mathcal{K}\left(1A^{l-2}1A^n\right)]\\ &\quad + E[\mathcal{K}\left(0A^{l-2}1A^n\right)] + E[\mathcal{K}(D)] + E[\mathcal{K}(F_3)]\\ &\ge \frac{1}{2} - \mathcal{K}(C) + \frac{1}{4} + \frac{1}{4} + \mathcal{K}(D) + \mathcal{K}(C)^2\\ &\quad - 2\mathcal{K}(C)\mathcal{K}(D) \qquad (47)\\ &= \frac{3}{4} + \left(\frac{1}{2} - \mathcal{K}(C)\right)\left(\frac{1}{2} - \mathcal{K}(C) + 2\mathcal{K}(D)\right)\\ &\ge \frac{3}{4} \qquad (48)\end{aligned}$$

  where (43) follows since $F_{3,1}$ and $F_{3,2}$ are disjoint; (44) and (45) follow from Lemma VIII.5; (46) follows from $F_{3,3} \cup F_{3,4} \cup F_{3,5} \subseteq F_{3,1} \cup F_{3,2}$; (47) follows from Lemma V.2; and (48) follows since $\mathcal{K}(C) \le \frac{1}{2}$ and $\mathcal{K}(D) \ge 0$.

  The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 2: $2l - k = 1$.
  Define the following sets:

$$\begin{aligned}F_{3,1} &= 1\ A^{l-2}0\ A^{n+l-1} \cap A^{k-1}C\\ F_{3,2} &= C_0A^{k-1} \cap A^{k-1}C\\ F_3 &= (F_{3,1} \cup F_{3,2}) - (DA^{k-l} \cap A^{k-1}C).\end{aligned}$$

  The sets $F_{3,1}$, $F_{3,2}$, and $(DA^{k-l} \cap A^{k-1}C)$ consist of words of length λ_3, and these sets are random since they involve the random sets C or D. It is easy to verify that none of the words of F_1 (respectively, F_2) are prefixes or suffixes of any words in F_2 or F_3 (respectively, F_3), and that $F_{3,1}$ and $F_{3,2}$ are disjoint. Also note that $DA^{k-l} \cap A^{k-1}C$ is the set of all words of $F_{3,1} \cup F_{3,2}$ that have some word of D as a prefix or suffix.

  Then we have

$$\begin{aligned}&E[\mathcal{K}(F_{3,1} \cup F_{3,2})]\\ &= E[1A^{l-2}0\ A^{n+l-1} \cap A^{k-1}C]\\ &\quad + E[C_0A^{k-1} \cap A^{k-1}C] \qquad (49)\\ &= \frac{\mathcal{K}(C)}{4} + \mathcal{K}(C_0)\mathcal{K}(C) = \mathcal{K}(C)^2 \qquad (50)\end{aligned}$$

$$E[\mathcal{K}\left(DA^{k-l} \cap A^{k-1}C\right)] = 2\left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(D) \qquad (51)$$

$$E[\mathcal{K}(F_3)] = \mathcal{K}(C)^2 - 2\left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(D) \quad (52)$$

$$\begin{aligned}
&E[\mathcal{K}(F_1 \cup F_2 \cup F_3)] \\
&= E[\mathcal{K}(F_1)] + E[\mathcal{K}(F_2)] + E[\mathcal{K}(F_3)] \\
&= E[\mathcal{K}(0A^n{-}C)] + E[\mathcal{K}(1A^{l-2}1A^n)] \\
&\quad + E[\mathcal{K}(0A^{l-2}1A^n)] + E[\mathcal{K}(D)] + E[\mathcal{K}(F_3)] \\
&= \left(\frac{1}{2} - \mathcal{K}(C)\right) + \frac{1}{4} + \frac{1}{4} + \mathcal{K}(D) + \mathcal{K}(C)^2 \\
&\quad - 2\left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(D) \qquad (53) \\
&= \frac{3}{4} + \left(\frac{1}{2} - \mathcal{K}(C)\right)^2 + 2\mathcal{K}(D)\left(\frac{3}{4} - \mathcal{K}(C)\right) \\
&\geq \frac{3}{4} \qquad (54)
\end{aligned}$$

where (49) follows since $F_{3,1}$ and $F_{3,2}$ are disjoint; (51) follows from Lemma V.11; (52) follows from $DA^{k-l} \cap A^{k-1}C \subseteq F_{3,1} \cup F_{3,2}$; (53) follows from Lemma V.2; and (54) follows since $\mathcal{K}(C) \leq \frac{1}{2} < \frac{3}{4}$.

The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1.

- Overlap Case 3: $2l - k > 1$.

If $n = 1$ then $\mu_1 2^{-\lambda_1} < \frac{1}{4}$ implies $\mu_1 2^{-\lambda_1} = 0$, in which case the proof is covered by Theorem III.1 since then there are codewords of only two distinct lengths λ_2 and λ_3. So assume $n \geq 2$.

Of the calculated expected values of the Kraft sums of $F_{3,1}$, $F_{3,2}$, $F_{3,3}$, $F_{3,4}$, and $F_{3,5}$ in Overlap Case 1, the only quantity that changes under the condition of Overlap Case 3 is $E[\mathcal{K}(F_{3,3})]$, as seen in Lemma V.11. In particular, since $2l - k > 1$ in this case, we have

$$\begin{aligned}
&E[\mathcal{K}(F_{3,3})] \\
&= E[\mathcal{K}(DA^{k-l} \cap A^{k-1}C)] \\
&\leq \mathcal{K}(C)\mathcal{K}(D) + \frac{\mathcal{K}(D)(\mathcal{K}(C) - \frac{1}{4})(\frac{1}{2} - \mathcal{K}(C))}{\mathcal{K}(C)(2^{n-1} - 1)}.
\end{aligned}$$

Therefore, in the calculation of $E[\mathcal{K}(F_1 \cup F_2 \cup F_3)]$, we get the lower bound

$$\begin{aligned}
&E[\mathcal{K}(F_1 \cup F_2 \cup F_3)] \\
&\geq \frac{3}{4} + \left(\frac{1}{2} - \mathcal{K}(C)\right)\left(\frac{1}{2} - \mathcal{K}(C) + 2\mathcal{K}(D)\right) \\
&\quad - \frac{\mathcal{K}(D)(\mathcal{K}(C) - \frac{1}{4})(\frac{1}{2} - \mathcal{K}(C))}{\mathcal{K}(C)(2^{n-1} - 1)} \\
&= \frac{3}{4} + \left(\frac{1}{2} - \mathcal{K}(C)\right)\left(\frac{1}{2} - \mathcal{K}(C)\right. \\
&\qquad \left. + 2\mathcal{K}(D) \cdot \frac{\mathcal{K}(C)(2^n - 3) + \frac{1}{4}}{\mathcal{K}(C)(2^n - 2)}\right) \\
&\geq \frac{3}{4} \qquad (55)
\end{aligned}$$

where (55) follows from $\mathcal{K}(C) \leq \frac{1}{2}$ and $2^n \geq 4$. The current case is then finished by applying the same reasoning used following (4) to the end of Part 1, Overlap Case 1. ■

Appendix
Proofs of Lemmas

Proof of Lemma V.1: Suppose a sequence of positive integers consists of $\mu_n > 0$ occurrences of integer l_n, for $1 \leq n \leq M$. Suppose its Kraft sum is less than $3/4$ and define

$$\mu'_n = \begin{cases} \mu_n & \text{if } 1 \leq n \leq M-1 \\ 3 \cdot 2^{l_M - 2} - \sum_{k=1}^{M-1} \mu_k 2^{l_M - k} & \text{if } n = M. \end{cases}$$

Note that

$$\mu'_M = \left(\frac{3}{4} - \sum_{k=1}^{M} \mu_k 2^{-k}\right) 2^{l_M} + \mu_M > \mu_M$$

and the new Kraft sum is

$$\begin{aligned}
\sum_{n=1}^{M} \mu'_n 2^{-l_n} &= \sum_{n=1}^{M-1} \mu_n 2^{-l_n} + \frac{3}{4} - \sum_{k=1}^{M} \mu_k 2^{-k} + \mu_M 2^{-l_M} \\
&= \frac{3}{4}.
\end{aligned}$$

If the sequence with multiplicities $\{\mu'_n\}$ has a fix-free code, then discarding any $\mu'_M - \mu_M$ codewords of length M yields a fix-free code for the sequence with multiplicities $\{\mu_n\}$. ■

Proof of Lemma V.3: Suppose $w \in R_l(U)$. Then the ith bit of the length-l prefix of w must be the same as the ith bit of the length-l suffix of w (which lies at position $i + m - l$). In other words, $w \in R_l(U)$ if and only if $w \in U$ and $w_i = w_{i+m-l}$ for all $1 \leq i \leq l$. The condition on w_i is equivalent to w_i being constant whenever i is congruent to p mod $(m - l)$, and $1 \leq i \leq m$, and $p \in \{1, \ldots, m - l\}$.

For any word $w \in R_l(U)$, the constant bit value w_i associated with each congruence class can be assigned independently of any other congruence class. Thus, the cardinality of $R_l(U)$ is equal to the product of the number N_p of allowable constant bit values for each congruence class. That is,

$$|R_l(U)| = \prod_{p=1}^{m-l} N_p.$$

Let $I_p = \{i \in \{1, \ldots, m\} \mid i \equiv p \bmod (m - l)\}$ be the set of positions in w that are in the pth congruence class. If $U_i = A$ for each $i \in I_p$, then $N_p = 2$, since any word $w \in R_l(U)$ could have either a 0 or 1 in the positions of I_p. If there exist $i, j \in I_p$ such that $U_i = 0$ and $U_j = 1$, then $N_p = 0$, since there is no way to label the positions in I_p with a constant bit value. Otherwise, $N_p = 1$, since then there exists at least one $i \in I_p$ such that $U_i \in \{0, 1\}$, and $U_j \in \{U_i, A\}$ for every other $j \in I_p$. Hence, N_p equals the cardinality of the intersection of the sets U_i taken over all $i \in I_p$. ■

Proof of Lemma V.4: For each $i \in \{1, 2\}$, let

$$g_i = |\{j : (X_i)_j = A\}|$$

be the number of positions in X_i that are not fixed points. Then for all $u \in U_1 \cap U_2$, by independence, we have

$$\begin{aligned}
&E\left[\mathcal{K}\left(W_1 \cap W_2\right)\right] \\
&= E\left[\sum_{u \in U_1 \cap U_2} 1_{W_1 \cap W_2}(u) 2^{-m}\right] \\
&= 2^{-m} \sum_{u \in U_1 \cap U_2} P(u \in (W_1 \cap W_2)) \\
&= 2^{-m} \sum_{u \in U_1 \cap U_2} P(u \in W_1) P(u \in W_2) \\
&= 2^{-m} \sum_{u \in U_1 \cap U_2} P(u \in A^a Y_1 A^b) P(u \in A^c Y_2 A^d) \\
&= 2^{-m} \sum_{u \in U_1 \cap U_2} \prod_{i=1}^{2} \frac{|Y_i| \cdot 2^{m-m_i}}{2^{g_i+m-m_i}} \\
&= \frac{|U_1 \cap U_2|}{2^m} \cdot \prod_{i=1}^{2} \frac{|Y_i|/2^m}{2^{g_i}/2^m} \\
&= \mathcal{K}\left(U_1 \cap U_2\right) \prod_{i=1}^{2} \frac{\mathcal{K}\left(Y_i\right)}{\mathcal{K}\left(X_i\right)}. \qquad (56)
\end{aligned}$$

