PAPER *Special Section on Information Theory and Its Applications*

Properties of k-Bit Delay Decodable Codes

Kengo HASHIMOTO[†a)], ***Nonmember and*** **Ken-ichi IWATA**[†], ***Senior Member***

SUMMARY The class of k-bit delay decodable codes, source codes allowing decoding delay of at most k bits for $k \geq 0$, can attain a shorter average codeword length than Huffman codes. This paper discusses the general properties of the class of k-bit delay decodable codes with a finite number of code tables and proves two theorems which enable us to limit the scope of codes to be considered when discussing optimal k-bit delay decodable codes.

key words: *data compression, source coding, decoding delay*

1. Introduction

It is known that one can achieve a shorter average codeword length than Huffman codes by allowing multiple code tables and some decoding delay. AIFV (almost instantaneous fixed-to-variable length) codes developed by Yamamoto, Tsuchihashi, and Honda [3] attain a shorter average codeword length than Huffman codes by using a time-variant encoder with two code tables and allowing decoding delay of at most two bits. AIFV codes are generalized to AIFV-m codes, which can achieve a shorter average codeword length than AIFV codes for $m \geq 3$, allowing m code tables and decoding delay of at most m bits [7]. The worst-case redundancy of AIFV-m codes is analyzed in [7], [8] for $m = 2, 3, 4, 5$. The literature [9]–[22] proposes the code construction and coding method of AIFV and AIFV-m codes. Extensions of AIFV-m codes are proposed in [23], [24].

The literature [4] formalizes a binary encoder with a finite number of code tables as a *code-tuple* and introduces the class of code-tuples decodable with a delay of at most k bits as the class of *k-bit delay decodable codes*, which includes the class of AIFV-k codes as a proper subclass. Also, [4] proves that Huffman codes achieve the optimal average codeword length in the class of 1-bit delay decodable code-tuples. The literature [5] indicates that the class of AIFV codes achieves the optimal average codeword length in the class of 2-bit delay decodable code-tuples with two code tables.

This paper discusses the general properties of k-bit delay decodable code-tuples for $k \geq 0$ and proves two theorems as the main results. The first theorem guarantees that it is not the case that one can achieve an arbitrarily small average codeword length by using arbitrarily many code tables. This leads to the existence of an *optimal* k-bit delay decodable code-tuple, which achieves an average codeword length shorter than or equal to any other k-bit delay code-tuple. The first theorem also gives an upper bound of the required number of code tables for an optimal k-bit delay decodable code-tuples. The second theorem gives a necessary condition for a k-bit delay decodable code-tuple to be optimal, which is a generalization of a property of Huffman codes that each internal node in a code tree has two child nodes. Both theorems enable us to limit the scope of code-tuples to be considered when discussing optimal k-bit delay decodable code-tuples. As an application of these theorems, we can prove of the optimality of AIFV codes in the class of 2-bit delay decodable codes with a finite number of code tables [6].

This paper is organized as follows. In Sect. 2, we prepare some notations, describe our data compression scheme, introduce some notions including k-bit delay decodable codes, and show their basic properties used to prove our main result. Then we prove two theorems in Sect. 3 as the main results of this paper. Lastly, we conclude this paper in Sect. 4. To clarify the flow of the discussion, we relegate the proofs of most of the lemmas to Appendix C. The main notations are listed in Appendix D.

2. Preliminaries

First, we define some notations as follows. Most of the notations in this paper are based on [4]. Let $\mathbb{R}$ denote the set of all real numbers, and let $\mathbb{R}^m$ denote the set of all m dimensional real row vectors for an integer $m \geq 1$. Let $|\mathcal{A}|$ denote the cardinality of a finite set $\mathcal{A}$. Let $\mathcal{A} \times \mathcal{B}$ denote the Cartesian product of $\mathcal{A}$ and $\mathcal{B}$, that is, $\mathcal{A} \times \mathcal{B} := \{(a, b) : a \in \mathcal{A}, b \in \mathcal{B}\}$. Let $\mathcal{A}^k$ (resp. $\mathcal{A}^{\leq k}$, $\mathcal{A}^{\geq k}$, $\mathcal{A}^*$, $\mathcal{A}^+$) denote the set of all sequences of length k (resp. of length less than or equal to k, of length greater than or equal to k, of finite length, of finite positive length) over a set $\mathcal{A}$. Thus, $\mathcal{A}^+ = \mathcal{A}^* \setminus \{\lambda\}$, where λ denotes the empty sequence. The length of a sequence $\boldsymbol{x}$ is denoted by $|\boldsymbol{x}|$, in particular, $|\lambda| = 0$. For a non-empty sequence $\boldsymbol{x} = x_1 x_2 \ldots x_n$, we define $\mathrm{pref}(\boldsymbol{x}) := x_1 x_2 \ldots x_{n-1}$ and $\mathrm{suff}(\boldsymbol{x}) := x_2 \ldots x_{n-1} x_n$. Namely, $\mathrm{pref}(\boldsymbol{x})$ (resp. $\mathrm{suff}(\boldsymbol{x})$) is the sequence obtained by deleting the last (resp. first) letter from $\boldsymbol{x}$. We say $\boldsymbol{x} \preceq \boldsymbol{y}$ if $\boldsymbol{x}$ is a prefix of $\boldsymbol{y}$, that is, there exists a sequence $\boldsymbol{z}$, possibly $\boldsymbol{z} = \lambda$, such that $\boldsymbol{y} = \boldsymbol{x}\boldsymbol{z}$. Also, we say $\boldsymbol{x} \prec \boldsymbol{y}$ if $\boldsymbol{x} \preceq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$. For sequences $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $\boldsymbol{x} \preceq \boldsymbol{y}$, let $\boldsymbol{x}^{-1}\boldsymbol{y}$

Manuscript received February 27, 2023.
Manuscript revised June 23, 2023.
Manuscript publicized September 7, 2023.
[†]The authors are with the University of Fukui, Fukui-shi, 910-8507 Japan.
a) E-mail: kengo3.1415@gmail.com
DOI: 10.1587/transfun.2023TAP0016

denote the unique sequence $\boldsymbol{z}$ such that $\boldsymbol{xz} = \boldsymbol{y}$. Note that a notation $\boldsymbol{x}^{-1}$ behaves like the "inverse element" of $\boldsymbol{x}$ as stated in the following statements (i)–(iii).

(i) For any $\boldsymbol{x}$, we have $\boldsymbol{x}^{-1}\boldsymbol{x} = \lambda$.
(ii) For any $\boldsymbol{x}$ and $\boldsymbol{y}$ such that $\boldsymbol{x} \preceq \boldsymbol{y}$, we have $\boldsymbol{xx}^{-1}\boldsymbol{y} = \boldsymbol{y}$.
(iii) For any $\boldsymbol{x}, \boldsymbol{y}$, and $\boldsymbol{z}$ such that $\boldsymbol{xy} \preceq \boldsymbol{z}$, we have $(\boldsymbol{xy})^{-1}\boldsymbol{z} = \boldsymbol{y}^{-1}\boldsymbol{x}^{-1}\boldsymbol{z}$.

The main notations used in this paper are listed in Appendix D.

We now describe the details of our data compression system. In this paper, we consider a data compression system consisting of a source, an encoder, and a decoder.

- Source: We consider an i.i.d. source, which outputs a sequence $\boldsymbol{x} = x_1x_2\dots x_n$ of symbols of the source alphabet $\mathcal{S} = \{s_1, s_2, \dots, s_\sigma\}$, where n and σ denote the length of $\boldsymbol{x}$ and the alphabet size, respectively. Each source output follows a fixed probability distribution $(\mu(s_1), \mu(s_2), \dots, \mu(s_\sigma))$, where $\mu(s_i)$ is the probability of occurrence of s_i for $i = 1, 2, \dots, \sigma$. In this paper, we assume $\sigma \geq 2$.
- Encoder: The encoder has m fixed code tables $f_0, f_1, \dots, f_{m-1} : \mathcal{S} \to \mathcal{C}^*$, where $\mathcal{C} := \{0, 1\}$ is the coding alphabet. The encoder reads the source sequence $\boldsymbol{x} \in \mathcal{S}^*$ symbol by symbol from the beginning of $\boldsymbol{x}$ and encodes them according to the code tables. For the first symbol x_1, we use an arbitrarily chosen code table from $f_0, f_1, \dots, f_{m-1}$. For $x_2, x_3, \dots, x_n$, we determine which code table to use to encode according to m fixed mappings $\tau_0, \tau_1, \dots, \tau_{m-1} : \mathcal{S} \to [m] := \{0, 1, 2, \dots, m-1\}$. More specifically, if the previous symbol x_{i-1} is encoded by the code table f_j, then the current symbol x_i is encoded by the code table $f_{\tau_j(x_{i-1})}$. Hence, if we use the code table f_i to encode x_1, then a source sequence $\boldsymbol{x} = x_1x_2\dots x_n$ is encoded to a codeword sequence $f(\boldsymbol{x}) := f_{i_1}(x_1)f_{i_2}(x_n)\dots f_{i_n}(x_n)$, where

$$i_j := \begin{cases} i & \text{if } j = 1, \\ \tau_{i_{j-1}}(x_{j-1}) & \text{if } j \geq 2 \end{cases} \tag{1}$$

  for $j = 1, 2, \dots, n$.
- Decoder: The decoder reads the codeword sequence $f(\boldsymbol{x})$ bit by bit from the beginning of $f(\boldsymbol{x})$. Each time the decoder reads a bit, the decoder recovers as long prefix of $\boldsymbol{x}$ as the decoder can uniquely identify from the prefix of $f(\boldsymbol{x})$ already read. We assume that the encoder and decoder share the index of the code table used to encode x_1 in advance.

2.1 Code-Tuples

The behavior of the encoder and decoder for a given source sequence is completely determined by m code tables $f_0, f_1, \dots, f_{m-1}$ and m mappings $\tau_0, \tau_1, \dots, \tau_{m-1}$ if we fix the index of code table used to encode x_1. Accordingly, we name a tuple $F(f_0, f_1, \dots, f_{m-1}, \tau_0, \tau_1, \dots, \tau_{m-1})$ as a *code-tuple* F

Table 1 Two examples of an code-tuple: $F^{(\alpha)}(f_0^{(\alpha)}, f_1^{(\alpha)}, f_2^{(\alpha)}, \tau_0^{(\alpha)}, \tau_1^{(\alpha)}, \tau_2^{(\alpha)})$ and $F^{(\beta)}(f_0^{(\beta)}, f_1^{(\beta)}, f_2^{(\beta)}, \tau_0^{(\beta)}, \tau_1^{(\beta)}, \tau_2^{(\beta)})$.

$s \in \mathcal{S}$	$f_0^{(\alpha)}$	$\tau_0^{(\alpha)}$	$f_1^{(\alpha)}$	$\tau_1^{(\alpha)}$	$f_2^{(\alpha)}$	$\tau_2^{(\alpha)}$
a	01	0	00	1	1100	1
b	10	1	λ	0	1110	2
c	0100	0	00111	1	111000	2
d	01	2	00111	2	110	2

$s \in \mathcal{S}$	$f_0^{(\beta)}$	$\tau_0^{(\beta)}$	$f_1^{(\beta)}$	$\tau_1^{(\beta)}$	$f_2^{(\beta)}$	$\tau_2^{(\beta)}$
a	λ	1	0110	1	λ	2
b	101	2	01	1	λ	2
c	1011	1	0111	1	λ	2
d	1101	2	01111	1	λ	2

and identify a source code with a code-tuple F.

Definition 1. *Let m be a positive integer. An m-*code-tuple *$F(f_0, f_1, \dots, f_{m-1}, \tau_0, \tau_1, \dots, \tau_{m-1})$ is a tuple of m mappings $f_0, f_1, \dots, f_{m-1} : \mathcal{S} \to \mathcal{C}^*$ and m mappings $\tau_0, \tau_1, \dots, \tau_{m-1} : \mathcal{S} \to [m]$.*

We define $\mathscr{F}^{(m)}$ as the set of all m-code-tuples. Also, we define $\mathscr{F} := \mathscr{F}^{(1)} \cup \mathscr{F}^{(2)} \cup \mathscr{F}^{(3)} \cup \cdots$. An element of $\mathscr{F}$ is called a code-tuple.

We write $F(f_0, f_1, \dots, f_{m-1}, \tau_0, \tau_1, \dots, \tau_{m-1})$ also as $F(f, \tau)$ or F for simplicity. For $F \in \mathscr{F}^{(m)}$, let $|F|$ denote the number of code tables of F, that is, $|F| := m$. We write $[|F|] = \{0, 1, 2, \dots, |F| - 1\}$ as $[F]$ for simplicity.