Let f_V denote the set of positions where V has a fixed point. Then $f_{U_1 \cap U_2} = f_{U_1} \cup f_{U_2}$, so using Lemma V.2,

$$\begin{aligned}
\mathcal{K}\left(U_1 \cap U_2\right) &= 2^{-|f_{U_1 \cap U_2}|} \\
&= 2^{-|f_{U_1} \cup f_{U_2}|} \\
&= 2^p \cdot 2^{-|f_{U_1}|} 2^{-|f_{U_2}|} \\
&= 2^p \cdot \mathcal{K}\left(U_1\right) \mathcal{K}\left(U_2\right).
\end{aligned}$$

Combining this with (56) proves the lemma. ■

Proof of Corollary V.5: By Lemma V.4,

$$\begin{aligned}
&E[\mathcal{K}\left(A^a Y A^b \cap U\right)] \\
&= E[\mathcal{K}\left((A^a Y A^b \cap U) \cap (A^{n+k} \cap A^{n+k})\right)] \\
&= \mathcal{K}(U) \cdot \frac{\mathcal{K}(Y)}{\mathcal{K}(X)} \cdot \mathcal{K}\left(A^{n+k}\right) \cdot \frac{\mathcal{K}\left(A^{n+k}\right)}{\mathcal{K}\left(A^{n+k}\right)} \\
&= \mathcal{K}(U) \cdot \frac{\mathcal{K}(Y)}{\mathcal{K}(X)}.
\end{aligned}$$

■

Proof of Lemma V.6: If $|C| = 0$, then clearly the lemma holds. Suppose $|C| \geq 1$. If $u \in X$, then the probability that $u \in C$ is

$$\frac{\binom{|X|-1}{|C|-1}}{\binom{|X|}{|C|}} = \frac{|C|}{|X|}.$$

Now suppose $|C| \geq 2$. If $u, v \in X$ are distinct, then the probability that $u, v \in C$ is

$$\frac{\binom{|X|-2}{|C|-2}}{\binom{|X|}{|C|}} = \frac{|C|(|C|-1)}{|X|(|X|-1)}.$$

Finally, note that this last equation also fits the $|C| = 1$ case, since then the probability that such particular distinct u and v lie in C is zero, as $|C|$ contains only one element. ■

Proof of Lemma V.7: Let

$$X = CA^{p+1} \cap bUbA^n \cap A^{p+1}C.$$

By Lemma V.3, $|R_{n+1}(bUbA^n)| = |U|$, and so

$$|bUbA^n - R_{n+1}(bUbA^n)| = |U| \cdot 2^n - |U| = |U| \cdot (2^n - 1).$$

A word of $R_{n+1}(bUbA^n)$ is in X if its $(n+1)$-bit prefix (which is also its $(n+1)$-bit suffix) is selected during the construction of C, and a word of $bUbA^n - R_{n+1}(bUbA^n)$ is in X if the distinct $(n+1)$-bit prefix and suffix are both selected during the construction of C. Thus the expected number of words of $bUbA^n$ with a prefix and a suffix in C is

$$\begin{aligned}
&E[|CA^{p+1} \cap bUbA^n \cap A^{p+1}C|] \\
&= E\left[\sum_{v \in bUbA^n} 1_{CA^{p+1} \cap A^{p+1}C}(v)\right] \\
&= \sum_{v \in bUbA^n} P\{v \in CA^{p+1} \cap A^{p+1}C\} \\
&= \sum_{v \in bUbA^n} P\{\exists w \in C : v \in wA^{p+1} \cap A^{p+1}w\} \\
&\quad + \sum_{v \in bUbA^n} P\left\{v \in CA^{p+1} \cap A^{p+1}C,\ \nexists w \in C : \right. \\
&\qquad\qquad \left. v \in wA^{p+1} \cap A^{p+1}w\right\} \\
&= |U| \cdot \frac{|C|}{2^n} + |U| \cdot (2^n - 1) \frac{|C| \cdot (|C| - 1)}{2^n(2^n - 1)} \qquad (57) \\
&= |U| \cdot \frac{|C|^2}{2^n},
\end{aligned}$$

where (57) follows using Lemma V.6. These words all have length $p + 2 + n$, so their expected Kraft sum is

$$\begin{aligned}
&E[\mathcal{K}\left(CA^{p+1} \cap bUbA^n \cap A^{p+1}C\right)] \\
&= \frac{E[|CA^{p+1} \cap bUbA^n \cap A^{p+1}C|]}{2^{p+n+2}} \\
&= \frac{1}{2^{p+2+n}} \cdot |U| \cdot \frac{|C|^2}{2^n} \\
&= \frac{|U|}{2^p} \cdot \left(\frac{|C|}{2^{n+1}}\right)^2 \\
&= \mathcal{K}(U) \mathcal{K}(C)^2.
\end{aligned}$$

■

Proof of Lemma V.8: First suppose $n = 0 = l - 2$. Then either $D = \varnothing$ or $D = \{11\}$. Since $k \geq 3$, we have $2l - k \leq 1$. If $D = \varnothing$, then clearly $E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] = 0$, and so the lemma holds. If $D = \{11\}$ and $2l - k = 1$, then

$$\begin{aligned}
E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] &= E[\mathcal{K}(111)] \\
&= \frac{1}{8} \\
&= 2 \cdot \frac{1}{16} = 2\mathcal{K}(D)^2,
\end{aligned}$$

and if $2l - k < 1$, then

$$\begin{aligned}
E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] &= E[\mathcal{K}\left(11A^{k-2l}11\right)] \\
&= \frac{1}{16} \\
&= \mathcal{K}(D)^2,
\end{aligned}$$

by Lemma V.2. Thus the lemma holds when $n = 0 = l - 2$.

Now suppose $n > 0$ or $l > 2$, so that the set $1A^{l-2}1\ A^n$ from which we choose D has at least 2 elements. We consider three cases, depending on the value of $2l-k$. In each of the cases, we will define a particular pattern $X \subseteq \{0,1,A\}^{k+n}$ such that the randomly created set $DA^{k-l} \cap A^{k-l}D$ is a subset of the deterministic set X. For each such case, let $G_1 = R_{n+l}(X)$ and $G_2 = X - G_1$. Then $|G_2| = |X| - |G_1|$, since $G_1 \subseteq X$. Note that a word of G_1 is in $DA^{k-l} \cap A^{k-l}D$ if and only if the common $(n+l)$-bit prefix and suffix is in D, and a word of G_2 is in $DA^{k-l} \cap A^{k-l}D$ if and only if the distinct $(n+l)$-bit prefix and suffix are in D. Therefore,

$$\begin{aligned} &E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\ &= E[\mathcal{K}\left(DA^{k-l} \cap X \cap A^{k-l}D\right)] \\ &= E\left[\sum_{u \in X} 1_{DA^{k-l} \cap A^{k-l}D}(u) \cdot 2^{-(n+k)}\right] \\ &= \frac{1}{2^{n+k}}\left(\sum_{u \in G_1} P\{u \in DA^{k-l} \cap A^{k-l}D\}\right. \\ &\qquad \left. + \sum_{u \in G_2} P\{u \in DA^{k-l} \cap A^{k-l}D\}\right) \\ &= \frac{1}{2^{n+k}}\left(|G_1| \cdot \frac{|D|}{2^{n+l-2}} + |G_2| \cdot \frac{|D|(|D|-1)}{2^{n+l-2}(2^{n+l-2}-1)}\right) \quad (58) \\ &= \frac{|D|}{2^{2n+k+l-2}}\left(|G_1| + (|XA^n| - |G_1|) \cdot \frac{|D|-1}{2^{n+l-2}-1}\right), \quad (59) \end{aligned}$$

where (58) follows from Lemma V.6.

- Case 1: $2l-k<1$.
  Let $X = 1A^{l-2}1A^{k-2l}1A^{l-2}1A^n$. By Lemma V.3, $|G_1| = 2^{k-l-2}$. Therefore, from (59),
  $$\begin{aligned} &E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\ &= \frac{|D|}{2^{2n+k+l-2}} \\ &\quad \cdot \left(2^{k-l-2} + \frac{(|D|-1)\left(2^{k+n-4} - 2^{k-l-2}\right)}{2^{n+l-2}-1}\right) \\ &= \left(\frac{|D|}{2^{n+l}}\right)^2 = \mathcal{K}(D)^2. \end{aligned}$$
- Case 2: $2l-k=1$.
  Let $X = 1A^{l-2}1A^{l-2}1A^n$. By Lemma V.3, $|G_1| = 2^{l-2}$. Therefore, from (59),
  $$\begin{aligned} &E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\ &= \frac{|D|}{2^{2n+k+l-2}} \\ &\quad \cdot \left(2^{l-2} + \frac{(|D|-1)\left(2^{2l-4+n} - 2^{l-2}\right)}{2^{n+l-2}-1}\right) \\ &= 2\left(\frac{|D|}{2^{n+l}}\right)^2 = 2\mathcal{K}(D)^2. \end{aligned}$$
- Case 3: $2l-k>1$.
  Let
  $$\begin{aligned} a &= k-l-1 \\ b &= 2l-k-2 \\ X &= 1A^a 1A^b 1A^a 1\ A^n. \end{aligned}$$
  By Lemma V.3, $|G_1| = \beta 2^a$, where
  $$\beta = \begin{cases} 1 & \text{if } (a+1) \mid (b+1) \\ 1/2 & \text{if } (a+1) \nmid (b+1). \end{cases}$$
  Therefore, from (59),
  $$\begin{aligned} &E[\mathcal{K}\left(DA^{k-l} \cap A^{k-l}D\right)] \\ &= \frac{|D|}{2^{2n+k+l-2}} \\ &\quad \cdot \left(|G_1| + (2^{k+n-4} - |G_1|) \cdot \frac{|D|-1}{2^{n+l-2}-1}\right) \\ &= \left(\frac{|D|}{2^{n+l}}\right)^2 + \left(\frac{|D|}{2^{n+l}}\right)\frac{(2\beta-1)(\frac{1}{4} - |D|2^{-l-n})}{2^{n+l-2}-1} \\ &= \mathcal{K}(D)^2 + \frac{\mathcal{K}(D)\,(\frac{1}{4} - \mathcal{K}(D))(2\beta-1)}{2^{n+l-2}-1}. \end{aligned}$$
  ∎

Proof of Lemma V.10: First suppose D is a set of a fixed size chosen uniformly at random from $1A^{l-2}C$. Then given a word is in $1A^{l-2}CA^{k-l} \cap g(C)$, the probability that that word is in DA^{k-l} is the probability that the $(n+l)$-bit prefix is in D, which is

$$|D|/|1A^{l-2}C| = \mathcal{K}(D)/(\mathcal{K}(C)/2).$$

Therefore, letting

$$X = 1A^{l-2}CA^{k-l} \cap g(C)$$

(and noting that $DA^{k-l} \cap g(C) = DA^{k-l} \cap X$),

$$\begin{aligned} &E[\mathcal{K}\left(DA^{k-l} \cap g(C)\right)] \\ &= \frac{1}{2^{n+k}} E\left[\sum_{u \in A^{n+k}} 1_{DA^{k-l} \cap X}(u)\right] \\ &= \frac{1}{2^{n+k}} \sum_{u \in A^{n+k}} E\left[1_{DA^{k-l} \cap X}(u)\right] \\ &= \frac{1}{2^{n+k}} \sum_{u \in A^{n+k}} P(u \in DA^{k-l} \cap X) \\ &= \frac{1}{2^{n+k}} \sum_{u \in A^{n+k}} P(u \in DA^{k-l} \mid u \in X) P(u \in X) \\ &= \frac{\mathcal{K}(D)}{\mathcal{K}(C)/2} \cdot \frac{1}{2^{n+k}} \sum_{u \in A^{n+k}} P(u \in X) \\ &= \frac{\mathcal{K}(D)}{\mathcal{K}(C)/2} E[\mathcal{K}\left(1A^{l-2}CA^{k-l} \cap g(C)\right)]. \end{aligned}$$

The other cases follow similarly. ∎

Proof of Lemma V.11: For any C as in the Lemma statement, we have

$$\begin{aligned} &\mathcal{K}\left(CA^{k-l} \cap A^{k-l}C\right) \\ &= \mathcal{K}\left(C_1A^{k-l} \cap A^{k-l}C_1\right) + \mathcal{K}\left(C_1A^{k-l} \cap A^{k-l}C_0\right) \\ &\quad - \mathcal{K}\left(C_0A^{k-l} \cap A^{k-l}C_1\right) + \mathcal{K}\left(C_0A^{k-l} \cap A^{k-l}C_0\right). \end{aligned}$$

Using Lemma V.4,

$$\mathcal{K}\left(C_1 A^{k-l} \cap A^{k-l} C_1\right) = \begin{cases} 0 & \text{if } k = 2l-1 \\ \frac{1}{16} & \text{otherwise} \end{cases}$$

$$E[\mathcal{K}\left(C_0 A^{k-l} \cap A^{k-l} C_1\right)] = \begin{cases} \frac{\mathcal{K}(C_0)}{2} & \text{if } k = 2l-1 \\ \frac{\mathcal{K}(C_0)}{4} & \text{otherwise} \end{cases}$$

$$E[\mathcal{K}\left(C_1 A^{k-l} \cap A^{k-l} C_0\right)] = \begin{cases} 0 & \text{if } k = 2l-1 \\ \frac{\mathcal{K}(C_0)}{4} & \text{otherwise,} \end{cases}$$

and by Corollary V.9,

$$E[\mathcal{K}\left(C_0 A^{k-l} \cap A^{k-l} C_0\right)] = \begin{cases} \mathcal{K}(C_0)^2 & \text{if } 2l-k<1 \\ 2\mathcal{K}(C_0)^2 & \text{if } 2l-k=1 \\ \mathcal{K}(C_0)^2 & \text{if } 2l-k>1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(C_0)^2 + \frac{\mathcal{K}(C_0)(\frac{1}{4}-\mathcal{K}(C_0))}{2^{n-1}-1} & \text{if } 2l-k>1 \text{ and } (k-l) \mid (2l-k-1). \end{cases}$$

Thus

$$\begin{aligned} &E[\mathcal{K}\left(C A^{k-l} \cap A^{k-l} C\right)] \\ &= \begin{cases} (\mathcal{K}(C_0)+\frac{1}{4})^2 & \text{if } 2l-k<1 \\ \mathcal{K}(C_0)(2\mathcal{K}(C_0)+\frac{1}{2}) & \text{if } 2l-k=1 \\ (\mathcal{K}(C_0)+\frac{1}{4})^2 & \text{if } 2l-k>1 \text{ and } (k-l) \nmid (2l-k-1) \\ (\mathcal{K}(C_0)+\frac{1}{4})^2 + \frac{\mathcal{K}(C_0)(\frac{1}{4}-\mathcal{K}(C_0))}{2^{n-1}-1} & \text{if } 2l-k>1 \text{ and } (k-l) \mid (2l-k-1) \end{cases} \\ &= \begin{cases} \mathcal{K}(C)^2 & \text{if } 2l-k<1 \\ 2\mathcal{K}(C)(\mathcal{K}(C)-\frac{1}{4}) & \text{if } 2l-k=1 \\ \mathcal{K}(C)^2 & \text{if } 2l-k>1 \text{ and } (k-l) \nmid (2l-k-1) \\ \mathcal{K}(C)^2 + \frac{(\mathcal{K}(C)-\frac{1}{4})(\frac{1}{2}-\mathcal{K}(C))}{2^{n-1}-1} & \text{if } 2l-k>1 \text{ and } (k-l) \mid (2l-k-1). \end{cases} \end{aligned}$$

The lemma now follows using Lemma V.10 since

$$\begin{aligned} &E[\mathcal{K}\left(D A^{k-l} \cap A^{k-1} C\right)] \\ &= E[\mathcal{K}\left(D A^{k-l} \cap 1 A^{l-2} C A^{k-l} \cap A^{k-1} C\right)] \\ &= \frac{\mathcal{K}(D)}{\mathcal{K}(C)/2} \cdot E[\mathcal{K}\left(1 A^{l-2} C A^{k-l} \cap A^{k-1} C\right)] \\ &= \frac{\mathcal{K}(D)}{\mathcal{K}(C)} \cdot E[\mathcal{K}\left(C A^{k-l} \cap A^{k-l} C\right)] \quad (60) \end{aligned}$$

where (60) follows by Lemma V.2. ■

Proof of Lemma V.12: First suppose $n \le a$. Then

$$\begin{aligned} &E[\mathcal{K}\left(1A^a C A^{a+b+2} \cap 1\ A^a 0\ A^b 0\ A^a 1\ A^n \cap A^{a+b+2} C A^{a+1}\right)] \\ &= E[\mathcal{K}\left(1A^a (C A^{b+1} \cap 0\ A^b 0\ A^n \cap A^{b+1} C) A^{a-n} 1 A^n\right)] \\ &= \mathcal{K}(1A^a)\, E[\mathcal{K}\left(C A^{b+1} \cap 0\ A^b 0\ A^n \cap A^{b+1} C\right)] \\ &\quad \cdot \mathcal{K}\left(A^{a-n} 1 A^n\right) \quad (61) \\ &= \frac{1}{2} \cdot \mathcal{K}(C)^2 \mathcal{K}\left(A^b\right) \cdot \frac{1}{2} \cdot \quad (62) \\ &= \frac{\mathcal{K}(C)^2}{4} \quad (63) \end{aligned}$$

where (61) follows from Lemma V.2; (62) follows from Lemma V.2 and Lemma V.7; and (63) follows from Lemma V.2.

Now suppose $n > a$. By Lemma V.3,

$$|R_{n+1}(0A^b 0 A^a 1 A^{n-(a+1)})| = \begin{cases} 2^{b-1} & \text{if } (b+1) \nmid (a+1) \\ 0 & \text{otherwise} \end{cases}. \quad (64)$$

Let $X = R_{n+1}(0A^b 0 A^a 1 A^{n-(a+1)})$. If $(b+1) \nmid (a+1)$, then the expected Kraft sum is

$$\begin{aligned} &E[\mathcal{K}\left(1A^a C A^{a+b+2} \cap 1\ A^a 0\ A^b 0\ A^a 1\ A^n \cap A^{a+b+2} C A^{a+1}\right)] \\ &= E\Big[\mathcal{K}\Big(1A^a (C A^{b+1} \cap 0\ A^b 0\ A^a 1\ A^{n-(a+1)} \cap A^{b+1} C) A^{a+1}\Big)\Big] \quad (65) \\ &= \mathcal{K}(1A^a) \\ &\quad \cdot E\left[\mathcal{K}\left(C A^{b+1} \cap 0\ A^b 0\ A^a 1\ A^{n-(a+1)} \cap A^{b+1} C\right)\right] \\ &\quad \cdot \mathcal{K}\left(A^{a+1}\right) \\ &= \frac{1}{2} E\left[\mathcal{K}\left(C A^{b+1} \cap 0\ A^b 0\ A^a 1\ A^{n-(a+1)} \cap A^{b+1} C\right)\right] \quad (66) \\ &= \frac{1}{2}\left(\frac{|X|}{2^{n+b+2}} \cdot \frac{|C|}{2^n} + \left(\frac{1}{8} - \frac{|X|}{2^{n+b+2}}\right) \frac{|C|(|C|-1)}{2^n(2^n-1)}\right) \quad (67) \\ &= \frac{1}{2^{n+4}}\left(\frac{|C|}{2^n} + \frac{2^n-1}{2^n} \cdot \frac{|C|(|C|-1)}{2^n-1}\right) \quad (68) \\ &= \frac{|C|^2}{2^{n+4}} \\ &= \frac{\mathcal{K}(C)^2}{4} \quad (69) \end{aligned}$$

where (66) follows from Lemma V.2; (67) follows from the fact that $\mathcal{K}\left(0A^b 0 A^a 1 A^{n-(a+1)}\right) = 1/8$ (by Lemma V.2) and from Lemma V.6; (68) follows from (64).

On the other hand, if $(b+1) \mid (a+1)$, then following the same Kraft sum calculation as in (65)-(67) gives

$$\begin{aligned} &E[\mathcal{K}\left(1A^a C A^{a+b+2} \cap 1\ A^a 0\ A^b 0\ A^a 1\ A^n \cap A^{a+b+2} C A^{a+1}\right)] \\ &= \frac{1}{2}\left(\frac{0}{2^{n+b+2}} \cdot \frac{|C|}{2^n} + \left(\frac{1}{8} - \frac{0}{2^{n+b+2}}\right) \frac{|C|(|C|-1)}{2^n(2^n-1)}\right) \quad (70) \end{aligned}$$

$$= \frac{\mathcal{K}(C)^2}{4} - \frac{1}{2^{n+4}} \cdot \frac{|C|}{2^n} + \frac{1}{2^{n+4}} \cdot \frac{|C|(|C|-1)}{2^n(2^n-1)} \tag{71}$$
$$= \frac{\mathcal{K}(C)^2}{4} - \frac{|C|}{2^{2n+4}}\left(1 - \frac{|C|-1}{2^n-1}\right)$$
$$= \frac{\mathcal{K}(C)^2}{4} - \frac{\mathcal{K}(C)}{2^{n+3}} \cdot \frac{2^n - |C|}{2^n - 1}$$
$$= \frac{\mathcal{K}(C)^2}{4} - \frac{\mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right)}{4(2^n-1)}$$

where (70) follows from (64); and (71) follows from (68) and (69). ■

Proof of Lemma V.13: Let $X = 0A^a0A^b0A^a1A^n$. If $n \leq b$, then using Lemma V.3,

$$\begin{aligned}&|R_{n+1}(X)|\\ &= |R_{n+1}(0A^a0A^n)A^{b-n}0A^a1A^n|\\ &= |R_{n+1}(0A^a0A^n)| \cdot |A^{b-n}0A^a1A^n|\\ &= 2^a \cdot 2^{a+b} = 2^{2a+b}.\end{aligned}$$

If $b < n \leq a+b+1$, then using Lemma V.3,

$$\begin{aligned}&|R_{n+1}(X)|\\ &= |R_{n+1}(0A^a0A^b0A^{n-(a+b+1)})A^{a+b+1-n}1A^n|\\ &= |R_{n+1}(0A^a0A^b0A^{n-(a+b+1)})| \cdot |A^{a+b+1-n}1A^n|\\ &= \begin{cases} 2^{2a+b+1} & \text{if } (a+1) \mid (b+1) \\ 2^{2a+b} & \text{otherwise.}\end{cases}\end{aligned}$$

If $n > a+b+1$, then

$$\begin{aligned}&|R_{n+1}(X)|\\ &= |R_{n+1}(0A^a0A^b0A^a1A^{n-(a+b+2)})A^{a+b+2}|\\ &= |R_{n+1}(0A^a0A^b0A^a1A^{n-(a+b+2)})| \cdot |A^{a+b+2}|\\ &= 0,\end{aligned}$$

using Lemma V.3, since $X_{a+b+3} = 0$, $X_{2a+b+4} = 1$, and $a+b+3 \equiv (2a+b+4) \bmod (a+1)$.