Example 1. *Table 1 shows two examples $F^{(\alpha)}$ and $F^{(\beta)}$ of a 3-code-tuple for $\mathcal{S} = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}\}$.*

Example 2. *We consider encoding of a source sequence $\boldsymbol{x} = x_1x_2x_3x_4 :=$ badb with the code-tuple $F(f, \tau) := F^{(\alpha)}(f^{(\alpha)}, \tau^{(\alpha)})$ in Table 1. If $x_1 = \mathrm{b}$ is encoded with the code table f_0, then the encoding process is as follows.*

- *$x_1 = \mathrm{b}$ is encoded to $f_0(\mathrm{b}) = 10$. The index of the next code table is $\tau_0(\mathrm{b}) = 1$.*
- *$x_2 = \mathrm{a}$ is encoded to $f_1(\mathrm{a}) = 00$. The index of the next code table is $\tau_1(\mathrm{a}) = 1$.*
- *$x_3 = \mathrm{d}$ is encoded to $f_1(\mathrm{d}) = 00111$. The index of the next code table is $\tau_1(\mathrm{d}) = 2$.*
- *$x_4 = \mathrm{b}$ is encoded to $f_2(\mathrm{b}) = 1110$. The index of the next code table is $\tau_2(\mathrm{b}) = 2$.*

As the result, we obtain a codeword sequence $\boldsymbol{c} := f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})f_2(\mathrm{b}) = 1000001111110$.

The decoding process of $\boldsymbol{c} = 1000001111110$ is as follows.

- *After reading the prefix 10 of $\boldsymbol{c}$, the decoder can uniquely identify $x_1 = \mathrm{b}$ and $10 = f_0(\mathrm{b})$. The decoder can also know that x_2 should be decoded with $f_{\tau_0(\mathrm{b})} = f_1$.*
- *After reading the prefix $1000 = f_0(\mathrm{c})f_0(\mathrm{a})$ of $\boldsymbol{c}$, the decoder still cannot uniquely identify $x_2 = \mathrm{a}$ because there remain three possible cases: the case $x_2 = \mathrm{a}$, the case $x_2 = \mathrm{c}$, and the case $x_2 = \mathrm{d}$.*
- *After reading the prefix 10000 of $\boldsymbol{c}$, the decoder can*

uniquely identify $x_2 = \mathrm{a}$ *and* $10000 = f_0(\mathrm{b})f_1(\mathrm{a})0$. *The decoder can also know that* x_3 *should be decoded with* $f_{\tau_1(\mathrm{a})} = f_1$.

- *After reading the prefix* $100000111 = f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})$ *of* $\boldsymbol{c}$, *the decoder still cannot uniquely identify* $x_3 = \mathrm{d}$ *because there remain two possible cases: the case* $x_3 = \mathrm{c}$ *and the case* $x_3 = \mathrm{d}$.
- *After reading the prefix* 10000011111 *of* $\boldsymbol{c}$, *the decoder can uniquely identify* $x_3 = \mathrm{d}$ *and* $10000011111 = f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})11$. *The decoder can also know that* x_4 *should be decoded with* $f_{\tau_1(\mathrm{d})} = f_2$.
- *After reading the prefix* $\boldsymbol{c} = 1000001111110$, *the decoder can uniquely identify* $x_4 = \mathrm{b}$ *and* $1000001111110 = f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})f_2(\mathrm{b})$.

As the result, the decoder recovers the original sequence $\boldsymbol{x} = \mathrm{badb}$.

In encoding $\boldsymbol{x} = x_1x_2\ldots x_n \in \mathcal{S}^*$ with $F(f,\tau) \in \mathscr{F}$, the m mappings $\tau_0, \tau_1, \ldots, \tau_{m-1}$ determine which code table to use to encode $x_2, x_3, \ldots, x_n$. However, there are choices of which code table to use for the first symbol x_1. For $i \in [F]$ and $\boldsymbol{x} \in \mathcal{S}^*$, we define $f_i^*(\boldsymbol{x}) \in \mathcal{C}^*$ as the codeword sequence in the case where x_1 is encoded with f_i. Also, we define $\tau_i^*(\boldsymbol{x}) \in [F]$ as the index of the code table used next after encoding $\boldsymbol{x}$ in the case where x_1 is encoded with f_i. We give formal definitions of f_i^* and τ_i^* in the following Definition 2 as recursive formulas.

Definition 2. *For* $F(f,\tau) \in \mathscr{F}$ *and* $i \in [F]$, *we define a mapping* $f_i^* : \mathcal{S}^* \to \mathcal{C}^*$ *and a mapping* $\tau_i^* : \mathcal{S}^* \to [F]$ *as*

$$f_i^*(\boldsymbol{x}) = \begin{cases} \lambda & \text{if } \boldsymbol{x} = \lambda, \\ f_i(x_1)f^*_{\tau_i(x_1)}(\mathrm{suff}(\boldsymbol{x})) & \text{if } \boldsymbol{x} \neq \lambda, \end{cases} \tag{2}$$

$$\tau_i^*(\boldsymbol{x}) = \begin{cases} i & \text{if } \boldsymbol{x} = \lambda, \\ \tau^*_{\tau_i(x_1)}(\mathrm{suff}(\boldsymbol{x})) & \text{if } \boldsymbol{x} \neq \lambda \end{cases} \tag{3}$$

for $\boldsymbol{x} = x_1x_2\ldots x_n \in \mathcal{S}^*$.

Example 3. *We consider* $F(f,\tau) := F^{(\alpha)}(f^{(\alpha)}, \tau^{(\alpha)})$ *of Table 1. Then* $f_0^*(\mathrm{badb})$ *and* $\tau_0^*(\mathrm{badb})$ *is given as follows (cf. Example 2):*

$$\begin{aligned} f_0^*(\mathrm{badb}) &= f_0(\mathrm{b})f_1^*(\mathrm{adb}) \\ &= f_0(\mathrm{b})f_1(\mathrm{a})f_1^*(\mathrm{db}) \\ &= f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})f_2^*(\mathrm{b}) \\ &= f_0(\mathrm{b})f_1(\mathrm{a})f_1(\mathrm{d})f_2(\mathrm{b})f_2^*(\lambda) \\ &= 1000001111110 \end{aligned}$$

and

$$\tau_0^*(\mathrm{badb}) = \tau_1^*(\mathrm{adb}) = \tau_1^*(\mathrm{db}) = \tau_2^*(\mathrm{b}) = \tau_2^*(\lambda) = 2.$$

The following Lemma 1 follows from Definition 2.

Lemma 1. *For any* $F(f,\tau) \in \mathscr{F}$, $i \in [F]$, *and* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}^*$, *the following statements (i)–(iii) hold.*

(i) $f_i^*(\boldsymbol{xy}) = f_i^*(\boldsymbol{x})f^*_{\tau_i^*(\boldsymbol{x})}(\boldsymbol{y})$.
(ii) $\tau_i^*(\boldsymbol{xy}) = \tau^*_{\tau_i^*(\boldsymbol{x})}(\boldsymbol{y})$.
(iii) If $\boldsymbol{x} \preceq \boldsymbol{y}$, *then* $f_i^*(\boldsymbol{x}) \preceq f_i^*(\boldsymbol{y})$.

2.2 k-Bit Delay Decodable Code-Tuples

In Example 2, despite $f_0(\mathrm{b})f_1(\mathrm{a}) = 1000$, to uniquely identify $x_1x_2 = \mathrm{ba}$, it is required to read 10000 including the additional 1 bit. Namely, a decoding delay of 1 bit occurs to decode $x_2 = \mathrm{a}$. Similarly, despite $f_0(\mathrm{b})f_0(\mathrm{a})f_1(\mathrm{d}) = 100000111$, to uniquely identify $x_1x_2x_3 = \mathrm{bad}$, it is required to read 10000011111 including the additional 2 bits. Namely, a decoding delay of 2 bits occurs to decode $x_3 = \mathrm{d}$. In general, in the decoding process with $F^{(\alpha)}$ in Table 1, it is required to read the additional at most 2 bits for the decoder to uniquely identify each symbol of a given source sequence. We say a code-tuple is *k-bit delay decodable* if the decoder can always uniquely identify each source symbol by reading the additional k bits of the codeword sequence. The code-tuple $F^{(\alpha)}$ is an example of a 2-bit delay decodable code-tuple. To state the formal definition of a k-bit delay decodable code-tuple, we introduce the following Definitions 3 and 4.

Definition 3. *For an integer* $k \geq 0$, $F(f,\tau) \in \mathscr{F}, i \in [F]$, *and* $\boldsymbol{b} \in \mathcal{C}^*$, *we define*

$$\mathcal{P}^k_{F,i}(\boldsymbol{b}) := \{\boldsymbol{c} \in \mathcal{C}^k : \boldsymbol{x} \in \mathcal{S}^+, f_i^*(\boldsymbol{x}) \succeq \boldsymbol{bc}, f_i(x_1) \succeq \boldsymbol{b}\}, \tag{4}$$

$$\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) := \{\boldsymbol{c} \in \mathcal{C}^k : \boldsymbol{x} \in \mathcal{S}^+, f_i^*(\boldsymbol{x}) \succeq \boldsymbol{bc}, f_i(x_1) \succ \boldsymbol{b}\}, \tag{5}$$

where x_1 *denotes the first symbol of* $\boldsymbol{x}$. *Namely,* $\mathcal{P}^k_{F,i}(\boldsymbol{b})$ *(resp.* $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b})$*) is the set of all* $\boldsymbol{c} \in \mathcal{C}^k$ *such that there exists* $\boldsymbol{x} = x_1x_2\ldots x_n \in \mathcal{S}^+$ *satisfying* $f_i^*(\boldsymbol{x}) \succeq \boldsymbol{bc}$ *and* $f_i(x_1) \succeq \boldsymbol{b}$ *(resp.* $f_i(x_1) \succ \boldsymbol{b}$*).*

Definition 4. *For* $F(f,\tau) \in \mathscr{F}, i \in [F]$, *and* $\boldsymbol{b} \in \mathcal{C}^*$, *we define*

$$\mathcal{P}^*_{F,i}(\boldsymbol{b}) := \mathcal{P}^0_{F,i}(\boldsymbol{b}) \cup \mathcal{P}^1_{F,i}(\boldsymbol{b}) \cup \mathcal{P}^2_{F,i}(\boldsymbol{b}) \cup \cdots, \tag{6}$$

$$\bar{\mathcal{P}}^*_{F,i}(\boldsymbol{b}) := \bar{\mathcal{P}}^0_{F,i}(\boldsymbol{b}) \cup \bar{\mathcal{P}}^1_{F,i}(\boldsymbol{b}) \cup \bar{\mathcal{P}}^2_{F,i}(\boldsymbol{b}) \cup \cdots. \tag{7}$$

We write $\mathcal{P}^k_{F,i}(\lambda)$ (resp. $\bar{\mathcal{P}}^k_{F,i}(\lambda)$) as $\mathcal{P}^k_{F,i}$ (resp. $\bar{\mathcal{P}}^k_{F,i}$) for simplicity. Also, we write $\mathcal{P}^*_{F,i}(\lambda)$ (resp. $\bar{\mathcal{P}}^*_{F,i}(\lambda)$) as $\mathcal{P}^*_{F,i}$ (resp. $\bar{\mathcal{P}}^*_{F,i}$). We have

$$\begin{aligned} \mathcal{P}^k_{F,i} &\overset{\text{(A)}}{=} \{\boldsymbol{c} \in \mathcal{C}^k : \boldsymbol{x} \in \mathcal{S}^+, f_i^*(\boldsymbol{x}) \succeq \boldsymbol{c}\} \\ &\overset{\text{(B)}}{=} \{\boldsymbol{c} \in \mathcal{C}^k : \boldsymbol{x} \in \mathcal{S}^*, f_i^*(\boldsymbol{x}) \succeq \boldsymbol{c}\}, \end{aligned} \tag{8}$$

where (A) follows from (4), and (B) is justified as follows. The relation "$\subseteq$" holds by $\mathcal{S}^+ \subseteq \mathcal{S}^*$. We show the relation "$\supseteq$". We choose $\boldsymbol{c} \in \mathcal{C}^k$ such that $f_i^*(\boldsymbol{x}) \succeq \boldsymbol{c}$ for some $\boldsymbol{x} \in \mathcal{S}^*$ arbitrarily and show that $f_i^*(\boldsymbol{x}') \succeq \boldsymbol{c}$ for some $\boldsymbol{x}' \in \mathcal{S}^+$. The case $\boldsymbol{x} \in \mathcal{S}^+$ is trivial. In the case $\boldsymbol{x} \in \{\lambda\} = \mathcal{S}^* \setminus \mathcal{S}^+$, then since $\boldsymbol{c} \preceq f_i^*(\boldsymbol{x}) = f_i^*(\lambda) = \lambda$ by (2), we have $\boldsymbol{c} = \lambda$, which leads to that any $\boldsymbol{x}' \in \mathcal{S}^+$ satisfies $f_i^*(\boldsymbol{x}') \succeq \lambda = \boldsymbol{c}$. Hence, the relation "$\supseteq$" holds.

Example 4. *For $F^{(\alpha)}$ in Table 1, we have*

$$\mathcal{P}^0_{F^{(\alpha)},1}(00) = \{\lambda\},$$
$$\mathcal{P}^1_{F^{(\alpha)},1}(00) = \{0, 1\},$$
$$\mathcal{P}^2_{F^{(\alpha)},1}(00) = \{00, 01, 10, 11\},$$
$$\mathcal{P}^3_{F^{(\alpha)},1}(00) = \{000, 001, 010, 011, 100, 101, 111\}$$

and

$$\bar{\mathcal{P}}^0_{F^{(\alpha)},1}(00) = \{\lambda\},$$
$$\bar{\mathcal{P}}^1_{F^{(\alpha)},1}(00) = \{1\},$$
$$\bar{\mathcal{P}}^2_{F^{(\alpha)},1}(00) = \{11\},$$
$$\bar{\mathcal{P}}^3_{F^{(\alpha)},1}(00) = \{111\}.$$

For $F^{(\beta)}$ in Table 1, we have

$$\mathcal{P}^0_{F^{(\beta)},2} = \{\lambda\}, \qquad \bar{\mathcal{P}}^0_{F^{(\beta)},2} = \emptyset,$$
$$\mathcal{P}^1_{F^{(\beta)},2} = \emptyset, \qquad \bar{\mathcal{P}}^1_{F^{(\beta)},2} = \emptyset.$$

We consider the situation where the decoder has already read the prefix $\boldsymbol{b}'$ of a given codeword sequence and identified a prefix $x_1x_2 \ldots x_l$ of the original sequence $\boldsymbol{x}$. Then we have $\boldsymbol{b}' = f_{i_1}(x_1)f_{i_2}(x_2) \ldots f_{i_l}(x_l)\boldsymbol{b}$ for some $\boldsymbol{b} \in C^*$. Put $i := i_{l+1}$ and let $\{s_1, s_2, \ldots, s_r\}$ be the set of all symbols $s \in \mathcal{S}$ such that $f_i(s) = \boldsymbol{b}$. Then there are the following $r + 1$ possible cases for the next symbol x_{l+1}: the case $x_{l+1} = s_1$, the case $x_{l+1} = s_2, \ldots,$ the case $x_{l+1} = s_r$, and the case $f_i(x_{l+1}) \succ \boldsymbol{b}$. For a code-tuple F to be k-bit delay decodable, the decoder must always be able to distinguish these $r+1$ cases by reading the following k bits of the codeword sequence. Namely, it is required that the $r + 1$ sets listed below are disjoint:

- $\mathcal{P}^k_{F,\tau_i(s_1)}$, the set of all possible following k bits in the case $x_{l+1} = s_1$,
- $\mathcal{P}^k_{F,\tau_i(s_2)}$, the set of all possible following k bits in the case $x_{l+1} = s_2$,
- $\cdots$,
- $\mathcal{P}^k_{F,\tau_i(s_r)}$, the set of all possible following k bits in the case $x_{l+1} = s_r$,
- $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b})$, the set of all possible following k bits in the case $f_i(x_{l+1}) \succ \boldsymbol{b}$.

This discussion leads to the following Definition 5.

Definition 5. *Let $k \geq 0$ be an integer. A code-tuple $F(f, \tau)$ is said to be* k-bit delay decodable *if the following conditions (i) and (ii) hold.*

(i) For any $i \in [F]$ and $s \in \mathcal{S}$, it holds that $\mathcal{P}^k_{F,\tau_i(s)} \cap \bar{\mathcal{P}}^k_{F,i}(f_i(s)) = \emptyset$.
(ii) For any $i \in [F]$ and $s, s' \in \mathcal{S}$, if $s \neq s'$ and $f_i(s) = f_i(s')$, then $\mathcal{P}^k_{F,\tau_i(s)} \cap \mathcal{P}^k_{F,\tau_i(s')} = \emptyset$.

For an integer $k \geq 0$, we define $\mathscr{F}_{k\text{-dec}}$ as the set of all k-bit delay decodable code-tuples, that is, $\mathscr{F}_{k\text{-dec}} := \{F \in \mathscr{F} : F$ is k-bit delay decodable$\}$.

Definition 5 is equivalent to the definition of k-bit delay decodable codes in [4]. See Appendix A for the proof.

Example 5. *We define $F(f, \tau)$ as $F^{(\alpha)}$ in Table 1. Then we have $F \in \mathscr{F}_{2\text{-dec}}$ while $F \notin \mathscr{F}_{1\text{-dec}}$ because*

$$\mathcal{P}^1_{F,\tau_0(\mathrm{a})} \cap \mathcal{P}^1_{F,\tau_0(\mathrm{d})} = \{0, 1\} \cap \{1\} = \{1\} \neq \emptyset,$$

that is, F does not satisfy Definition 5 (ii) for $k = 1$.

Next, we define $F(f, \tau)$ as $F^{(\beta)}$ in Table 1. Then we have $F \in \mathscr{F}_{1\text{-dec}}$ while $F \notin \mathscr{F}_{0\text{-dec}}$ because

$$\mathcal{P}^0_{F,\tau_1(\mathrm{c})} \cap \bar{\mathcal{P}}^0_{F,1}(f_1(\mathrm{c})) = \{\lambda\} \cap \{\lambda\} = \{\lambda\} \neq \emptyset,$$

that is, F does not satisfy Definition 5 (i) for $k = 0$.

Remark 1. *A k-bit delay decodable code-tuple F is not necessarily uniquely decodable, that is, the mappings $f_0^*, f_1^*, \ldots, f_{|F|-1}^*$ are not necessarily injective. For example, for $F^{(\alpha)} \in \mathscr{F}_{2\text{-dec}}$ in Table 1, we have $f_0^{(\alpha)^*}(\mathrm{bc}) = 1000111 = f_0^{(\alpha)^*}(\mathrm{bd})$. In general, it is possible that the decoder cannot uniquely recover the last few symbols of the original source sequence in the case where the rest of the codeword sequence is less than k bits. In such a case, we should append additional information for practical use.*

The classes $\mathscr{F}_{k\text{-dec}}, k = 0, 1, 2, \ldots$ form a hierarchical structure: $\mathscr{F}_{0\text{-dec}} \subseteq \mathscr{F}_{1\text{-dec}} \subseteq \mathscr{F}_{2\text{-dec}} \subseteq \cdots$ [4, Lem. 2].

For $F(f, \tau) \in \mathscr{F}$ and $i \in [F]$, the mapping f_i is said to be *prefix-free* if for any $s, s' \in \mathcal{S}$, if $f_i(s) \preceq f_i(s')$, then $s = s'$. A 0-bit delay decodable code-tuple is characterized as a code-tuple all of which code tables are prefix-free [4, Lem. 4].

Lemma 2. *A code-tuple $F(f, \tau) \in \mathscr{F}$ satisfies $F \in \mathscr{F}_{0\text{-dec}}$ if and only if the code tables $f_0, f_1, \ldots, f_{|F|-1}$ are prefix-free.*

2.3 Extendable Code-Tuples

For the code-tuple $F^{(\beta)}$ in Table 1, we can see that $f_2^{(\beta)^*}(\boldsymbol{x}) = \lambda$ for any $\boldsymbol{x} \in \mathcal{S}^*$. To exclude such abnormal and useless code-tuples, we introduce a class $\mathscr{F}_{\mathrm{ext}}$ in the following Definition 6.

Definition 6. *A code-tuple F is said to be* extendable *if $\mathcal{P}^1_{F,i} \neq \emptyset$ for any $i \in [F]$. We define $\mathscr{F}_{\mathrm{ext}}$ as the set of all extendable code-tuples, that is, $\mathscr{F}_{\mathrm{ext}} := \{F \in \mathscr{F} : {}^\forall i \in [F]; \mathcal{P}^1_{F,i} \neq \emptyset\}$.*

Example 6. *For $F^{(\alpha)}$ in Table 1, we have*

$$\mathcal{P}^1_{F^{(\alpha)},0} = \{0, 1\}, \quad \mathcal{P}^1_{F^{(\alpha)},1} = \{0, 1\}, \quad \mathcal{P}^1_{F^{(\alpha)},2} = \{1\}.$$

Therefore, we have $F^{(\alpha)} \in \mathscr{F}_{\mathrm{ext}}$. For $F^{(\beta)}$ in Table 1, we have

$$\mathcal{P}^1_{F^{(\beta)},0} = \{0, 1\}, \quad \mathcal{P}^1_{F^{(\beta)},1} = \{0\}, \quad \mathcal{P}^1_{F^{(\beta)},2} = \emptyset.$$

Since $\mathcal{P}^1_{F^{(\beta)},2} = \emptyset$, we have $F^{(\beta)} \notin \mathscr{F}_{\mathrm{ext}}$.

The following Lemma 3 shows that for an extendable code-tuple F, we can extend the length of $f^*_i(\boldsymbol{x})$ as long as we want by appending symbols to $\boldsymbol{x}$ appropriately.

Lemma 3. *A code-tuple $F(f,\tau)$ is extendable if and only if for any $i \in [F]$ and integer $l \geq 0$, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $|f^*_i(\boldsymbol{x})| \geq l$.*

Proof of Lemma 3. (Sufficiency) Fix $i \in [F]$ arbitrarily. Applying the assumption with $l = 1$, we see that there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $|f^*_i(\boldsymbol{x})| \geq 1$. Then there exists $c \in C$ such that $f^*_i(\boldsymbol{x}) \geq c$, which leads to $c \in \mathcal{P}^1_{F,i}$ by (8), that is, $\mathcal{P}^1_{F,i} \neq \emptyset$ as desired.

(Necessity) Assume $F \in \mathscr{F}_{\mathrm{ext}}$. We prove by induction for l. The base case $l = 0$ is trivial. We consider the induction step for $l \geq 1$. By the induction hypothesis, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that

$$|f^*_i(\boldsymbol{x})| \geq l - 1. \tag{9}$$

Also, by $F \in \mathscr{F}_{\mathrm{ext}}$, there exists $c \in \mathcal{P}^1_{F,\tau^*_i(\boldsymbol{x})}$. By (8), there exists $\boldsymbol{y} \in \mathcal{S}^*$ such that

$$f^*_{\tau^*_i(\boldsymbol{x})}(\boldsymbol{y}) \geq c. \tag{10}$$

Thus, we obtain

$$|f^*_i(\boldsymbol{xy})| \overset{\text{(A)}}{=} |f^*_i(\boldsymbol{x})| + |f^*_{\tau^*_i(\boldsymbol{x})}(\boldsymbol{y})| \overset{\text{(B)}}{\geq} (l-1) + 1 = l, \tag{11}$$

where (A) follows from Lemma 1 (i), and (B) follows from (9) and (10). This completes the induction. □

This property leads to the following Lemma 4 and Corollary 1.

Lemma 4. *Let k, k' be two integers such that $0 \leq k \leq k'$. For any $F(f,\tau) \in \mathscr{F}_{\mathrm{ext}}, i \in [F]$, $\boldsymbol{b} \in C^*$, and $\boldsymbol{c} \in C^k$, the following statements (i) and (ii) hold.*

(i) $\boldsymbol{c} \in \mathcal{P}^k_{F,i}(\boldsymbol{b}) \iff {}^\exists \boldsymbol{c}' \in C^{k'-k}; \boldsymbol{cc}' \in \mathcal{P}^{k'}_{F,i}(\boldsymbol{b})$.
(ii) $\boldsymbol{c} \in \bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) \iff {}^\exists \boldsymbol{c}' \in C^{k'-k}; \boldsymbol{cc}' \in \bar{\mathcal{P}}^{k'}_{F,i}(\boldsymbol{b})$.

Proof of Lemma 4. We prove (i) only because (ii) follows by the similar argument.