Then using a similar probability calculation as in the proof of Lemma V.7, when $|R_{n+1}(X)| = 2^{2a+b}$ we have

$$\begin{aligned}&E[\mathcal{K}(X)]\\ &= E[\mathcal{K}(R_{n+1}(X))] \cdot \frac{|C|}{2^n}\\ &\quad + E[\mathcal{K}\left(A^{2a+b+n} - R_{n+1}(X)\right)] \cdot \frac{|C|(|C|-1)}{2^n(2^n-1)}\\ &= \frac{2^{2a+b}}{2^{2a+b+n+4}} \cdot \frac{|C|}{2^n}\\ &\quad + \left(\frac{2^{2a+b+n}}{2^{2a+b+n+4}} - \frac{2^{2a+b}}{2^{2a+b+n+4}}\right) \cdot \frac{|C|(|C|-1)}{2^n(2^n-1)}\\ &= \frac{1}{4} \cdot \frac{|C|^2}{2^{2(n+1)}}\\ &= \frac{\mathcal{K}(C)^2}{4}.\end{aligned}$$

Otherwise, when $|R_{n+1}(X)| = 2^{2a+b} + \beta 2^{2a+b}$ for $\beta \in \{-1, 1\}$, we have, using the previous calculation,

$$\begin{aligned}&E[\mathcal{K}(X)]\\ &= \frac{\mathcal{K}(C)^2}{4} + \beta\frac{2^{2a+b}}{2^{2a+b+n+4}}\left(\frac{|C|}{2^n} - \frac{|C|(|C|-1)}{2^n(2^n-1)}\right)\\ &= \frac{\mathcal{K}(C)^2}{4} + \beta\frac{1}{4(2^n-1)}\mathcal{K}(C)\left(\frac{1}{2} - \mathcal{K}(C)\right).\end{aligned}$$

■

Proof of Lemma V.14: Let

$$\begin{aligned}G_1 &= 1A^a1A^bR_{n+1}(0A^a0A^n)\\ G_2 &= 1A^a1A^b0A^a0A^n - G_1\\ H_1 &= R_{n+a+b+3}(1A^a1A^b0A^a0A^n)\\ H_2 &= 1A^a1A^b0A^a0A^n - H_1.\end{aligned}$$

Then $H_1 \subseteq G_1$ and $G_2 \subseteq H_2$. Then Lemma V.3 implies

$$\begin{aligned}|G_1 \cap H_1| = |H_1| &= \begin{cases} 2^{a-1} & \text{if } (a+1) \nmid (b+1) \\ 0 & \text{otherwise}\end{cases}\\ |G_1 \cap H_2| &= |G_1 - H_1|\\ &= \begin{cases} 2^{2a+b} - 2^{a-1} & \text{if } (a+1) \nmid (b+1) \\ 0 & \text{otherwise}\end{cases}\\ |G_2 \cap H_2| = |G_2| &= 2^{2a+b+n} - |G_1|\\ &= 2^{2a+b+n} - 2^{a+b}2^a\\ &= 2^{2a+b}(2^n - 1).\end{aligned}$$

Let $S = DA^{a+1} \cap 1A^a1A^b0A^a0A^n \cap A^{a+1}D$. If C is chosen uniformly at random from $0A^n$, and D is chosen uniformly at random from $1A^{a+b+1}C \subseteq 1A^{a+b+1}0A^n$, then for any word of length $n+l+a+1$, the probability it lies in $S \cap G_1 \cap H_1$ is

$$\frac{|C|}{2^n} \cdot \frac{|D|}{|C| \cdot 2^{a+b+1}},$$

the probability it lies in $S \cap G_1 \cap H_2$ is

$$\frac{|C|}{2^n} \cdot \frac{|D| \cdot (|D|-1)}{|C| \cdot 2^{a+b+1} \cdot (|C| \cdot 2^{a+b+1} - 1)},$$

and the probability it lies in $S \cap G_2 \cap H_2$ is

$$\frac{|C| \cdot (|C|-1)}{2^n(2^n-1)} \cdot \frac{|D| \cdot (|D|-1)}{|C| \cdot 2^{a+b+1} \cdot (|C| \cdot 2^{a+b+1} - 1)}.$$

Therefore, if $(a+1) \nmid (b+1)$, then

$$\begin{aligned}&E[\mathcal{K}(S)]\\ &= E[\mathcal{K}(S \cap G_1 \cap H_1)] + E[\mathcal{K}(S \cap G_1 \cap H_2)]\\ &\quad + E[\mathcal{K}(S \cap G_2 \cap H_2)]\\ &= \frac{2^{a-1}}{2^{n+2a+b+4}} \cdot \frac{|C|}{2^n} \cdot \frac{|D|}{|C| \cdot 2^{a+b+1}}\\ &\quad + \frac{2^{2a+b} - 2^{a-1}}{2^{n+2a+b+4}} \cdot \frac{|C|}{2^n}\\ &\quad \cdot \frac{|D| \cdot (|D|-1)}{|C| \cdot 2^{a+b+1}(|C| \cdot 2^{a+b+1} - 1)}\end{aligned}$$

$$
\begin{aligned}
&+ \frac{2^{2a+b}(2^n-1)}{2^{n+2a+b+4}} \cdot \frac{|C|\cdot(|C|-1)}{2^n(2^n-1)} \\
&\quad \cdot \frac{|D|\cdot(|D|-1)}{|C|\cdot 2^{a+b+1}(|C|\cdot 2^{a+b+1}-1)} \\
&= \mathcal{K}(D)^2 \qquad (72)
\end{aligned}
$$

where (72) follows from $\mathcal{K}(D) = \frac{|D|}{2^{n+a+b+3}}$.

On the other hand, if $(a+1) \mid (b+1)$, then

$$
\begin{aligned}
&E[\mathcal{K}(S)] \\
&= \mathcal{K}(D)^2 - \frac{2^{a-1}}{2^{n+2a+b+4}} \cdot \frac{|C|}{2^n} \cdot \frac{|D|}{|C|\cdot 2^{a+b+1}} \\
&\quad + \frac{2^{a-1}}{2^{n+2a+b+4}} \cdot \frac{|C|}{2^n} \\
&\quad \cdot \frac{|D|\cdot(|D|-1)}{|C|\cdot 2^{a+b+1}(|C|\cdot 2^{a+b+1}-1)} \\
&= \mathcal{K}(D)^2 - \frac{\mathcal{K}(D)}{|C|\cdot 2^{a+b+1}-1}\left(\frac{\mathcal{K}(C)}{2} - \mathcal{K}(D)\right).
\end{aligned}
$$

■

Proof of Lemma V.15: Let $r = 2^n$ and $s = 2^l$. Then

$$
\begin{aligned}
f(x,y) = y^2 &\left(\frac{xrs}{xrs-2}\right) \\
&- y\left(\frac{1}{2} - \frac{\frac{1}{2}-x}{2(r-1)} + \frac{x}{xrs-2}\right) \\
&+ \frac{1}{16} + \frac{x(\frac{1}{2}-x)}{2(r-1)}.
\end{aligned}
$$

Since

$$
\begin{aligned}
f\left(x, \frac{x}{2}\right) &= \frac{r}{4(r-1)}\left(x-\frac{1}{2}\right)\left(x-\frac{1}{2}+\frac{1}{2r}\right) \\
&\ge 0 \qquad \forall x \le \frac{1}{2} - \frac{1}{2r} \\
f\left(\frac{1}{2}, y\right) &= \left(\frac{rs}{rs-4}\right)\left(y-\frac{1}{4}\right)\left(y-\frac{1}{4}+\frac{1}{rs}\right) \\
&\ge 0 \qquad \forall y \le \frac{1}{4} - \frac{1}{rs}
\end{aligned}
$$

and $f(1/2, 1/4) = 0$, the lemma holds when $y = x/2$ and also when $x = 1/2$. Note that

$$
\begin{aligned}
f\left(x, \frac{x}{2} - \frac{1}{rs}\right) = &\frac{1}{4}\left(\frac{1}{2}-x\right)\left(\frac{1}{2}-x-\frac{x}{r-1}\right) \\
&+ \frac{\frac{1}{2}-x}{rs}\left(1-\frac{1}{2(r-1)}\right)
\end{aligned}
$$

which is 0 when $x = \frac{1}{2}$, and when $x \le \frac{1}{2} - \frac{1}{2r}$, satisfies

$$
\begin{aligned}
f\left(x, \frac{x}{2}-\frac{1}{rs}\right) &\ge \frac{1}{4}\cdot\frac{1}{2r}\left(\frac{1}{2r}-\frac{1}{2r}\right) + \frac{1}{2r^2 s}\left(1-\frac{1}{2}\right) \\
&> 0.
\end{aligned}
$$

Thus $f(x,y) \ge 0$ when $y = \frac{x}{2} - \frac{1}{rs}$ and $x \in [0, \frac{1}{2} - \frac{1}{2r}] \cup \{\frac{1}{2}\}$, i.e., the lemma holds when $y = \frac{x}{2} - \frac{1}{rs}$.

For all $x \in [\frac{1}{2r}, \frac{1}{2} - \frac{1}{2r}]$, since $xrs \ge (1/2r)rs \ge 4$, we have

$$
\begin{aligned}
\left.\frac{\partial f}{\partial y}\right|_{y=\frac{x}{2}-\frac{1}{rs}} &= -\left(\frac{1}{2}-x\right)\left(\frac{2r-3}{r-1}\right) - \frac{2x}{xrs-2} \\
&< 0.
\end{aligned}
$$

Thus, for any fixed $x \in [0, \frac{1}{2} - \frac{1}{2^{n+1}}]$, the function f is a convex parabola in y, which at $y = \frac{x}{2} - \frac{1}{rs}$ is both non-negative and has a negative slope, and is therefore non-negative for all $y \le \frac{x}{2} - \frac{1}{rs}$. ■

Proof of Lemma VI.1: For cases (i)–(ix), we will assume $2l - k \ne 1$. For these cases, the set

$$Z_1 A^{l-2} Z_2 A^{n+k-l} \cap A^{k-l} Z_3 A^{l-2} Z_4 A^n$$

has bits Z_2 and Z_3 in different positions, so it is either the pattern

$$Z_1 A^{l-2} Z_2 A^{k-2l} Z_3 A^{l-2} Z_4 A^n$$

(when $2l - k < 1$) or the pattern

$$Z_1 A^{k-l-1} Z_3 A^{2l-k-2} Z_2 A^{k-l-1} Z_4 A^n$$

(when $2l - k > 1$), which in both cases has exactly four fixed bits.