($\Longrightarrow$): Assume $\boldsymbol{c} \in \mathcal{P}^k_{F,i}(\boldsymbol{b})$. Then by (4), there exists $\boldsymbol{x} \in \mathcal{S}^+$ such that

$$f^*_i(\boldsymbol{x}) \geq \boldsymbol{bc}, \tag{12}$$
$$f_i(x_1) \geq \boldsymbol{b}. \tag{13}$$

By $F \in \mathscr{F}_{\mathrm{ext}}$ and Lemma 3, there exists $\boldsymbol{y} \in \mathcal{S}^*$ such that

$$|f^*_{\tau^*_i(\boldsymbol{x})}(\boldsymbol{y})| \geq k' - k. \tag{14}$$

Hence, we have

$$|f^*_i(\boldsymbol{xy})| \overset{\text{(A)}}{=} |f^*_i(\boldsymbol{x})| + |f^*_{\tau^*_i(\boldsymbol{x})}(\boldsymbol{y})| \overset{\text{(B)}}{\geq} |\boldsymbol{bc}| + k' - k, \tag{15}$$

where (A) follows from Lemma 1 (i), and (B) follows from (12) and (14). By (12) and (15), there exists $\boldsymbol{c}' \in C^{k'-k}$ such that

$$f^*_i(\boldsymbol{xy}) \geq \boldsymbol{bcc}'. \tag{16}$$

Equations (13) and (16) lead to $\boldsymbol{cc}' \in \mathcal{P}^{k'}_{F,i}(\boldsymbol{b})$ by (4).

($\Longleftarrow$): Assume that there exists $\boldsymbol{c}' \in C^{k'-k}$ such that $\boldsymbol{cc}' \in \mathcal{P}^{k'}_{F,i}(\boldsymbol{b})$. Then by (4), there exists $\boldsymbol{x} = x_1x_2\dots x_n \in \mathcal{S}^+$ such that $f^*_i(\boldsymbol{x}) \geq \boldsymbol{bcc}'$ and $f_i(x_1) \geq \boldsymbol{b}$. This clearly implies $f^*_i(\boldsymbol{x}) \geq \boldsymbol{bc}$ and $f_i(x_1) \geq \boldsymbol{b}$, which leads to $\boldsymbol{c} \in \mathcal{P}^k_{F,i}(\boldsymbol{b})$ by (4). □

Corollary 1. *For any integer $k \geq 0$, $F \in \mathscr{F}_{\mathrm{ext}}$, $i \in [F]$, and $\boldsymbol{b} \in C^*$, we have $\mathcal{P}^k_{F,i}(\boldsymbol{b}) = \emptyset$ (resp. $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) = \emptyset$) if and only if $\mathcal{P}^0_{F,i}(\boldsymbol{b}) = \emptyset$ (resp. $\bar{\mathcal{P}}^0_{F,i}(\boldsymbol{b}) = \emptyset$).*

The following Lemma 5 gives a lower bound of the length of a codeword sequence for $F \in \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$. See Appendix C.1 for the proof of Lemma 5.

Lemma 5. *For any integer $k \geq 0$, $F(f,\tau) \in \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}, i \in [F]$, and $\boldsymbol{x} \in \mathcal{S}^*$, we have $|f^*_i(\boldsymbol{x})| \geq \lfloor |\boldsymbol{x}|/|F| \rfloor$.*

2.4 Average Codeword Length of Code-Tuple

We introduce the average codeword length $L(F)$ of a code-tuple F. From now on, we fix an arbitrary probability distribution μ of the source symbols, that is, a real-valued function $\mu : \mathcal{S} \to \mathbb{R}$ such that $\sum_{s \in \mathcal{S}} \mu(s) = 1$ and $0 < \mu(s) \leq 1$ for any $s \in \mathcal{S}$. Note that we exclude the case where $\mu(s) = 0$ for some $s \in \mathcal{S}$ without loss of generality.

First, for $F(f,\tau) \in \mathscr{F}$ and $i, j \in [F]$, we define the transition probability $Q_{i,j}(F)$ as the probability of using the code table f_j next after using the code table f_i in the encoding process.

Definition 7. *For $F(f,\tau) \in \mathscr{F}$ and $i, j \in [F]$, we define the* transition probability $Q_{i,j}(F)$ *as*

$$Q_{i,j}(F) := \sum_{s \in \mathcal{S}, \tau_i(s) = j} \mu(s). \tag{17}$$

We also define the transition probability matrix $Q(F)$ *as the following $|F| \times |F|$ matrix:*

$$\begin{bmatrix} Q_{0,0}(F) & Q_{0,1}(F) & \cdots & Q_{0,|F|-1}(F) \\ Q_{1,0}(F) & Q_{1,1}(F) & \cdots & Q_{1,|F|-1}(F) \\ \vdots & \vdots & \ddots & \vdots \\ Q_{|F|-1,0}(F) & Q_{|F|-1,1}(F) & \cdots & Q_{|F|-1,|F|-1}(F) \end{bmatrix}. \tag{18}$$

We fix $F \in \mathscr{F}$ and consider the encoding process with F. Let $I_i \in [F]$ be the index of the code table used to encode the i-th symbol of a source sequence for $i = 1, 2, 3, \dots$. Then $\{I_i\}_{i=1,2,3,\dots}$ is a Markov process with the transition probability matrix $Q(F)$. We consider a stationary distribution of the Markov process $\{I_i\}_{i=1,2,3,\dots}$, formally defined as follows.

Definition 8. *For $F \in \mathscr{F}$, a solution $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1}) \in \mathbb{R}^{|F|}$ of the following simultaneous equations (19) and (20) is called a* stationary distribution of F*:*

$$\begin{cases} \boldsymbol{\pi} Q(F) = \boldsymbol{\pi}, & (19) \\ \sum_{i \in [F]} \pi_i = 1. & (20) \end{cases}$$

A code-tuple has at least one stationary distribution without a negative element as shown in the following Lemma 6. See Appendix C.2 for the proof of Lemma 6.

Lemma 6. *For any $F \in \mathscr{F}$, there exists a stationary distribution $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1})$ of F such that $\pi_i \geq 0$ for any $i \in [F]$.*

As stated later in Definition 10, the average codeword length $L(F)$ of F is defined depending on the stationary distribution $\boldsymbol{\pi}$ of F. However, it is possible that a code-tuple has multiple stationary distributions. Therefore, we limit the scope of consideration to a class $\mathscr{F}_{\mathrm{reg}}$ defined as the following Definition 9, which is the class of code-tuples with a unique stationary distribution.

Definition 9. *A code-tuple F is said to be* regular *if F has a unique stationary distribution. We define $\mathscr{F}_{\mathrm{reg}}$ as the set of all regular code-tuples, that is, $\mathscr{F}_{\mathrm{reg}} := \{F \in \mathscr{F} : F \text{ is regular}\}$. For $F \in \mathscr{F}_{\mathrm{reg}}$, we define $\boldsymbol{\pi}(F) = (\pi_0(F), \pi_1(F), \ldots, \pi_{|F|-1}(F))$ as the unique stationary distribution of F.*

Since the transition probability matrix $Q(F)$ depends on μ, it might seem that the class $\mathscr{F}_{\mathrm{reg}}$ also depends on μ. However, we can see later that in fact $\mathscr{F}_{\mathrm{reg}}$ is independent of μ. More precisely, whether a code-tuple $F(f, \tau)$ belongs to $\mathscr{F}_{\mathrm{reg}}$ depends only on $\tau_0, \tau_1, \ldots, \tau_{|F|-1}$.

We also note that for any $F \in \mathscr{F}_{\mathrm{reg}}$, the unique stationary distribution $\boldsymbol{\pi}(F)$ of F satisfies $\pi_i(F) \geq 0$ for any $i \in [F]$ by Lemma 6.

The asymptotical performance (i.e., average codeword length per symbol) of a regular code-tuple does not depend on which code table we start encoding: the average codeword length $L(F)$ of a regular code-tuple F is the weighted sum of the average codeword lengths of the code tables $f_0, f_1, \ldots, f_{|F|-1}$ weighted by the stationary distribution $\boldsymbol{\pi}(F)$. Namely, $L(F)$ is defined as the following Definition 10.

Definition 10. *For $F(f, \tau) \in \mathscr{F}$ and $i \in [F]$, we define the* average codeword length $L_i(F)$ of the single code table $f_i : \mathcal{S} \to \mathcal{C}^*$ *as*

$$L_i(F) := \sum_{s \in \mathcal{S}} |f_i(s)| \cdot \mu(s). \tag{21}$$

For $F \in \mathscr{F}_{\mathrm{reg}}$, we define the average codeword length $L(F)$ of the code-tuple F *as*

$$L(F) := \sum_{i \in [F]} \pi_i(F) L_i(F). \tag{22}$$

Example 7. *We consider $F := F^{(\alpha)}$ of Table 1, where $(\mu(\mathrm{a}), \mu(\mathrm{b}), \mu(\mathrm{c}), \mu(\mathrm{d})) = (0.1, 0.2, 0.3, 0.4)$. We have*

$$Q(F) = \begin{bmatrix} 0.4 & 0.2 & 0.4 \\ 0.2 & 0.4 & 0.4 \\ 0 & 0.1 & 0.9 \end{bmatrix}.$$

The code-tuple F has a unique stationary distribution $\boldsymbol{\pi}(F) = (\pi_0(F), \pi_1(F), \pi_2(F)) = (1/20, 3/20, 16/20)$. Hence, we have $F \in \mathscr{F}_{\mathrm{reg}}$. Also, we have

$$L_0(F) = 2.6, \quad L_1(F) = 3.7, \quad L_2(F) = 4.2.$$

Therefore, $L(F)$ is given as

$$\begin{aligned} L(F) &= \pi_0(F)L_0(F) + \pi_1(F)L_1(F) + \pi_2(F)L_2(F) \\ &= 4.045. \end{aligned}$$

Remark 2. *Note that $Q(F), L_i(F), L(F)$, and $\boldsymbol{\pi}(F)$ depend on μ. However, since we are now discussing on a fixed μ, the average codeword length $L_i(F)$ of f_i (resp. the transition probability matrix $Q(F)$) is determined only by the mapping f_i (resp. $\tau_0, \tau_1, \ldots, \tau_{|F|-1}$) and therefore the stationary distribution $\boldsymbol{\pi}(F)$ of a regular code-tuple F is also determined only by $\tau_0, \tau_1, \ldots, \tau_{|F|-1}$.*

2.5 Irreducible Parts of Code-Tuple

As we can see from (22), the code tables f_i of $F(f, \tau) \in \mathscr{F}_{\mathrm{reg}}$ such that $\pi_i(F) = 0$ does not contribute to $L(F)$. It is useful to remove such non-essential code tables and obtain an *irreducible* code-tuple: we say that a regular code-tuple F is *irreducible* if $\pi_i(F) > 0$ for any $i \in [F]$ as formally defined later in Definition 13. In this subsection, we introduce an *irreducible part* of $F \in \mathscr{F}_{\mathrm{reg}}$, which is an irreducible code-tuple obtained by removing all the code tables f_i such that $\pi_i(F) = 0$ from F. The formal definition of an irreducible part of F is stated using a notion of *homomorphism* defined in the following Definition 11.

Definition 11. *For $F(f, \tau), F'(f', \tau') \in \mathscr{F}$, a mapping $\varphi : [F'] \to [F]$ is called a* homomorphism from F' to F *if*

$$f'_i(s) = f_{\varphi(i)}(s), \tag{23}$$

$$\varphi(\tau'_i(s)) = \tau_{\varphi(i)}(s) \tag{24}$$

for any $i \in [F']$ and $s \in \mathcal{S}$.

Given a homomorphism of code-tuples, the following Lemma 7 holds between the two code-tuples. See Appendix C.3 for the proof of Lemma 7.

Lemma 7. *For any $F(f, \tau), F'(f', \tau') \in \mathscr{F}$ and a homomorphism $\varphi : [F'] \to [F]$ from F' to F, the following statements (i)–(vi) hold.*

(i) For any $i \in [F']$ and $\boldsymbol{x} \in \mathcal{S}^*$, we have $f'^*_i(\boldsymbol{x}) = f^*_{\varphi(i)}(\boldsymbol{x})$

and $\varphi(\tau_i'^*(\boldsymbol{x})) = \tau^*_{\varphi(i)}(\boldsymbol{x})$.

(ii) For any $i \in [F']$ *and* $\boldsymbol{b} \in C^*$, *we have* $\mathcal{P}^*_{F',i}(\boldsymbol{b}) = \mathcal{P}^*_{F,\varphi(i)}(\boldsymbol{b})$ *and* $\bar{\mathcal{P}}^*_{F',i}(\boldsymbol{b}) = \bar{\mathcal{P}}^*_{F,\varphi(i)}(\boldsymbol{b})$.

(iii) For any stationary distribution $\boldsymbol{\pi}' = (\pi'_0, \pi'_1, \ldots, \pi'_{|F'|-1})$ *of* F', *the vector* $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1}) \in \mathbb{R}^{|F|}$ *defined as*

$$\pi_j = \sum_{j' \in \mathcal{A}_j} \pi'_{j'} \text{ for } j \in [F] \tag{25}$$

is a stationary distribution of F, *where*

$$\mathcal{A}_i := \{i' \in [F'] : \varphi(i') = i\} \tag{26}$$

for $i \in [F]$.