- (i) The set
$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n$$
is a pattern with exactly four fixed bits, and thus, by Lemma V.2, its Kraft sum is $1/16$.
- (ii),(iii) The set
$$CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n$$
equals $CA^{k-1} \cap 0U$ (where $U \in \{0,1,A\}^{n+k-1}$ is a pattern with exactly three fixed bits), and thus, by Corollary V.5, its expected Kraft sum is $\mathcal{K}(C)/8$. Similar reasoning proves case (iii).
- (iv) The set
$$CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}b_2A^{l-2}C$$
equals $CA^{k-1} \cap 0U0A^n \cap A^{k-1}C$ (where $U \in \{0,1,A\}^{k-2}$ is a pattern with exactly two fixed bits) and thus, by Lemma V.7, its expected Kraft sum is $\mathcal{K}(C)^2/4$.
- (v),(vi) The set
$$DA^{k-l} \cap A^{k-l}0A^{l-2}1A^n$$
equals $DA^{k-l} \cap U$, where $U \in \{0,1,A\}^{n+k}$ is either the pattern $1A^{l-2}1A^{k-2l}0A^{l-2}1A^n$ (when $2l-k<1$) or the pattern $1A^{k-l-1}0\ A^{2l-k-2}1\ A^{k-l-1}1\ A^n$ (when $2l-k>1$), both of which have exactly four fixed bits. Thus, by Corollary V.5, the set's expected Kraft sum is $\mathcal{K}(D)/4$. Similar reasoning proves case (vi).
- (vii),(viii) The set
$$CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}D,$$
by Lemma V.4, has expected Kraft sum is $\mathcal{K}(C)\mathcal{K}(D)/2$. Similar reasoning proves case (viii).
- (ix) This case follows immediately from Lemma V.8.

For cases (x)–(xvi), we will assume $2l - k = 1$. For these cases, the set

$$Z_1 A^{l-2} Z_2 A^{n+k-l} \cap A^{k-l} Z_3 A^{l-2} Z_4 A^n$$

is empty if $Z_2 \ne Z_3$, and otherwise is a pattern with exactly three fixed bits.

- (x) The set
$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n$$
is a pattern with exactly three fixed bits, and thus, by Lemma V.2. its Kraft sum is $1/8$.
- (xi),(xii) The set
$$CA^{k-1} \cap 0A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n$$
equals $CA^{k-1} \cap 0U$, (where $U \in \{0,1,A\}^{n+k-1}$ is a pattern with exactly two fixed bits), and thus, by Corollary V.5, its expected Kraft sum is $\mathcal{K}(C)/4$. Similar reasoning proves case (xii).
- (xiii) The set
$$CA^{k-1} \cap 0A^{l-2}b_1A^{n+k-l} \cap A^{k-l}b_1A^{l-2}C$$
equals $CA^{k-1} \cap 0U0A^n \cap A^{k-1}C$ (where $U \in \{0,1,A\}^{k-2}$ is a pattern with exactly one fixed bit), and thus, by Lemma V.7, its expected Kraft sum is $\mathcal{K}(C)^2/2$.
- (xiv),(xv) The set
$$CA^{k-1} \cap 0A^{l-2}1A^{n+k-l} \cap A^{k-l}D$$
equals $CA^{k-1} \cap A^{k-l}D$, and thus, by Lemma V.4, its expected Kraft sum is $\mathcal{K}(C)\mathcal{K}(D)$. Similar reasoning proves case (xv).
- (xvi) This case follows directly from Lemma V.8.

■

Proof of Lemma VII.1: The proof is similar to that of Lemma VI.1. For cases (i)–(iii), we will assume $2l-k \neq 1$.

- (i) The set
$$1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}1A^n$$
equals the set $A^{l-1}(0A^n - C)A^{k-l} \cap U$ (where $U \in \{0,1,A\}^{n+k}$ is a pattern with exactly four fixed bits), and thus, by Corollary V.5, its expected Kraft sum is $\mathcal{K}(0A^n - C)/8 = (\frac{1}{2} - \mathcal{K}(C))/8$, since $0A^n - C$ is chosen uniformly at random from $0A^n$ (by Lemma V.2).
- (ii) The expected Kraft sum of the set
$$\begin{aligned} &1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0\ A^{l-2}C \\ &= (1A^{l-2}0A^{n+k-l} \cap A^{k-l}0\ A^{l-2}C) \\ &\quad - (1A^{l-2}CA^{k-l} \cap A^{k-l}0\ A^{l-2}C) \end{aligned}$$
is $\frac{1}{8}\mathcal{K}(C) - \frac{1}{4}\mathcal{K}(C)^2 = \frac{1}{4}\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))$, since the expected Kraft sums of its two parts are $\frac{1}{8}\mathcal{K}(C)$ (by Lemma VI.1) and $\frac{1}{4}\mathcal{K}(C)^2$ (by Lemma V.7).
- (iii) The sets C and D_1 are chosen independently of each other, and the locations of the fixed bits of the sets from which they are drawn do not overlap (since $2l-1 \neq k$). Therefore, the expected Kraft sum of $1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}D_1$, by Lemma V.4, is $(1/2)(\frac{1}{2} - \mathcal{K}(C))\mathcal{K}(D_1)$, since the probability that it contains any particular word of length $n+k$ is the product of the probabilities that the word lies in each of the two intersected sets.

For cases (iv)–(ix), we will assume $2l-k=1$.

- (iv) The set
$$1A^{l-2}(0A^n - C)A^{k-l} \cap A^{k-l}0A^{l-2}1A^n$$
equals the set $A^{l-1}(0A^n - C)A^{k-l} \cap U$ (where $U \in \{0,1,A\}^{n+k}$ is a pattern with exactly three fixed bits), and thus, by Corollary V.5, its expected Kraft sum is $\frac{1}{4}\mathcal{K}(0A^n - C) = \frac{1}{4}(\frac{1}{2} - \mathcal{K}(C))$, since $0A^n - C$ is chosen uniformly at random from $0A^n$.
- (v) By Lemma VI.1,
$$\begin{aligned} &E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}1A^{n+k-l} \cap A^{k-l}1A^{l-2}C\right)] \\ &= \frac{\mathcal{K}(C)^2}{2}, \end{aligned}$$
and so by Lemma V.10, using
$$g(C) = CA^{k-1} \cap 0A^{l-2}1A^{n+k-l}$$
and
$$A^{k-l}D_2 \subseteq A^{k-l}1A^{l-2}C,$$
we get
$$\begin{aligned} &E[\mathcal{K}\left(CA^{k-1} \cap 0A^{l-2}1A^{n+k-l} \cap A^{k-l}D_2\right)] \\ &= \frac{\mathcal{K}(C)^2}{2} \cdot \frac{\mathcal{K}(D_2)}{\mathcal{K}(C)/2} \\ &= \mathcal{K}(C)\mathcal{K}(D_2). \end{aligned}$$
- (vi) The expected Kraft sum of the set
$$\begin{aligned} &1A^{l-2}(0A^n - C)A^{k-l} \\ &\quad \cap A^{k-l}0\ A^{l-2}C \\ &= 1A^{l-2}0A^{n+k-l} \\ &\quad \cap A^{k-l}0\ A^{l-2}C - 1A^{l-2}CA^{k-l} \cap A^{k-l}0\ A^{l-2}C. \end{aligned}$$
is $\frac{1}{4}\mathcal{K}(C) - \frac{1}{2}\mathcal{K}(C)^2 = \frac{1}{2}\mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))$, since the expected Kraft sums of its two parts are $\mathcal{K}(C)/4$ (by Lemma VI.1) and $\mathcal{K}(C)^2/2$ (by Lemma V.2 and Lemma V.7).
- (vii) Since $1A^{l-2}CA^{k-l} \cap A^{k-l}0A^{l-2}1A^n$ equals $A^{l-1}CA^{k-l} \cap U$ (where $U \in \{0,1,A\}^{n+k}$ is a pattern with exactly three fixed bits), by Corollary V.5. its expected Kraft is $\mathcal{K}(C)/4$. Then by Lemma V.10,
$$\begin{aligned} &E[\mathcal{K}\left(D_2A^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)] \\ &= \frac{(\mathcal{K}(C)/4)\mathcal{K}(D_2)}{\mathcal{K}(C)/2} \\ &= \mathcal{K}(D_2)/2. \end{aligned}$$
- (viii) By Lemma V.2 and Lemma V.7, we have
$$\begin{aligned} &E[\mathcal{K}\left(1A^{l-2}CA^{k-l} \cap A^{k-l}0A^{l-2}C\right)] \\ &= \mathcal{K}\left(1A^{l-2}\right) E[\mathcal{K}\left(CA^{k-l} \cap A^{l-1}C\right)] \\ &= \mathcal{K}(C)^2/2 \end{aligned}$$
so by Lemma V.10, the claimed expected Kraft sum is
$$\frac{(\mathcal{K}(C)^2/2)\mathcal{K}(D_2)}{\mathcal{K}(C)/2} = \mathcal{K}(C)\mathcal{K}(D_2).$$

■

Proof of Lemma VIII.2: The proof is similar to that of Lemma VI.1. For cases (i)–(iii), we will assume $2l-k<1$.

- (i) The set
$$\begin{aligned}&CA^{l-2-n}0A^{n+k-l} \cap A^{k-l}(0A^n{-}C)A^{l-2-n}1A^n\\&= CA^{l-2-n}0A^{k-2l}0A^{l-2}1A^n\\&\quad - CA^{l-2-n}0A^{k-2l}CA^{l-2-n}1A^n\end{aligned}$$
has expected Kraft sum
$$\frac{1}{8}\mathcal{K}(C) - \frac{1}{4}\mathcal{K}(C)^2 = \mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))/4,$$
since its first term has expected Kraft sum $\mathcal{K}(C)/8$ (by Corollary V.5) and its second term has expected Kraft sum $\mathcal{K}(C)^2/4$ (by Lemma V.2 and Lemma V.7).
- (ii) This case follows from Lemma V.12, since $0A^n{-}C$ is chosen uniformly at random from $0A^n$, and since $n \leq l-2$.
- (iii) Since $0A^n{-}C$ and D_1 are chosen independently and the fixed bits of $1A^{l-2}1A^{n+k-l}$ and $A^{k-l}0A^{l-2}1A^n$ do not overlap, Lemma V.4 implies the claimed expected Kraft sum is
$$\begin{aligned}&\mathcal{K}\left(1A^{l-2}1A^{n+k-l}\right) \cdot \frac{\mathcal{K}(D_1)}{\mathcal{K}(1A^{l-2}1A^n)}\\&\cdot \mathcal{K}\left(A^{k-l}0A^{l-2}1A^n\right) \cdot \frac{\mathcal{K}(0A^n{-}C)}{\mathcal{K}(0A^n)}\\&= \frac{\mathcal{K}(D_1)(\frac{1}{2} - \mathcal{K}(C))}{2}.\end{aligned}$$

For cases (iv)–(v), we will assume $2l{-}k = 1$.

- (iv) The set
$$\begin{aligned}&CA^{l-2-n}0A^{n+k-l} \cap A^{k-l}(0A^n{-}C)A^{l-2-n}1A^n\\&= CA^{l-2-n}0A^{l-2}1A^n - CA^{l-2-n}CA^{l-2-n}1A^n\end{aligned}$$
has expected Kraft sum
$$\frac{1}{4}\mathcal{K}(C) - \frac{1}{2}\mathcal{K}(C)^2 = \mathcal{K}(C)(\frac{1}{2} - \mathcal{K}(C))/2,$$
where its first term has expected Kraft sum $\mathcal{K}(C)/4$ (by Lemma VI.1) and its second term has expected Kraft sum $\mathcal{K}(C)^2/2$ (by Lemma V.2 and Lemma V.7).
- (v) The expected Kraft sum of
$$\begin{aligned}&1A^{l-2}(0A^n{-}C)A^{k-l} \cap A^{k-l}(0A^n{-}C)A^{l-2-n}1A^n\\&= 1A^{l-2}(0A^n{-}C)A^{l-2-n}1A^n\end{aligned}$$
is $(\frac{1}{2} - \mathcal{K}(C))/4$ by Lemma V.2.