(iv) If $F \in \mathscr{F}_{\text{ext}}$, *then* $F' \in \mathscr{F}_{\text{ext}}$.

(v) If $F, F' \in \mathscr{F}_{\text{reg}}$, *then* $L(F') = L(F)$.

(vi) For any integer $k \geq 0$, *if* $F \in \mathscr{F}_{k\text{-dec}}$, *then* $F' \in \mathscr{F}_{k\text{-dec}}$.

We also introduce the set $\mathcal{R}_F$ for $F \in \mathscr{F}$ as the following Definition 12. We state in Lemma 8 that we can characterize a regular code-tuple F by $\mathcal{R}_F$.

Definition 12. *For* $F(f, \tau) \in \mathscr{F}$, *we define* $\mathcal{R}_F$ *as*

$$\mathcal{R}_F := \{i \in [F] : {}^\forall j \in [F]; {}^\exists \boldsymbol{x} \in \mathcal{S}^*; \tau_j^*(\boldsymbol{x}) = i\}. \tag{27}$$

Namely, $\mathcal{R}_F$ *is the set of indices* i *of the code tables such that for any* $j \in [F]$, *there exists* $\boldsymbol{x} \in \mathcal{S}^*$ *such that* $\tau_j^*(\boldsymbol{x}) = i$.

Example 8. *For* $F^{(\alpha)}$ *and* $F^{(\beta)}$ *in Table 1, we have* $\mathcal{R}_{F^{(\alpha)}} = \{0, 1, 2\}$ *and* $\mathcal{R}_{F^{(\beta)}} = \emptyset$.

Lemma 8. *For any* $F \in \mathscr{F}$, *the following statements (i) and (ii) hold.*

(i) $F \in \mathscr{F}_{\text{reg}}$ *if and only if* $\mathcal{R}_F \neq \emptyset$.

(ii) If $F \in \mathscr{F}_{\text{reg}}$, *then for any* $i \in [F]$, *the following equivalence relation holds:* $\pi_i(F) > 0 \iff i \in \mathcal{R}_F$.

The proof of Lemma 8 is given in Appendix C.4.

Since $\mathcal{R}_F$ does not depend on μ, we can see from Lemma 8 (i) that the class $\mathscr{F}_{\text{reg}}$ is determined independently of μ as mentioned before.

By Lemma 8 (ii), a regular code-tuple $F(f, \tau)$ satisfies $\pi_i(F) > 0$ for any $i \in [F]$ if and only if F is an irreducible code-tuple defined as follows.

Definition 13. *A code-tuple* F *is said to be* irreducible *if* $\mathcal{R}_F = [F]$. *We define* $\mathscr{F}_{\text{irr}}$ *as the set of all irreducible code-tuples, that is,* $\mathscr{F}_{\text{irr}} := \{F \in \mathscr{F} : \mathcal{R}_F = [F]\}$.

Note that $\mathscr{F}_{\text{irr}} \subseteq \mathscr{F}_{\text{reg}}$ since $F \in \mathscr{F}_{\text{reg}}$ is equivalent to $\mathcal{R}_F \neq \emptyset$ by Lemma 8 (i).

Now we define an *irreducible part* $\bar{F}$ of a code-tuple F as the following Definition 14.

Definition 14. *An irreducible code-tuple* $\bar{F}$ *is called an* irreducible part *of a code-tuple* F *if there exists an injective homomorphism* $\varphi : [\bar{F}] \to [F]$ *from* $\bar{F}$ *to* F.

The following property of $\bar{F}$ is immediately from Definition 14 and Lemma 7 (iv)–(vi).

Lemma 9. *For any integer* $k \geq 0$, $F \in \mathscr{F}_{\text{reg}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$, *and an irreducible part* $\bar{F}$ *of* F, *we have* $\bar{F} \in \mathscr{F}_{\text{irr}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$ *and* $L(\bar{F}) = L(F)$.

The existence of an irreducible part is guaranteed as the following Lemma 10. See Appendix C.5 for the proof of Lemma 10.

Lemma 10. *For any* $F \in \mathscr{F}_{\text{reg}}$, *there exists an irreducible part* $\bar{F}$ *of* F.

3. Main Results

In this section, we discuss the average codeword length for code-tuples of the class $\mathscr{F}_{\text{reg}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$ for $k \geq 0$ and prove Theorems 1 and 2 as the main results of this paper.

3.1 Theorem 1

The first theorem claims that for any $F \in \mathscr{F}_{\text{reg}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$, there exists $F^\dagger \in \mathscr{F}_{\text{irr}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$ such that $L(F^\dagger) \leq L(F)$ and $\mathcal{P}^k_{F^\dagger,0}, \mathcal{P}^k_{F^\dagger,1}, \ldots, \mathcal{P}^k_{F^\dagger,|F^\dagger|-1}$ are distinct. Namely, Theorem 1 guarantees that it suffices to consider only irreducible code-tuples with at most $2^{(2^k)}$ code tables to achieve a short average codeword length. In particular, it is not the case that one can achieve an arbitrarily small average codeword length by using arbitrarily many code tables. To state Theorem 1, we prepare the following Definition 15.

Definition 15. *For an integer* $k \geq 0$ *and* $F \in \mathscr{F}$, *we define* $\mathscr{P}^k_F$ *as*

$$\mathscr{P}^k_F := \{\mathcal{P}^k_{F,i} : i \in [F]\}. \tag{28}$$

Example 9. *For* $F^{(\alpha)}$ *in Table 1, we have*

$$\mathscr{P}^0_{F^{(\alpha)}} = \{\{\lambda\}\},$$
$$\mathscr{P}^1_{F^{(\alpha)}} = \{\{0, 1\}, \{1\}\},$$
$$\mathscr{P}^2_{F^{(\alpha)}} = \{\{01, 10\}, \{00, 01, 10\}, \{11\}\}.$$

Note that $\mathcal{P}^k_{F,0}, \mathcal{P}^k_{F,1}, \ldots, \mathcal{P}^k_{F,|F|-1}$ are distinct if and only if $|\mathscr{P}^k_F| = |F|$. Also, note that the following Lemma 11 holds by Lemma 7 (ii).

Lemma 11. *For any integer* $k \geq 0$, $F \in \mathscr{F}_{\text{reg}}$, *and an irreducible part* $\bar{F}$ *of* F, *we have* $\mathscr{P}^k_{\bar{F}} \subseteq \mathscr{P}^k_F$.

Using Definition 15, we state Theorem 1 as follows.

Theorem 1. *For any integer* $k \geq 0$ *and* $F \in \mathscr{F}_{\text{reg}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$, *there exists* $F^\dagger \in \mathscr{F}$ *satisfying the following conditions (a)–(d).*

(a) $F^\dagger \in \mathscr{F}_{\text{irr}} \cap \mathscr{F}_{\text{ext}} \cap \mathscr{F}_{k\text{-dec}}$.

(b) $L(F^\dagger) \leq L(F)$.

Table 2 The code-tuple $F^{(\delta)}$ is an optimal 2-bit delay decodable code-tuple satisfying Theorem 1 (a)–(d) with $F = F^{(\gamma)}$, where $(\mu(\mathrm{a}), \mu(\mathrm{b}), \mu(\mathrm{c}), \mu(\mathrm{d})) = (0.1, 0.2, 0.3, 0.4)$.

$s \in \mathcal{S}$	$f_0^{(\gamma)}$	$\tau_0^{(\gamma)}$	$f_1^{(\gamma)}$	$\tau_1^{(\gamma)}$	$f_2^{(\gamma)}$	$\tau_2^{(\gamma)}$	$f_3^{(\gamma)}$	$\tau_3^{(\gamma)}$
a	0010	2	100	1	1100	1	010	0
b	0011	0	00	0	11	2	011	1
c	000	1	01	1	01	1	100	0
d	λ	2	1	2	10	0	1	2

$s \in \mathcal{S}$	$f_0^{(\delta)}$	$\tau_0^{(\delta)}$	$f_1^{(\delta)}$	$\tau_1^{(\delta)}$
a	100	0	1100	0
b	00	0	11	1
c	01	0	01	0
d	1	1	10	0

(c) $\mathscr{P}^k_{F^\dagger} \subseteq \mathscr{P}^k_F$.
(d) $|\mathscr{P}^k_{F^\dagger}| = |F^\dagger|$.

Example 10. *Let* $(\mu(\mathrm{a}), \mu(\mathrm{b}), \mu(\mathrm{c}), \mu(\mathrm{d})) = (0.1, 0.2, 0.3, 0.4)$ *and* $F := F^{(\gamma)}$ *in Table 2. Then we have* $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{2\text{-dec}}$, $L(F) \approx 1.98644$, *and* $\mathscr{P}^2_F = \{\{00, 01, 10, 11\}, \{01, 10, 11\}\}$. *The code-tuple* $F^\dagger := F^{(\delta)}$ *in Table 2 satisfies Theorem 1 (a)–(d) because* $\mathcal{R}_{F^\dagger} = \{0, 1\} = [F^\dagger]$, $L(F^\dagger) \approx 1.8667 \leq L(F)$, *and* $\mathscr{P}^2_{F^\dagger} = \{\{00, 01, 10, 11\}, \{01, 10, 11\}\}$.

Example 11. *We confirm that Theorem 1 holds for* $k = 0$. *Choose* $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{0\text{-dec}}$ *arbitrarily and define* $F^\dagger(f^\dagger, \tau^\dagger) \in \mathscr{F}^{(1)}$ *as*

$$f_0^\dagger(s) = f_p(s), \tag{29}$$

$$\tau_0^\dagger(s) = 0 \tag{30}$$

for $s \in \mathcal{S}$, *where*

$$p \in \arg\min_{i \in [F]} L_i(F). \tag{31}$$

Namely, $F^\dagger$ *is the* 1*-code-tuple consisting of the most efficient code table of* F.

We can see that $F^\dagger$ *satisfies Theorem 1 (a)–(d) as follows.*

(a) *We obtain* $F^\dagger \in \mathscr{F}_{\mathrm{irr}}$ *directly from* $|F^\dagger| = 1$. *By* $F \in \mathscr{F}_{0\text{-dec}}$ *and Lemma 2, all code tables of* F *are prefix-free. In particular,* $f_0^\dagger = f_p$ *is prefix-free and thus* $F^\dagger \in \mathscr{F}_{0\text{-dec}}$. *Moreover, since* $f_0^\dagger$ *is prefix-free and* $\sigma \geq 2$, *we have* $f_0^\dagger(s) \neq \lambda$ *for some* $s \in \mathcal{S}$, *which shows* $F^\dagger \in \mathscr{F}_{\mathrm{ext}}$.
(b) *We have*

$$\begin{aligned} L(F^\dagger) = L_0(F^\dagger) &\overset{\text{(A)}}{=} L_p(F) \\ &= \sum_{i \in [F]} \pi_i(F) L_p(F) \overset{\text{(B)}}{\leq} \sum_{i \in [F]} \pi_i(F) L_i(F) \\ &= L(F), \end{aligned}$$

where (A) follows from (29), and (B) follows from (31).
(c) *By* $\mathscr{P}^0_{F^\dagger} = \{\{\lambda\}\} = \mathscr{P}^0_F$.
(d) *By* $|\mathscr{P}^0_{F^\dagger}| = |\{\{\lambda\}\}| = 1 = |F^\dagger|$.

As a preparation for the proof of Theorem 1, we state the following Lemmas 12–15. See Appendix C.6–C.8 for the proofs of Lemmas 12, 13, and 15.

Lemma 12. *Let* $k \geq 0$ *be an integer and let* $F(f, \tau)$ *and* $F'(f', \tau')$ *be code-tuples such that* $|F| = |F'|$. *Assume that the following conditions (a) and (b) hold.*

(a) $f_i(s) = f'_i(s)$ *for any* $i \in [F]$ *and* $s \in \mathcal{S}$.
(b) $\mathcal{P}^k_{F, \tau_i(s)} = \mathcal{P}^k_{F, \tau'_i(s)}$ *for any* $i \in [F]$ *and* $s \in \mathcal{S}$.

Then the following statements (i)–(iii) hold.

(i) *For any* $i \in [F']$ *and* $\boldsymbol{b} \in \mathcal{C}^*$, *we have* $\mathcal{P}^k_{F,i}(\boldsymbol{b}) = \mathcal{P}^k_{F',i}(\boldsymbol{b})$ *and* $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) = \bar{\mathcal{P}}^k_{F',i}(\boldsymbol{b})$.
(ii) *If* $F \in \mathscr{F}_{\mathrm{ext}}$, *then* $F' \in \mathscr{F}_{\mathrm{ext}}$.
(iii) *If* $F \in \mathscr{F}_{k\text{-dec}}$, *then* $F' \in \mathscr{F}_{k\text{-dec}}$.

Lemma 13. *For any* $F(f, \tau) \in \mathscr{F}_{\mathrm{irr}}$, $\mathcal{I} \subseteq [F]$, *and* $p \in \mathcal{I}$, *the code-tuple* $F'(f', \tau') \in \mathscr{F}^{(|F|)}$ *defined as (32) and (33) satisfies* $F' \in \mathscr{F}_{\mathrm{reg}}$:

$$f'_i(s) = f_i(s), \tag{32}$$

$$\tau'_i(s) = \begin{cases} p & \text{if } \tau_i(s) \in \mathcal{I}, \\ \tau_i(s) & \text{if } \tau_i(s) \notin \mathcal{I} \end{cases} \tag{33}$$

for $i \in [F']$ *and* $s \in \mathcal{S}$.

Lemma 14. *For any* $F \in \mathscr{F}$, *there exists* $(h_0, h_1, \ldots, h_{|F|-1}) \in \mathbb{R}^{|F|}$ *satisfying*

$$^\forall i \in [F]; L(F) = L_i(F) + \sum_{j \in [F]} (h_j - h_i) Q_{i,j}(F). \tag{34}$$

See [25, Sec. 8.2] for proof of Lemma 14. The vector h called "bias" defined as [25, (8.2.2)] satisfies (34) of this paper. This fact is shown as [25, (8.2.12)] in [25, Theorem 8.2.6], where g, r, and P in [25, (8.2.12)] correspond to the notations of this paper as follows:

$$g = \begin{bmatrix} L(F) \\ L(F) \\ \vdots \\ L(F) \end{bmatrix}, \quad r = \begin{bmatrix} L_0(F) \\ L_1(F) \\ \vdots \\ L_{|F|-1}(F) \end{bmatrix}, \quad P = Q(F).$$

A real vector $(h_0, h_1, \ldots, h_{|F|-1})$ satisfying (34) is not unique. We refer to arbitrarily chosen one of them as $h(F) = (h_0(F), h_1(F), \ldots, h_{|F|-1}(F))$.

Lemma 15. *For any* $F(f, \tau), F'(f', \tau') \in \mathscr{F}_{\mathrm{reg}}$ *such that* $|F| = |F'|$, *if the following conditions (a) and (b) hold, then* $L(F') \leq L(F)$.

(a) $L_i(F) = L_i(F')$ *for any* $i \in [F]$.
(b) $h_{\tau_i(s)}(F) \geq h_{\tau'_i(s)}(F)$ *for any* $i \in [F]$ *and* $s \in \mathcal{S}$.

Using these lemmas, we now prove Theorem 1.

Proof of Theorem 1. We fix an integer $k \geq 0$ arbitrarily and prove Theorem 1 by induction for $|F|$. For the base case $|F| = 1$, the code-tuple $F^\dagger := F$ satisfies (a)–(d) of Theorem 1 as desired. We now consider the induction step for $|F| \geq 2$.

We consider an irreducible part $\bar{F}(\bar{f}, \bar{\tau})$ of F. By Lemmas 9 and 11, the following statements ($\bar{\text{a}}$)–($\bar{\text{c}}$) hold (cf. (a)–(c) of Theorem 1).

($\bar{\text{a}}$) $\bar{F} \in \mathscr{F}_{\mathrm{irr}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$.
($\bar{\text{b}}$) $L(\bar{F}) = L(F)$.
($\bar{\text{c}}$) $\mathscr{P}^k_{\bar{F}} \subseteq \mathscr{P}^k_F$.

Therefore, if $|\mathscr{P}^k_{\bar{F}}| = |\bar{F}|$, then $F^\dagger := \bar{F}$ satisfies (a)–(d) of Theorem 1 as desired. Thus, we now assume $|\mathscr{P}^k_{\bar{F}}| < |\bar{F}|$. Then we can choose $i', j' \in [\bar{F}]$ such that $i' \neq j'$ and $\mathcal{P}^k_{\bar{F},i'} = \mathcal{P}^k_{\bar{F},j'}$ by pigeonhole principle. We define $F'(f', \tau') \in \mathscr{F}^{(|\bar{F}|)}$ as

$$f'_i(s) = \bar{f}_i(s), \tag{35}$$

$$\tau'_i(s) = \begin{cases} p & \text{if } \bar{\tau}_i(s) \in \mathcal{I}, \\ \bar{\tau}_i(s) & \text{if } \bar{\tau}_i(s) \notin \mathcal{I} \end{cases} \tag{36}$$

for $i \in [F']$ and $s \in \mathcal{S}$, where

$$\mathcal{I} := \{i \in [\bar{F}] : \mathcal{P}^k_{\bar{F},i} = \mathcal{P}^k_{\bar{F},i'} (= \mathcal{P}^k_{\bar{F},j'})\} \tag{37}$$

and we choose

$$p \in \arg\min_{i \in \mathcal{I}} h_i(\bar{F}) \tag{38}$$

arbitrarily.

Then we obtain $F' \in \mathscr{F}_{\mathrm{reg}}$ by applying Lemma 13 since $\bar{F} \in \mathscr{F}_{\mathrm{irr}}$. Also, we obtain $F' \in \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ and

$$\mathscr{P}^k_{F'} = \mathscr{P}^k_{\bar{F}} \tag{39}$$

for any $i \in [F']$ by applying Lemma 12 (i)–(iii) since $\bar{f}_i(s) = f'_i(s)$ and $\mathcal{P}^k_{\bar{F},\bar{\tau}_i(s)} = \mathcal{P}^k_{\bar{F},\tau'_i(s)}$ for any $i \in [\bar{F}]$ and $s \in \mathcal{S}$ by (35) and (36). Moreover, we can see

$$L(F') \leq L(\bar{F}) \tag{40}$$

by applying Lemma 15 because F' satisfies (a) (resp. (b)) of Lemma 15 by (35) (resp. (36)–(38)).

Since $|\mathcal{I}| \geq |\{i', j'\}| \geq 2$, we have $\mathcal{I} \setminus \{p\} \neq \emptyset$. Also, for any $i \in \mathcal{I} \setminus \{p\}$, we have $i \notin \mathcal{R}_{F'}$ since for any $j \in [F'] \setminus \{i\}$, there exists no $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau'^*_j(\boldsymbol{x}) = i$ by (36). Therefore, we have

$$\mathcal{R}_{F'} \subsetneq [F']. \tag{41}$$

For an irreducible part $\bar{F}'$ of F', we have

$$|\bar{F}'| = |\mathcal{R}_{F'}| \overset{\text{(A)}}{<} |F'| = |\bar{F}| = |\mathcal{R}_F| \leq |F|, \tag{42}$$

where (A) follows from (41). Therefore, by applying the induction hypothesis to $\bar{F}'$, we can see that there exists $F^\dagger \in \mathscr{F}$ satisfying the following conditions (a†)–(d†).

(a†) $F^\dagger \in \mathscr{F}_{\mathrm{irr}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$.
(b†) $L(F^\dagger) \leq L(\bar{F}')$.
(c†) $\mathscr{P}^k_{F^\dagger} \subseteq \mathscr{P}^k_{\bar{F}'}$.
(d†) $|\mathscr{P}^k_{F^\dagger}| = |F^\dagger|$.

We can see that $F^\dagger$ is a desired code-tuple, that is, $F^\dagger$ satisfies (a)–(d) of Theorem 1 as follows. First, (a) and (d) are directly from (a†) and (d†), respectively. We obtain (b) as follows:

$$L(F^\dagger) \overset{\text{(A)}}{\leq} L(\bar{F}') \overset{\text{(B)}}{=} L(F') \overset{\text{(C)}}{\leq} L(\bar{F}) \overset{\text{(D)}}{=} L(F), \tag{43}$$

where (A) follows from (b†), (B) follows from Lemma 9, (C) follows from (40), and (D) follows from Lemma 9. The condition (c) holds because

$$\mathscr{P}^k_{F^\dagger} \overset{\text{(A)}}{\subseteq} \mathscr{P}^k_{\bar{F}'} \overset{\text{(B)}}{\subseteq} \mathscr{P}^k_{F'} \overset{\text{(C)}}{=} \mathscr{P}^k_{\bar{F}} \overset{\text{(D)}}{\subseteq} \mathscr{P}^k_F,$$

where (A) follows from (c†), (B) follows from Lemma 11, (C) follows from (39), and (D) follows from Lemma 11. □

As a consequence of Theorem 1, we can prove the existence of an *optimal* k-bit delay decodable code-tuple, that is, $F^* \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ such that $L(F^*) \leq L(F)$ for any $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$. We prove this fact in Appendix B.

We define $\mathscr{F}_{k\text{-opt}}$ as the set of all optimal k-bit delay decodable code-tuples as following Definition 16.

Definition 16. *For an integer $k \geq 0$, we define*

$$\mathscr{F}_{k\text{-opt}} := \underset{F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}}{\arg\min} L(F). \tag{44}$$

Note that $\mathscr{F}_{k\text{-opt}}$ depends on the probability distribution μ of the source symbols, and we are now discussing on an arbitrarily fixed μ.

Example 12. *Let $(\mu(\mathrm{a}), \mu(\mathrm{b}), \mu(\mathrm{c}), \mu(\mathrm{d})) = (0.1, 0.2, 0.3, 0.4)$. Then the code-tuple $F^{(\delta)}$ in Table 2 is an optimal 2-bit delay decodable code-tuple with $L(F^{(\delta)}) \approx 1.8667$.*

3.2 Theorem 2

Theorem 2 gives a necessary condition for $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ to be optimal. Recall that every internal node in a code-tree of Huffman code has two child nodes because of its optimality. This leads to that any bit sequence is a prefix of codeword sequence of some source sequence. More formally,

$$^\forall \boldsymbol{b} \in \mathcal{C}^*; {}^\exists \boldsymbol{x} \in \mathcal{S}^*; f_{\mathrm{Huff}}(\boldsymbol{x}) \succeq \boldsymbol{b}, \tag{45}$$

where $f_{\mathrm{Huff}}(\boldsymbol{x})$ is the codeword sequence of $\boldsymbol{x}$ with the Huffman code. The following Theorem 2 is a generalization of this property of Huffman codes to k-bit delay decodable code-tuples for $k \geq 0$.

Theorem 2. *For any integer $k \geq 0$, $F \in \mathscr{F}_{k\text{-opt}}$, $i \in \mathcal{R}_F$, and $\boldsymbol{b} = b_1 b_2 \ldots b_l \in \mathcal{C}^{\geq k}$, if $b_1 b_2 \ldots b_k \in \mathcal{P}^k_{F,i}$, then $\boldsymbol{b} \in \mathcal{P}^*_{F,i}$.*

Remark 3. *A Huffman code is represented by a 1-code-tuple*

Table 3 An example of $\boldsymbol{x} \in \mathcal{S}^*$ such that $f_i^*(\boldsymbol{x}) \succeq \boldsymbol{b}$, where $F(f,\tau) := F^{(\delta)}$ in Table 2, $i \in \{0,1\}$, and $\boldsymbol{b} \in C^3$.

i \ $\boldsymbol{b}$	000	001	010	011	100	101	110	111
0	bb	ba	cb	ca	a	dc	dd	db
1	-	-	cb	ca	db	da	a	ba

$F \in \mathscr{F}^{(1)}$. We have $F \in \mathscr{F}_{0\text{-opt}}$ by the optimality of Huffman codes. Applying Theorem 2 to F with $k = 0$, we obtain

$$^{\forall}\boldsymbol{b} \in C^*; \boldsymbol{b} \in \mathcal{P}^*_{F,0},$$

which is equivalent to (45), and thus Theorem 2 is indeed a generalization of the property (45) of Huffman codes.

Example 13. *For $F(f,\tau) := F^{(\delta)}$ in Table 2, we have $F \in \mathscr{F}_{2\text{-opt}}$ for $(\mu(\mathrm{a}), \mu(\mathrm{b}), \mu(\mathrm{c}), \mu(\mathrm{d})) = (0.1, 0.2, 0.3, 0.4)$ (cf. Example 12). Theorem 2 claims that for any $i \in \mathcal{R}_F = \{0,1\}$ and $\boldsymbol{b} \in C^{\geq 2}$ such that $b_1b_2 \in \mathcal{P}^2_{F,i}$, it holds that $\boldsymbol{b} \in \mathcal{P}^*_{F,i}$, that is, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $f_i^*(\boldsymbol{x}) \succeq \boldsymbol{b}$.*

For $i \in \{0,1\}$ and $\boldsymbol{b} \in C^3$ such that $b_1b_2 \in \mathcal{P}^2_{F,i}$, Table 3 shows an example of $\boldsymbol{x} \in \mathcal{S}^*$ such that $f_i^*(\boldsymbol{x}) \succeq \boldsymbol{b}$. For example, we have $f_0^*(\mathrm{ca}) \succeq 011$ and $f_1^*(\mathrm{ba}) \succeq 111$. Note that $b_1b_2 \in \mathcal{P}^2_{F,i}$ does not hold for $(i,\boldsymbol{b}) = (1, 000)$ and $(i,\boldsymbol{b}) = (1, 001)$.