For cases (vi)–(xii), we will assume $2l{-}k > 1$.

- (vi) Since
$$\mathcal{K}\left(CA^{l-2-n}0A^{n+k-l} \cap A^{k-l}1A^{l-2}C\right) = \frac{1}{4}\mathcal{K}(C)^2$$
by Lemma VI.1, the desired expected Kraft sum is
$$\frac{(\mathcal{K}(C)^2/4)\mathcal{K}(D_2)}{\mathcal{K}(C)/2} = \mathcal{K}(C)\mathcal{K}(D_2)/2.$$
- (vii) By Lemma V.10, the expected Kraft sum of the set
$$\begin{aligned}&1A^{l-2}(0A^n{-}C)A^{k-l} \cap A^{k-l}1A^{l-2}C\\&= 1\ A^{k-l-1}1\ A^{2l-k-2}0\ A^{k-l-1}C\\&\quad - 1\ A^{k-l-1}1\ A^{2l-k-2}\left(CA^{k-l} \cap 0\ A^{k-l-1}0\ A^n\right.\\&\qquad\qquad \left.\cap A^{k-l}C\right)\end{aligned}$$
is
$$\begin{aligned}&\frac{\mathcal{K}(D_2)}{\mathcal{K}(C)/2}\left(\frac{\mathcal{K}(C)}{8} - \frac{\mathcal{K}(C)^2}{4}\right)\\&= \frac{(\frac{1}{2} - \mathcal{K}(C))\mathcal{K}(D_2)}{2}\end{aligned}$$
since the expected Kraft sum of the first term is $\mathcal{K}(C)/8$ (by Lemma V.2) and the expected Kraft sum of the second term is $\mathcal{K}(C)^2/4$ (by Lemma V.2 and Lemma V.7).
- (viii) Lemma V.2 and Lemma V.7 imply
$$\begin{aligned}&\mathcal{K}\left(1A^{l-2-n}CA^{k-l} \cap A^{k-l}0A^{l-2}C\right)\\&= \mathcal{K}\left(1A^{k-l-1}0\ A^{2l-k-2}(CA^{k-l} \cap 0\ A^{k-l-1}0\ A^n\right.\\&\qquad\qquad \left.\cap A^{k-l}C)\right)\\&= \mathcal{K}(C)^2/4\end{aligned}$$
so the claimed expected Kraft sum is
$$\frac{1}{4}\mathcal{K}(C)^2\,\frac{\mathcal{K}(D_2)}{\mathcal{K}(C)/2} = \mathcal{K}(C)\mathcal{K}(D_2)/2.$$
- (ix) We have
$$\begin{aligned}&\mathcal{K}\left(CA^{l-2-n}0A^{n+k-l}\right.\\&\qquad \left.\cap A^{k-l}(0A^n{-}C)A^{l-2-n}1A^n\right)\\&= \mathcal{K}\left(CA^{l-2-n}0A^{n+k-l} \cap A^{k-l}0A^{l-2}1A^n\right)\\&\quad - \mathcal{K}\left(CA^{k-l} \cap 0\ A^{k-l-1}0\ A^{2l-k-2}0\ A^{k-l-1}1A^n\right.\\&\qquad \left.\cap A^{k-l}C\right)\\&= \frac{\mathcal{K}(C)}{8} - \frac{\mathcal{K}(C)^2}{4}\\&\quad - \begin{cases}\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } n \geq 2l-k-1\\ & \quad\text{and } (k-l) \mid (2l-k-1)\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$
by Lemma V.2 and Lemma V.13 (with $a = k-l-1$ and $b = 2l-k-2$).
- (x) This case follows immediately by Lemma V.12 (since $0A^n{-}C$ is drawn uniformly at random from $0A^n$), with $a = k-l-1$ and $b = 2l-k-2$.
- (xi) We have
$$\begin{aligned}&\mathcal{K}\left(1A^{l-2}CA^{k-l} \cap A^{k-l}(0A^n{-}C)A^{l-2-n}1A^n\right)\\&= \mathcal{K}\left(1A^{l-2}CA^{k-l} \cap A^{k-l}0A^{l-2}1A^n\right)\\&\quad - \mathcal{K}\left(1A^{l-2}CA^{k-l} \cap A^{k-l}CA^{l-2-n}1A^n\right)\\&= \frac{\mathcal{K}(C)}{8} - \frac{\mathcal{K}(C)^2}{4}\\&\quad + \begin{cases}\frac{\mathcal{K}(C)(\frac{1}{2}-\mathcal{K}(C))}{4(2^n-1)} & \text{if } n \geq k-l\\ & \quad\text{and } (2l{-}k{-}1) \mid (k{-}l)\\ 0 & \text{otherwise}\end{cases},\end{aligned}$$

so by Lemma V.10,

$$
\begin{aligned}
&E[\mathcal{K}\left(D_2A^{k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1A^n\right)] \\
&= \frac{\mathcal{K}(D_2)}{\mathcal{K}(C)/2} \\
&\quad \cdot \mathcal{K}\left(1A^{l-2}CA^{k-l} \cap A^{k-l}(0A^n - C)A^{l-2-n}1A^n\right) \\
&= \frac{(\frac{1}{2} - \mathcal{K}(C))\mathcal{K}(D_2)}{2} \\
&\quad + \begin{cases} \frac{\mathcal{K}(D_2)(\frac{1}{2} - \mathcal{K}(C))}{2(2^n - 1)} & \text{if } n \geq k-l \\ & \quad \text{and } (2l-k-1) \mid (k-l) \cdot \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

- (xii) This case follows directly from Lemma V.14.

■

Proof of Lemma VIII.3: The proof is similar to that of Lemma VI.1. For cases (i)–(iv), we will assume $2l - k \neq 1$.

- (i) The set

$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}bA^{l-2}C$$

equals the set $U0A^n \cap A^{k-1}C$ (where $U \in \{0, 1, A\}^{k-1}$ is a pattern with exactly three fixed bits), so by Lemma V.2, its expected Kraft sum is $\mathcal{K}(C)/8$.

- (ii) The claimed expected Kraft sum is

$$(1/4)\mathcal{K}(G) = \left(\frac{1}{4} - \mathcal{K}(C)\right)/4,$$

by Lemma V.4.

- (iii) The claimed expected Kraft sum is

$$\mathcal{K}(D_1) \cdot (1/4)\mathcal{K}(C)/(1/2) = \mathcal{K}(C)\mathcal{K}(D_1)/2,$$

by Lemma V.4.

- (iv) The claimed expected Kraft sum is

$$\mathcal{K}(D_1)\mathcal{K}(G) = \mathcal{K}(D_1)\left(\frac{1}{4} - \mathcal{K}(C)\right),$$

by Lemma V.4.

For cases (v)–(viii), we will assume $2l - k = 1$.

- (v) The set

$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}C$$

equals the set $1A^{l-2}0A^{l-2}C$, which has expected Kraft sum $\mathcal{K}(C)/4$ by Lemma V.2.

- (vi) The set

$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}GA^{l-2}$$

equals the set $1A^{l-2}GA^{l-2}$, which has expected Kraft sum $\mathcal{K}(G)/2 = (\frac{1}{4} - \mathcal{K}(C))/2$ by Lemma V.2.

- (vii) We have

$$
\begin{aligned}
&E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}CA^{l-1}\right)] \\
&= E[\mathcal{K}\left(1A^{l-2}CA^{l-1}\right)] = \mathcal{K}(C)/2
\end{aligned}
$$

by Lemma V.2, so the claimed expected Kraft sum is $(\mathcal{K}(D_2)/\mathcal{K}(C))(\mathcal{K}(C)/2) = \mathcal{K}(D_2)/2$, by Lemma V.10.

- (viii) We have

$$
\begin{aligned}
&R_{n+1}(0A^{l-2}1A^{l-2}0A^{l-2}1A^{n-(l-1)}) \\
&= 2^{2(l-2)} = 2^{2l-4} = 2^{k-3}
\end{aligned}
$$

by Lemma V.3. Therefore, using Lemma V.6, the expected Kraft sum of

$$
\begin{aligned}
&CA^{k-1} \cap A^{k-l}1A^{l-2}C \\
&\quad \subseteq 0A^{l-2}1A^{l-2}0A^{l-2}1A^{n-(l-1)}
\end{aligned}
$$

is

$$
\begin{aligned}
&\frac{2^{k-3}}{2^{n+k}} \cdot \frac{|C|}{2^{n-1}} + \left(\frac{1}{16} - \frac{2^{k-3}}{2^{n+k}}\right) \cdot \frac{|C|(|C|-1)}{2^{n-1}(2^{n-1}-1)} \\
&= \frac{|C|}{2^{2(n+1)}} + \frac{1}{8}\left(\frac{2^{n-1}-1}{2^n}\right)\frac{|C|(|C|-1)}{2^{n-1}(2^{n-1}-1)} \\
&= \frac{|C|^2}{2^{2(n+1)}} = \mathcal{K}(C)^2.
\end{aligned}
$$

Thus the claimed expected Kraft sum is

$$(\mathcal{K}(D_2)/\mathcal{K}(C))\mathcal{K}(C)^2 = \mathcal{K}(C)\mathcal{K}(D_2),$$

by Lemma V.10.

■

Proof of Lemma VIII.4: The proof is similar to that of Lemma VI.1. For cases (i)–(iv), we will assume $2l - k < 1$.

- (i) The set

$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}bA^{l-2}C$$

equals the set $U0A^n \cap A^{k-1}C$ (where $U \in \{0, 1, A\}^{k-1}$ is a pattern with exactly three fixed bits), so, by Lemma V.2, its expected Kraft sum is $\mathcal{K}(C)/8$.

- (ii) We have

$$
\begin{aligned}
&R_{n+1}(0A^{l-2}0A^{k-1} \cap A^{k-l}bA^{l-2}0A^{l-2}0A^{n-(l-1)}) \\
&= 2^{k-4}
\end{aligned}
$$

by Lemma V.3, since exactly 3 of the first $k-1$ positions in the set above are fixed bits. Therefore, using Lemma V.6,

$$
\begin{aligned}
&E[\mathcal{K}\left(C_0A^{k-l} \cap A^{k-1}bA^{l-2}C_0\right)] \\
&= \frac{2^{k-4}}{2^{n+k}}\frac{|C_0|}{2^{n-1}} + \left(\frac{1}{2^5} - \frac{2^{k-4}}{2^{n+k}}\right)\frac{|C_0|(|C_0|-1)}{2^{n-1}(2^{n-1}-1)} \\
&= \frac{|C_0|}{2^{2n+3}} + \frac{1}{16}\left(\frac{2^{n-1}-1}{2^n}\right)\frac{|C_0|(|C_0|-1)}{2^{n-1}(2^{n-1}-1)} \\
&= \frac{1}{2} \cdot \frac{|C_0|^2}{2^{2(n+1)}} \\
&= \frac{\mathcal{K}(C_0)^2}{2}.
\end{aligned}
$$

Corollary V.5 then implies

$$\begin{aligned}
&E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}bA^{l-2}C\right)]\\
&= E\left[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}bA^{l-2}0A^{l-2}1A^{n-(l-1)}\right)\right]\\
&\quad + E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}bA^{l-2}C_0\right)]\\
&= \frac{\mathcal{K}(C_0)}{8} + \frac{\mathcal{K}(C_0)^2}{2}\\
&= \frac{\mathcal{K}(C_0)\left(\frac{1}{4} + \mathcal{K}(C_0)\right)}{2}\\
&= \frac{\mathcal{K}(C)\left(\frac{1}{4} - \mathcal{K}(C)\right)}{2}.
\end{aligned}$$

- (iii) By Lemma V.4, the claimed expected Kraft sum is

$$\mathcal{K}(C_0)\mathcal{K}(D_1) = \left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(D_1).$$

- (iv) By Lemma V.4, the claimed expected Kraft sum is

$$\mathcal{K}(D_1) \cdot (1/4)\mathcal{K}(C)/(1/2) = \mathcal{K}(C)\mathcal{K}(D_1)/2.$$

For cases (v)–(ix), we will assume $2l{-}k = 1$.

- (v) The set

$$1A^{l-2}0A^{n+k-l} \cap A^{k-l}0A^{l-2}C$$

equals the set $1A^{l-2}0A^{l-2}C$, so the claimed expected Kraft sum is $\mathcal{K}(C)/4$ by Lemma V.2.

- (vi) We have

$$\begin{aligned}
&E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}C\right)]\\
&= E\left[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}0A^{l-1}1A^{n-(l-1)}\right)\right]\\
&\quad + E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}C_0\right)].
\end{aligned}$$

The first expected Kraft sum on the right equals $\mathcal{K}(C_0)/4$ by Corollary V.5, and the second expected Kraft sum on the right equals $\mathcal{K}(C_0)^2$ by Corollary V.9. Thus the claimed expected Kraft sum is

$$\begin{aligned}
\mathcal{K}(C_0)/4 + \mathcal{K}(C_0)^2 &= \mathcal{K}(C_0)\left(\frac{1}{4} + \mathcal{K}(C_0)\right)\\
&= \left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(C).
\end{aligned}$$

- (vii) The set $1A^{l-2}0A^{n+k-l} \cap A^{k-l}D_0$ equals the set $1A^{l-2}D_0$, and so the claimed expected Kraft sum is $\mathcal{K}(D_0)/2$ by Lemma V.2.
- (viii) By Lemma V.4, the claimed expected Kraft sum is

$$2\mathcal{K}(C_0)\mathcal{K}(D_0) = 2\left(\mathcal{K}(C) - \frac{1}{4}\right)\mathcal{K}(D_0).$$

- (ix) We have

$$\begin{aligned}
&E[\mathcal{K}\left(D_0A^{k-l} \cap A^{k-l}1A^{l-2}C\right)]\\
&= E[\mathcal{K}\left(D_0A^{k-l} \cap A^{k-l}1A^{l-2}C_0\right)]\\
&\quad + E\left[\mathcal{K}\left(D_0A^{k-l} \cap A^{k-l}1A^{l-2}0A^{l-2}1A^{n-(l-1)}\right)\right].
\end{aligned}$$

The first expected Kraft sum on the right equals

$$2\mathcal{K}(D_0) \cdot (1/8)\mathcal{K}(C_0)/(1/4) = \mathcal{K}(C_0)\mathcal{K}(D_0)$$

by Lemma V.4, and the second expected Kraft sum on the right equals $2\mathcal{K}(D_0)(1/8) = \mathcal{K}(D_0)/4$ by Lemma V.4. Thus the claimed expected Kraft sum is

$$\begin{aligned}
&\mathcal{K}(C_0)\mathcal{K}(D_0) + \frac{\mathcal{K}(D_0)}{4}\\
&= (\mathcal{K}(C) - 1/4)\mathcal{K}(D_0) + \frac{\mathcal{K}(D_0)}{4}\\
&= \mathcal{K}(C)\mathcal{K}(D_0).
\end{aligned}$$

■

Proof of Lemma VIII.5: The proof is similar to that of Lemma VI.1.

- (i) We have

$$1A^{l-2}0A^{n+k-l} \cap A^{k-1}C = 1A^{l-2}0A^{k-l-1}C,$$

so its expected Kraft sum is $\mathcal{K}(C)/4$ by Lemma V.2.

- (ii) If $2l - k \neq 1$, then

$$\begin{aligned}
&E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-1}C\right)]\\
&= E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}C\right)]\\
&\quad + E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}1A^{l-2}C\right)]\\
&= \mathcal{K}(C)\left(\mathcal{K}(C) - \frac{1}{4}\right)
\end{aligned}$$

by Lemma VIII.4. If $2l - k = 1$, then

$$\begin{aligned}
&E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-1}C\right)]\\
&= E[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}0A^{l-2}C\right)]\\
&= \mathcal{K}(C)\left(\mathcal{K}(C) - \frac{1}{4}\right)
\end{aligned}$$

by Lemma VIII.4.

- (iii) This case follows directly from Lemma V.11.

For cases (iv) and (v), we will assume $2l{-}k < 1$.

- (iv) We have

$$\begin{aligned}
&E[\mathcal{K}\left(1A^{l-2}0A^{n+k-l} \cap A^{k-l}1A^{l-2}C\right)]\\
&= E[\mathcal{K}\left(1A^{l-2}0A^{k-2l}1A^{l-2}C\right)] = \mathcal{K}(C)/8
\end{aligned}$$

by Lemma V.2. Then by Lemma V.10, the claimed expected Kraft sum is

$$(\mathcal{K}(D)/(\mathcal{K}(C)/2))(\mathcal{K}(C)/8) = \mathcal{K}(D)/4.$$

- (v) We have

$$\begin{aligned}
&E\left[\mathcal{K}\left(C_0A^{k-1} \cap A^{k-l}1A^{l-2}C\right)\right]\\
&= E\left[\mathcal{K}\left(C_0A^{k-2l}1A^{l-2}0A^{l-2}1A^{n-(l-1)}\right)\right]\\
&\quad + E[\mathcal{K}\left(C_0A^{k-2l}1A^{l-2}C_0\right)]\\
&= \frac{\mathcal{K}(C_0)}{8} + \frac{\mathcal{K}(C_0)^2}{2}\\
&= \frac{(\mathcal{K}(C) - \frac{1}{4})\mathcal{K}(C)}{2}
\end{aligned}$$

by Lemma V.2 and Corollary V.9, and using the fact that $C_0A^{k-2l}1A^{l-2}C_0$ contains exactly half of the words of $C_0 A^{k-l-1} C_0$. Then by Lemma V.10, the claimed expected Kraft sum is

$$\begin{aligned}(\mathcal{K}(D)/(\mathcal{K}(C)/2))\left(\left(\mathcal{K}(C)-\frac{1}{4}\right)\mathcal{K}(C)/2\right)\\ =\left(\mathcal{K}(C)-\frac{1}{4}\right)\mathcal{K}(D).\end{aligned}$$

■

ACKNOWLEDGMENT

The authors would like to thank Sergey Yekhanin for his unpublished notes [78], and Matthew Ekaireb for helpful computer simulations.

REFERENCES

[1] N. Abedini, S. P. Khatri, and S. A. Savari, “A SAT-based scheme to determine optimal fix-free codes,” in *Proc. Data Compress. Conf.*, 2010, pp. 169–178.

[2] A. Aghajan and M. Khosravifard, “93% of the $\frac{3}{4}$-conjecture is already verified,” *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8182–8194, Dec. 2013.

[3] A. Aghajan and M. Khosravifard, “Weakly symmetric fix-free codes,” *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5500–5515, Sep. 2014.

[4] R. Ahlswede, B. Balkenhol, and L. Khachatrian, “Some properties of fix-free codes,” in *Proc. 1st Intas Seminar Coding Theory Combinatorics*, Thahkadzor, Armenia, pp. 20–33, 1996.

[5] R. Bauer and J. Hagenauer, “Iterative source/channel-decoding using reversible variable length codes,” in *Proc. Data Compress. Conf.*, 2000, pp. 93–102.

[6] M.-P. Béal, J. Berstel, B. H. Marcus, D. Perrin, C. Reutenauer, and P. Siegel, “Variable-length codes and finiteautomata,” in *Selected Topics in Information and Coding Theory* (Series on Coding Theory and Cryptology). Singapore: World Scientific, pp. 505–584, 2010.

[7] J. Berstel and D. Perrin, *Theory of Codes*. New York, NY, USA: Academic, 1985.

[8] J. Berstel, C. De Felice, D. Perrin, C. Reutenauer, and G. Rindone, “Bifix codes and sturmian words,” *J. Algebra*, vol. 369, pp. 146–202, Nov. 2012.

[9] V. Berthé et al., “Acyclic, connected and tree sets,” *Monatshefte für Math.*, vol. 176, pp. 521–550, 2015.

[10] V. Berthé et al., “The finite index basis property,” *J. Pure Appl. Algebra*, vol. 219, pp. 2521–2537, Jul. 2015.

[11] V. Berthé et al., “Bifix codes and interval exchanges,” *J. Pure Appl. Algebra*, vol. 219, pp. 2781–2798, Jul. 2015.

[12] V. Berthé et al., “Maximal bifix decoding,” *Discrete Math.*, vol. 338, no. 5, pp. 725–742, 2015.

[13] M. Bodewig, “Multiplied complete fix-free codes and shiftings regarding the $\frac{3}{4}$-conjecture,” in *Information Theory, Combinatorics, and Search Theory* (Lecture Notes in Computer Science), vol. 7777, R. Ahlswede, Ed. Berlin, Germany: Springer, 2013, pp. 694–710.

[14] M. Bodewig, “An introduction of greedy extension sets for the application on fix-free codes,” Ph.D. thesis, Dept. Math., Aachen Univ., Aachen, Germany, 2015.

[15] M. Bystrom, S. Kaiser, and A. Kopansky, “Soft source decoding with applications,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 10, pp. 1108–1120, Oct. 2001.

[16] Y. Césari, “Propriétés combinatoires des codes biprefixes,” in *Théorie des Codes*, D. Perrin, ed. Paris, France: LITP, 1979, pp. 20–46.

[17] Y. Césari, “Sur un algorithme donnant les codes bipréfixes finis,” *Math. Syst. Theory*, vol. 6, pp. 221–225, Mar. 1982.

[18] T. Cover and J. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.

[19] C. Deppe and H. Schnettler, “On the $\frac{3}{4}$-conjecture for fix-free codes,” in *Proc. Eur. Conf. Combinatorics, Graph Theory Appl. (EuroComb)*, Berlin, Germany, 2005, pp. 111–116.

[20] C. Deppe and H. Schnettler, “On q-ary fix-free codes and directed De Bruijn graphs,” in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 1482–1485.

[21] A. S. Fraenkel, “Bidirectional Huffman coding,” *Comput. J.*, vol. 33, no. 4, pp. 296–307, Apr. 1990.

[22] S.-S. Gao and G.-F. Tu, “Robust H.263+ video transmission using partial backward decodable bit stream (PBDBS),” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 2, pp. 182–187, Feb. 2003.

[23] E. N. Gilbert and E. F. Moore, “Variable-length binary encodings,” *Bell Syst. Tech. J.*, vol. 38, no. 4, pp. 933–967, Jul. 1959.

[24] D. Gillman and R. L. Rivest, “Complete variable-length ‘fix-free’ codes,” *Des., Codes Cryptogr.*, vol. 5, no. 2, pp. 109–114, Mar. 1995.

[25] B. Girod, “Bidirectionally decodable streams of prefix code-words,” *IEEE Commun. Lett.*, vol. 3, no. 8, pp. 245–247, Aug. 1999.

[26] X. Guang, F. Fu, and L. Chen, “The existence and synchronization properties of symmetric fix-free codes,” *Sci. China Inf. Sci.*, vol. 56, no. 9, pp. 1–9, Sep. 2013.