Proof of Theorem 2. We prove by contradiction assuming that there exist $p \in \mathcal{R}_F$ and $\boldsymbol{b} = b_1b_2\ldots b_l \in C^{\geq k}$ such that

$$\boldsymbol{b} \notin \mathcal{P}^*_{F,p}, \quad b_1b_2\ldots b_k \in \mathcal{P}^k_{F,p}. \tag{46}$$

Without loss of generality, we assume $p = |F|-1$ and $\boldsymbol{b}$ is the shortest sequence satisfying (46). Because we have $l > k$ by (46), we have $\mathrm{pref}(\boldsymbol{b}) \succeq b_1b_2\ldots b_k \in \mathcal{P}^k_{F,|F|-1}$. Since $\boldsymbol{b}$ is the shortest sequence satisfying (46), it must hold that $\mathrm{pref}(\boldsymbol{b}) \in \mathcal{P}^*_{F,|F|-1}$. Hence, by $F \in \mathscr{F}_{\mathrm{ext}}$ and Lemma 4 (i), we have $\boldsymbol{d} = d_1d_2\ldots d_l := \mathrm{pref}(\boldsymbol{b})\bar{b_l} \in \mathcal{P}^*_{F,|F|-1}$, where $\bar{c}$ denotes the negation of $c \in C$, that is, $\bar{0} := 1$ and $\bar{1} := 0$. Namely, we have

$$\boldsymbol{d} \in \mathcal{P}^*_{F,|F|-1}, \quad \mathrm{pref}(\boldsymbol{d})\bar{d_l} = \boldsymbol{b} \notin \mathcal{P}^*_{F,|F|-1}. \tag{47}$$

We state the key idea of the proof as follows. By (47), whenever the decoder reads a prefix $\mathrm{pref}(\boldsymbol{d})$ of the codeword sequence, the decoder can know that the following bit is d_l without reading it. Hence, the bit d_l gives no information and is unnecessary for the k-bit delay decodability of the mapping $f^*_{|F|-1}$. We consider obtaining another code-tuple $F'' \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ such that $L(F'') < L(F)$ by removing this redundant bit, which leads to a contradiction to $F \in \mathscr{F}_{k\text{-opt}}$ as desired. However, naive removing a bit may impair the k-bit delay decodability of the other mappings f_i^* for $i \in [|F|-1]$. Accordingly, we first define a code-tuple F' which is essentially equivalent to F by adding some duplicates of the code tables to F. Then by making changes to the replicated code tables instead of the original code tables, we obtain the desired F'' without affecting the k-bit delay decodability of f_i^* for $i \in [|F|-1]$.

We define the code-tuple F' as follows. Put $L := |F|(|\boldsymbol{d}|+1)$ and $M := |\mathcal{S}^{\leq L}|$. We number all the sequences of $\mathcal{S}^{\leq L}$ as $\boldsymbol{z}^{(0)}, \boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}, \ldots, \boldsymbol{z}^{(M-1)}$ in any order but $\boldsymbol{z}^{(0)} := \lambda$. For $\boldsymbol{z}' \in \mathcal{S}^{\leq L}$, we define $\langle \boldsymbol{z}' \rangle := |F|-1+t$, where t is the integer such that $\boldsymbol{z}^{(t)} = \boldsymbol{z}'$. Note that $\langle\lambda\rangle = |F|-1$ since $\boldsymbol{z}^{(0)} = \lambda$. We define the code-tuple $F' \in \mathscr{F}^{(|F|-1+M)}$ consisting of $f'_0, f'_1, \ldots, f'_{|F|-1}, f'_{\langle \boldsymbol{z}^{(1)}\rangle}, f'_{\langle \boldsymbol{z}^{(2)}\rangle}, \ldots f'_{\langle \boldsymbol{z}^{(M-1)}\rangle}$ and $\tau'_0, \tau'_1, \ldots, \tau'_{|F|-1}, \tau'_{\langle \boldsymbol{z}^{(1)}\rangle}, \tau'_{\langle \boldsymbol{z}^{(2)}\rangle}, \ldots \tau'_{\langle \boldsymbol{z}^{(M-1)}\rangle}$ as

$$f'_i(s) = \begin{cases} f_{\tau^*_{\langle\lambda\rangle}(\boldsymbol{z})}(s) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{\leq L}, \\ f_i(s) & \text{otherwise,} \end{cases} \tag{48}$$

$$\tau'_i(s) = \begin{cases} \langle \boldsymbol{z}s \rangle & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{\leq L-1}, \\ \tau^*_{\langle\lambda\rangle}(\boldsymbol{z}s) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{L}, \\ \tau_i(s) & \text{otherwise} \end{cases} \tag{49}$$

for $i \in [F']$ and $s \in \mathcal{S}$. Then F' satisfies the following Lemma 16. See Appendix C.9 for the proof of Lemma 16.

Lemma 16. *For any $\boldsymbol{z} \in \mathcal{S}^{\leq L}$, the following statements (i) and (ii) hold.*

(i) $\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}) = \langle \boldsymbol{z} \rangle$.
(ii) $\langle \boldsymbol{z} \rangle \in \mathcal{R}_{F'}$.

Lemma 16 (i) claims that the code table in F' used next after encoding $\boldsymbol{z} \in \mathcal{S}^{\leq L}$ starting from $f'_{\langle\lambda\rangle}$ is $f'_{\langle \boldsymbol{z}\rangle}$, which is a duplicate of the code table in F used next after encoding $\boldsymbol{z}$ starting from $f_{\langle\lambda\rangle}$. This leads to the equivalency of F and F' shown next.

We confirm that F' is equivalent to F, that is, $F' \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ and $L(F') = L(F)$. We obtain $F' \in \mathscr{F}_{\mathrm{reg}}$ from Lemma 16 (ii) and Lemma 8 (i). To prove $F' \in \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ and $L(F') = L(F)$ by using Lemma 7, we show that a mapping $\varphi : [F'] \to [F]$ defined as the following (50) is a homomorphism:

$$\varphi(i) = \begin{cases} i & \text{if } i \in [F], \\ \tau^*_{\langle\lambda\rangle}(\boldsymbol{z}) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{\leq L} \end{cases} \tag{50}$$

for $i \in [F']$. The case $i = |F|-1 = \langle\lambda\rangle$ applies to both of the first and second cases of (50). However, this case is consistent since $\tau^*_{\langle\lambda\rangle}(\boldsymbol{z}) = \tau^*_{\langle\lambda\rangle}(\lambda) = \langle\lambda\rangle = i$. We see that φ satisfies (23) directly from (48) and (50). We confirm that φ satisfies also (24) as follows:

$$\begin{aligned} &\varphi(\tau'_i(s)) \\ &\overset{\text{(A)}}{=} \begin{cases} \varphi(\langle \boldsymbol{z}s \rangle) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{\leq L-1}, \\ \varphi(\tau^*_{\langle\lambda\rangle}(\boldsymbol{z}s)) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{L}, \\ \varphi(\tau_i(s)) & \text{otherwise,} \end{cases} \\ &\overset{\text{(B)}}{=} \begin{cases} \tau^*_{\langle\lambda\rangle}(\boldsymbol{z}s) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{\leq L-1}, \\ \tau^*_{\langle\lambda\rangle}(\boldsymbol{z}s) & \text{if } i = \langle \boldsymbol{z} \rangle \text{ for some } \boldsymbol{z} \in \mathcal{S}^{L}, \\ \tau_i(s) & \text{otherwise,} \end{cases} \end{aligned}$$

$$\overset{\text{(C)}}{=} \begin{cases} \tau_{\tau^*_{\langle\lambda\rangle}(z)}(s) & \text{if } i = \langle z\rangle \text{ for some } z \in \mathcal{S}^{\leq L}, \\ \tau_i(s) & \text{otherwise,} \end{cases}$$
$$\overset{\text{(D)}}{=} \tau_{\varphi(i)}(s),$$

where (A) follows from (49), (B) follows from (50), (C) follows from Lemma 1 (ii), and (D) follows from (50). Hence, by Lemma 7 (iv)–(vi), we obtain $F' \in \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ and $L(F') = L(F)$.

Now, we define a code-tuple $F'' \in \mathscr{F}^{(|F'|)}$ as

$$f''_i(s) = \begin{cases} f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(zs)) \\ \quad \text{if } i = \langle z\rangle \text{ and } f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs) \\ \quad \text{for some } z \in \mathcal{S}^{\leq L}, \\ f'_i(s) \qquad \text{otherwise,} \end{cases} \tag{51}$$

$$\tau''_i(s) = \tau'_i(s) \tag{52}$$

for $i \in [F'']$ and $s \in \mathcal{S}$.

Intuitively, (51) means that F'' is obtained by removing the bit d_l from codeword sequences of F' such that $f'^*_{\langle\lambda\rangle}(z) \succeq \boldsymbol{d}$.

Then F'' satisfies the following Lemma 17. See Appendix C.10 for the proof of Lemma 17.

Lemma 17. *The following statements (i)–(iii) hold.*

(i) For any $z \in \mathcal{S}^{\leq L}$ and $\boldsymbol{x} \in \mathcal{S}^{\leq L-|z|}$, we have

$$f''^*_{\langle z\rangle}(\boldsymbol{x}) = \begin{cases} f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})) \\ \quad \text{if } f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}), \\ f'^*_{\langle z\rangle}(\boldsymbol{x}) \qquad \text{otherwise.} \end{cases} \tag{53}$$

(ii) For any $z \in \mathcal{S}^{\leq L}$ and $s, s' \in \mathcal{S}$, if $f''_{\langle z\rangle}(s) \prec f''_{\langle z\rangle}(s')$, then $f'_{\langle z\rangle}(s) \prec f'_{\langle z\rangle}(s')$.

(iii) For any $\boldsymbol{x} \in \mathcal{S}^{\geq L}$, we have $|f^*_{\langle\lambda\rangle}(\boldsymbol{x})| = |f'^*_{\langle\lambda\rangle}(\boldsymbol{x})| \geq |\boldsymbol{d}|+1$ and $|f''^*_{\langle\lambda\rangle}(\boldsymbol{x})| \geq |\boldsymbol{d}|$.

We show that $F'' \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ and $L(F'') < L(F')$ $(= L(F)$ as shown above), which conflicts with $F \in \mathscr{F}_{k\text{-opt}}$ and completes the proof of Theorem 2.

(Proof of $F'' \in \mathscr{F}_{\mathrm{reg}}$): From $F' \in \mathscr{F}_{\mathrm{reg}}$ and (52).

(Proof of $F'' \in \mathscr{F}_{\mathrm{ext}}$): Choose $j \in [F'']$ arbitrarily. Since $\langle\lambda\rangle \in \mathcal{R}_{F'} = \mathcal{R}_{F''}$ by Lemma 16 (ii) and (52), there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that

$$\tau''^*_j(\boldsymbol{x}) = \langle\lambda\rangle. \tag{54}$$

Also, we can choose $\boldsymbol{x}' \in \mathcal{S}^L$ such that

$$f'^*_{\langle\lambda\rangle}(\boldsymbol{x}') \succeq \boldsymbol{d} \tag{55}$$

by Lemma 17 (iii). We have

$$\begin{aligned} |f''^*_j(\boldsymbol{x}\boldsymbol{x}')| &\overset{\text{(A)}}{=} |f''^*_j(\boldsymbol{x})| + |f''^*_{\tau''^*_j(\boldsymbol{x})}(\boldsymbol{x}')| \\ &\geq |f''^*_{\tau''^*_j(\boldsymbol{x})}(\boldsymbol{x}')| \\ &\overset{\text{(B)}}{=} |f''^*_{\langle\lambda\rangle}(\boldsymbol{x}')| \\ &\overset{\text{(C)}}{=} |f'^*_{\langle\lambda\rangle}(\lambda)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(\boldsymbol{x}')| \\ &= |f'^*_{\langle\lambda\rangle}(\boldsymbol{x}')| - 1 \\ &\overset{\text{(D)}}{\geq} |\boldsymbol{d}| \\ &\geq 1, \end{aligned}$$

where (A) follows from Lemma 1 (i), (B) follows from (54), (C) follows from (55) and the first case of (53), and (D) follows from Lemma 17 (iii). Hence, by (8), $\mathcal{P}^1_{F'',j} \neq \emptyset$ holds for any $j \in [F'']$, which leads to $F'' \in \mathscr{F}_{\mathrm{ext}}$ as desired.