[27] R. Gupta and R. Goel, “A necessary and sufficient condition for the existence of asymmetrical reversible VLCs,” *Int. J. Innov. Technol. Exploring Eng. (IJITEE)*, vol. 8, no. 4, pp. 314–317, Feb. 2019.

[28] K. Harada and K. Kobayashi, “A note on the fix-free property,” *IEICE Trans. Fundam.*, vol. 82, no. 10, pp. 2121–2128, Oct. 1999.

[29] D. A. Huffman, “A method for the construction of minimum-redundancy codes,” *Proc. IRE*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952.

[30] J.-Y. Huo, Y.-L. Chang, L.-H. Ma, and Z. Luo, “On constructing symmetrical reversible variable-length codes independent of the Huffman code,” *J. Zhejiang Univ.-Sci. A*, vol. 7, no. S1, pp. 59–62, Jan. 2006.

[31] *Information Technology—Coding of Audio/Visual Objects*, Final Draft International Standard, Part 2, Visual, Standard ISO/IEC 14496-2, Oct. 1998.

[32] ITU-T Recommendation H.263, *Video Coding for Low Bit Rate Communications*, Annex V, 2000.

[33] M. Javad-Kalbasi and M. Khosravifard, “Some tight lower bounds on the redundancy of optimal binary prefix-free and fix-free codes,” *IEEE Trans. Inf. Theory*, vol. 66, no. 7, pp. 4419–4430, Jul. 2020.

[34] W.-H. Jeong, Y.-S. Yoon, and Y.-S. Ho, “Design of reversible variable-length codes using properties of the Huffman code and average length function,” in *Proc. Int. Conf. Image Process. (ICIP)*, Singapore, vol. 2, Oct. 2004, pp. 817–820.

[35] S. Kaiser and M. Bystrom, “Soft decoding of variable-length codes,” in *Proc. IEEE Int. Conf. Commun. Global Converg. Through Commun. Conf. Rec. (ICC)*, New Orleans, LA, USA, Jun. 2000, pp. 1203–1207.

[36] A. Kakhbod, A. Nazari, and M. Zadimoghaddam, “Some notes on fix-free codes,” in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2008, pp. 1015–1018.

[37] A. Kakhbod and M. Zadimoghaddam, “On the construction of prefix-free and fix-free codes with specified codeword compositions,” *Discrete Appl. Math.*, vol. 159, no. 18, pp. 2269–2275, Dec. 2011.

[38] M. Khosravifard and T. A. Gulliver, “On the capability of the Harada-Kobayashi algorithm in finding fix-free codewords,” in *Proc. Int. Symp. Inf. Theory Appl.*, Auckland, New Zealand, Dec. 2008, pp. 1–4.

[39] S. Kheradmand, M. Khosravifard, and T. Gulliver, “The redundancy of an optimal binary fix-free code is not greater than 1 bit,” *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3549–3558, Jun. 2015.

[40] M. Khosravifard, H. Halabian, and T. Gulliver, “A kraft-type sufficient condition for the existence of D-ary fix-free codes,” *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2920–2927, Jun. 2010.

[41] M. Khosravifard and S. Kheradmand, “Some upper bounds on the redundancy of optimal binary fix-free codes,” *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 4049–4057, Jun. 2012.

[42] J. Kliewer and R. Thobaben, “Iterative joint source-channel decoding of variable-length codes using residual source redundancy,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 919–929, May 2005.

[43] L. G. Kraft, “A device for quantizing, grouping, and coding amplitude modulated pulses,” M.S. thesis, Dept. Elect. Eng., MIT, Cambridge, MA, USA, 1949.

[44] Zs. Kukorelly and K. Zeger, “Sufficient conditions for existence of binary fix-free codes,” *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3433–3444, Oct. 2005.

[45] K. Lakovic and J. Villasenor, “On design of error-correcting reversible variable length codes,” *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 337–339, Aug. 2002.

[46] K. Lakovic and J. Villasenor, “An algorithm for construction of efficient fix-free codes,” *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 391–393, Aug. 2003.

[47] M. Leonard, “A property of biprefix codes,” *RAIRO-Theor. Informat. Appl.*, vol. 22, no. 3, pp. 311–318, 1988.

[48] C.-W. Lin, J.-L. Wu, and Y.-J. Chuang, “Two algorithms for constructing efficient Huffman-code based reversible variable length codes,” *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 81–89, Jan. 2008.

[49] D. Perrin, "Codes bipréfixes et groupes de permutations," M.S. thesis, Dept. Math., Université Paris, Paris, France, 1975.

[50] D. Perrin, "La transitivite du groupe d'un code biprefixe fini," *Mathematische Zeitschrift*, vol. 153, no. 3, pp. 283–287, Oct. 1977.

[51] D. Perrin, "Le degré minimal du groupe d'un code bipréfixe fini," *J. Combinat. Theory, A*, vol. 25, pp. 163–173, Sep. 1978.

[52] D. Perrin, "Completing biprefix codes," in *Automata, Languages and Programming* (Lecture Notes in Computer Science), vol. 140. Aarhus, Denmark: Springer-Verlag, 1982, pp. 397–406.

[53] J. L. Peterson, "Computer programs for detecting and correcting spelling errors," *Commun. ACM*, vol. 23, no. 12, pp. 676–687, Dec. 1980.

[54] C. Reutenauer, "Semisimplicity of the algebra associated to a biprefix code," *Semigroup Forum*, vol. 23, no. 1, pp. 327–342, Jan. 1981.

[55] S. A. Savari, S. M. H. T. Yazdi, N. Abedini, and S. P. Khatri, "On optimal and achievable fix-free codes," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5112–5129, Aug. 2012.

[56] H. Schnettler, "On the $\frac{3}{4}$-conjecture for fix-free codes—A survey," 2007, *arXiv:0709.2598*.

[57] M. P. Schützenberger, "On an application of semi groups methods to some problems in coding," *IRE Trans. Inf. Theory*, vol. 2, no. 3, pp. 47–60, Sep. 1956.

[58] M. P. Schutzenberger, "On a special class of recurrent events," *Ann. Math. Statist.*, vol. 32, no. 4, pp. 1201–1213, Dec. 1961.

[59] M. P. Schützenberger, "On a family of submonoids," Math. Inst. Hungarian Acad. Sci., Budapest, Hungary, 1961, pp. 381–391, vol. 6.

[60] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–666, Oct. 1948.

[61] Y. Takishima, M. Wada, and H. Murakami, "Reversible variable length codes," *IEEE Trans. Commun.*, vol. 43, nos. 2–4, pp. 158–162, Feb. 1995.

[62] C.-W. Tsai, T.-J. Huang, K.-L. Fang, and J.-L. Wu, "A hybrid and flexible H.263-based error resilient and testing system," in *Proc. IEEE Region 10 Int. Conf. Electr. Electron. Technol. (TENCON)*, vol. 1, Aug. 2001, pp. 122–128.

[63] C.-W. Tsai and J.-L. Wu, "On constructing the Huffman-code-based reversible variable-length codes," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1506–1509, Sep. 2001.

[64] C.-W. Tsai and J.-L. Wu, "Modified symmetrical reversible variable-length code and its theoretical bounds," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2543–2548, Sep. 2001.

[65] C. W. Tsai, J. L. Wu, and S. W. Liu, "Modified symmetrical reversible variable length code and its theoretical bounds," *Proc. SPIE*, vol. 3974, pp. 606–616, Apr. 2000.

[66] H.-W. Tseng, "Construction of symmetrical reversible variable length codes using backtracking," *Comput. J.*, vol. 46, no. 1, pp. 100–105, Jan. 2003.

[67] H. Wang, S. N. Koh, and W. W. Chang, "Application of reversible variable-length codes in robust speech coding," *IEE Proc.-Commun.*, vol. 152, no. 3, pp. 272–276, Jun. 2005.

[68] J. Wang, L.-L. Yang, and L. Hanzo, "Iterative construction of reversible variable-length codes and variable-length error-correcting codes," *IEEE Commun. Lett.*, vol. 8, no. 11, pp. 671–673, Nov. 2004.

[69] J. L. H. Webb, "Efficient table access for reversible variable-length decoding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 981–985, Aug. 2001.

[70] J. Wen and J. D. Villasenor, "A class of reversible variable length codes for robust image and video coding," in *Proc. Int. Conf. Image Process. (ICIP)*, Santa Barbara, CA, USA, vol. 2, Oct. 1997, pp. 65–68.

[71] J. Wen and J. D. Villasenor, "Reversible variable length codes for efficient and robust image and video coding," in *Proc. Data Compress. Conf. (DCC)*, Snowbird, UT, USA, Mar./Apr. 1998, pp. 471–480.

[72] N. Yadav, K. C. Roy, and Y. Krishan, "Construction of reversible variable length code for digital image processing," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 10, pp. 5332–5336, Oct. 2010.

[73] Z. Yan, S. Kumar, J. Li, and C. C. J. Kuo, "Reversible variable length codes (RVLC) for robust coding of 3D topological mesh data," in *Proc. DCC Data Compress. Conf.*, Snowbird, UT, USA, 1999, p. 560.

[74] S. M. H. T. Yazdi and S. A. Savari, "On the relationships among optimal symmetric fix-free codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4567–4583, Aug. 2014.

[75] C. Ye and R. W. Yeung, "Some basic properties of fix-free codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 72–87, Jan. 2001.

[76] S. Yekhanin, "Sufficient conditions of existence of fix-free codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Washington, DC, USA, Jun. 2001, p. 284.

[77] S. Yekhanin, "Improved upper bound for the redundancy of fix-free codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2815–2818, Nov. 2004.

[78] S. Yekhanin, "Sufficient conditions of existence of fix-free codes," unpublished manuscript.

[79] A. Zaghian, A. Aghajan, and T. Gulliver, "The optimal fix-free code for anti-uniform sources," *Entropy*, vol. 17, no. 3, pp. 1379–1386, Mar. 2015.

[80] S. J. Zahabi, A. Aghajan, and M. Khosravifard, "Sequentially-constructible reversible variable length codes," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2605–2614, Aug. 2014.

[81] S. J. Zahabi and M. Khosravifard, "On the penalty of optimal fix-free codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2776–2787, May 2015.

[82] A. Zammit, "Reversible variable-length codes," M.S. dissertation, Dept. Commun. Comput. Eng., Univ. Malta, Msida, Malta, 2007.

Spencer Congero (Student Member, IEEE) was born in Hartford, Connecticut, in 1994. He received the bachelor's degree in electrical engineering from the University of Southern California in 2016. He is currently pursuing the Ph.D. degree with the University of California at San Diego.

Kenneth Zeger was born in Boston, MA, USA, in 1963. He received the S.B. and S.M. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1984, and the M.A. degree in mathematics and the Ph.D. degree in electrical engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 1989 and 1990, respectively. He was an Assistant Professor of electrical engineering at the University of Hawaii from 1990 to 1992. He was with the Department of Electrical and Computer Engineering and the Co-ordinated Science Laboratory, University of Illinois at Urbana–Champaign, as an Assistant Professor from 1992 to 1995, and an Associate Professor from 1995 to 1996. From 1996 to 1998, he was with the Department of Electrical and Computer Engineering, University of California at San Diego, as an Associate Professor, where he is currently a Professor. He received an NSF Presidential Young Investigator Award in 1991. He served as an Associate Editor At-Large for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1995 to 1998 and a member of the Board of Governors for the IEEE Information Theory Society from 1998 to 2000, from 2005 to 2007, and from 2008 to 2010.