(Proof of $L(F'') < L(F')$): For any $i \in [F'']$ and $s \in \mathcal{S}$, we have $|f''_i(s)| \leq |f'_i(s)|$ by (51). Hence, for any $i \in [F'']$, we have

$$\pi_i(F')L_i(F'') \leq \pi_i(F')L_i(F'). \tag{56}$$

By Lemma 17 (iii), we can choose $\boldsymbol{x} = x_1x_2\ldots x_L \in \mathcal{S}^L$ such that $f'^*_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \boldsymbol{d}$. Since $f'^*_{\langle\lambda\rangle}(\lambda) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{x})$, there exists exactly one integer r such that

$$f'^*_{\langle\lambda\rangle}(x_1x_2\ldots x_{r-1}) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(x_1x_2\ldots x_r), \tag{57}$$

which leads to

$$\begin{aligned} |f''_{\langle z\rangle}(x_r)| &\overset{\text{(A)}}{=} |f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(zx_r))| \\ &= |f'_{\langle z\rangle}(x_r)| - 1 \\ &< |f'_{\langle z\rangle}(x_r)|, \end{aligned} \tag{58}$$

where $z \coloneqq x_1x_2\ldots x_{r-1}$, and (A) follows from (57) and the first case of (51). This leads to

$$\pi_{\langle z\rangle}(F')L_{\langle z\rangle}(F'') < \pi_{\langle z\rangle}(F')L_{\langle z\rangle}(F') \tag{59}$$

because $\pi_{\langle z\rangle}(F') > 0$ by Lemma 16 (ii) and Lemma 8 (ii).

Hence, we have

$$\begin{aligned} &L(F'') \\ &= \sum_{i\in[F'']} \pi_i(F'')L_i(F'') \\ &\overset{\text{(A)}}{=} \sum_{i\in[F'']} \pi_i(F')L_i(F'') \\ &= \sum_{i\in[F'']\setminus\{\langle z\rangle\}} \pi_i(F')L_i(F'') + \pi_{\langle z\rangle}(F')L_{\langle z\rangle}(F'') \\ &\overset{\text{(B)}}{\leq} \sum_{i\in[F'']\setminus\{\langle z\rangle\}} \pi_i(F')L_i(F') + \pi_{\langle z\rangle}(F')L_{\langle z\rangle}(F'') \\ &\overset{\text{(C)}}{<} \sum_{i\in[F'']\setminus\{\langle z\rangle\}} \pi_i(F')L_i(F') + \pi_{\langle z\rangle}(F')L_{\langle z\rangle}(F') \\ &= \sum_{i\in[F']} \pi_i(F')L_i(F') \\ &= L(F') \end{aligned}$$

as desired, where (A) follows from (52), (B) follows from

(56), and (C) follows from (59).

(Proof of $F'' \in \mathscr{F}_{k\text{-dec}}$): To prove $F'' \in \mathscr{F}_{k\text{-dec}}$, we use the following Lemma 18, where $\mathcal{J} := ([F'] \setminus \langle\lambda\rangle) \cup \{\langle z\rangle : z \in \mathcal{S}^L\} = [F'] \setminus \{\langle z\rangle : z \in \mathcal{S}^{\leq L-1}\}$. See Appendix C.11 for the proof of Lemma 18.

Lemma 18. *The following statements (i)–(iii) hold.*

(i) For any $\boldsymbol{x} \in \mathcal{S}^*$ and $\boldsymbol{c} \in \mathcal{C}^{\leq k}$, if $f''^*_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \boldsymbol{c}$, then $f'^*_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \boldsymbol{c}$. Therefore, we have $\mathcal{P}^k_{F',\langle\lambda\rangle} \supseteq \mathcal{P}^k_{F'',\langle\lambda\rangle}$ by (8).

(ii) For any $i \in \mathcal{J}$ and $s \in \mathcal{S}$, we have $f''_i(s) = f'_i(s)$.

(iii) For any $i \in \mathcal{J}$ and $\boldsymbol{b} \in \mathcal{C}^*$, we have $\mathcal{P}^k_{F'',i}(\boldsymbol{b}) \subseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$ and $\bar{\mathcal{P}}^k_{F'',i}(\boldsymbol{b}) \subseteq \bar{\mathcal{P}}^k_{F',i}(\boldsymbol{b})$.

Also, for $z \in \mathcal{S}^*$, we define a mapping $\psi_z : \mathcal{C}^* \to \mathcal{C}^*$ as

$$\psi_z(\boldsymbol{b}) = \begin{cases} f'^*_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\,\mathrm{pref}(\boldsymbol{d})^{-1}(f'^*_{\langle\lambda\rangle}(z)\boldsymbol{b}) \\ \quad \text{if } f'^*_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z)\boldsymbol{b}, \\ \boldsymbol{b} \qquad\qquad \text{otherwise} \end{cases} \tag{60}$$

for $\boldsymbol{b} \in \mathcal{C}^*$. Then ψ_z satisfies the following Lemma 19.

Lemma 19. *The following statements (i)–(iii) hold.*

(i) For any $z \in \mathcal{S}^*$ and $\boldsymbol{b}, \boldsymbol{b}' \in \mathcal{C}^*$, if $\boldsymbol{b} \preceq \boldsymbol{b}'$, then $\psi_z(\boldsymbol{b}) \preceq \psi_z(\boldsymbol{b}')$.

(ii) For any $z \in \mathcal{S}^{\leq L}$, $\boldsymbol{x} \in \mathcal{S}^{\leq L-|z|}$, and $\boldsymbol{c} \in \mathcal{C}^*$, we have

$$\psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) = \begin{cases} \mathrm{pref}(f'^*_{\langle z\rangle}(\boldsymbol{x})) \\ \quad \textit{if } f'^*_{\langle\lambda\rangle}(z) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}, \boldsymbol{c} = \lambda, \\ f'^*_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(\boldsymbol{c}) \ \textit{otherwise}. \end{cases} \tag{61}$$

(iii) For any $z \in \mathcal{S}^L$ and $\boldsymbol{b} \in \mathcal{C}^*$, we have $\psi_z(\boldsymbol{b}) = \boldsymbol{b}$.

See Appendix C.12 for the proof of Lemma 19.

By Lemma 19 (ii) with $\boldsymbol{c} = \lambda$, it holds that $\psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x})) = f'^*_{\langle z\rangle}(\boldsymbol{x})$ in most cases. Thus, we can intuitively interpret the mapping ψ_z as a kind of an inverse transformation of (53). We prove k-bit delay decodability of F'' later by attributing it to k-bit delay decodability of F' using ψ_z.

Now we prove $F'' \in \mathscr{F}_{k\text{-dec}}$. We first show that F'' satisfies Definition 5 (i). Namely, we show that $\mathcal{P}^k_{F'',\tau''_i(s)} \cap \bar{\mathcal{P}}^k_{F'',i}(f''_i(s)) = \emptyset$ for any $i \in [F'']$ and $s \in \mathcal{S}$ dividing into the following two cases: the case $i \in \mathcal{J}$ and the case $i \in [F''] \setminus \mathcal{J}$.

- The case $i \in \mathcal{J}$: Then for any $i \in \mathcal{J}$ and $s \in \mathcal{S}$, we have

$$\begin{aligned}\mathcal{P}^k_{F'',\tau''_i(s)} \cap \bar{\mathcal{P}}^k_{F'',i}(f''_i(s)) &\overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F',\tau''_i(s)} \cap \bar{\mathcal{P}}^k_{F',i}(f''_i(s)) \\ &\overset{\text{(B)}}{=} \mathcal{P}^k_{F',\tau'_i(s)} \cap \bar{\mathcal{P}}^k_{F',i}(f'_i(s)) \\ &\overset{\text{(C)}}{=} \emptyset,\end{aligned}$$

where (A) follows from Lemma 18 (i) and (iii) since $\tau''_i(s) \in [F]$, (B) follows from Lemma 18 (ii) and (52), and (C) follows from $F' \in \mathscr{F}_{k\text{-dec}}$.

- The case $i \in [F''] \setminus \mathcal{J}$: We prove by contradiction assuming that there exist $z \in \mathcal{S}^{\leq L-1}$, $s \in \mathcal{S}$, and $\boldsymbol{c} \in \bar{\mathcal{P}}^k_{F'',\langle z\rangle}(f''_{\langle z\rangle}(s)) \cap \mathcal{P}^k_{F'',\langle zs\rangle}$. By $\boldsymbol{c} \in \bar{\mathcal{P}}^k_{F'',\langle z\rangle}(f''_{\langle z\rangle}(s))$ and (5), there exist $\boldsymbol{x} \in \mathcal{S}^{L-|z|}$ and $\boldsymbol{y} \in \mathcal{S}^*$ such that

$$f''^*_{\langle z\rangle}(\boldsymbol{x}\boldsymbol{y}) \succeq f''_{\langle z\rangle}(s)\boldsymbol{c} \tag{62}$$

and

$$f''_{\langle z\rangle}(x_1) \succ f''_{\langle z\rangle}(s). \tag{63}$$

By Lemma 3, we may assume

$$|f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})| \geq \max\{k, 1\}. \tag{64}$$

By (63) and Lemma 17 (ii), we obtain

$$f'_{\langle z\rangle}(x_1) \succ f'_{\langle z\rangle}(s). \tag{65}$$

This shows that $f'_{\langle z\rangle}$ is not prefix-free, which conflicts with $F' \in \mathscr{F}_{k\text{-dec}}$ in the case $k = 0$ by Lemma 2. Thus, we consider the case $k \geq 1$, that is,

$$\boldsymbol{c} \neq \lambda. \tag{66}$$

Equation (62) leads to

$$\begin{aligned} & f''^*_{\langle z\rangle}(\boldsymbol{x}\boldsymbol{y}) \succeq f''_{\langle z\rangle}(s)\boldsymbol{c} \\ \overset{\text{(A)}}{\Longrightarrow}\ & \psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x}\boldsymbol{y})) \succeq \psi_z(f''_{\langle z\rangle}(s)\boldsymbol{c}) \\ \overset{\text{(B)}}{\Longleftrightarrow}\ & \psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x})f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})) \succeq \psi_z(f''_{\langle z\rangle}(s)\boldsymbol{c}) \\ \overset{\text{(C)}}{\Longleftrightarrow}\ & f'^*_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})) \succeq f'_{\langle z\rangle}(s)\psi_{zs}(\boldsymbol{c}) \\ \overset{\text{(D)}}{\Longleftrightarrow}\ & f'^*_{\langle z\rangle}(\boldsymbol{x})f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y}) \succeq f'_{\langle z\rangle}(s)\psi_{zs}(\boldsymbol{c}), \end{aligned} \tag{67}$$

where (A) follows from Lemma 19 (i), (B) follows from Lemma 1 (i) and Lemma 16 (i), (C) follows from (64), (66), and the second case of (61), and (D) follows from Lemma 19 (iii) and $|z\boldsymbol{x}| = L$.

Now, for $\boldsymbol{b} \in \mathcal{C}^{\geq k}$, let $[\boldsymbol{b}]_k$ denote the prefix of length k of $\boldsymbol{b}$. Then by (65) and (67), we have

$$f'^*_{\langle z\rangle}(\boldsymbol{x})[f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \succeq f'_{\langle z\rangle}(s)[\psi_{zs}(\boldsymbol{c})]_k. \tag{68}$$

Also, we have

$$[f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \in \mathcal{P}^k_{F'',\langle z\boldsymbol{x}\rangle} \overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F',\langle z\boldsymbol{x}\rangle}, \tag{69}$$

where (A) follows from Lemma 18 (iii) and $\langle z\boldsymbol{x}\rangle \in \mathcal{S}^L \subseteq \mathcal{J}$. Hence, by (8) there exists $\boldsymbol{y}' \in \mathcal{S}^*$ such that $f'^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y}') \succeq [f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})]_k$, which leads to

$$\begin{aligned} f'^*_{\langle z\rangle}(\boldsymbol{x}\boldsymbol{y}') &= f'^*_{\langle z\rangle}(\boldsymbol{x})f'^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y}') \\ &\succeq f'^*_{\langle z\rangle}(\boldsymbol{x})[f''^*_{\langle z\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \\ &\overset{\text{(A)}}{\succeq} f'_{\langle z\rangle}(s)[\psi_{zs}(\boldsymbol{c})]_k, \end{aligned} \tag{70}$$

where (A) follows from (68). Equations (65) and (70) show

$$[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k \in \bar{\mathcal{P}}^k_{F',\langle \boldsymbol{z}\rangle}(f'_{\langle \boldsymbol{z}\rangle}(s)) \tag{71}$$

by (5).

On the other hand, by $\boldsymbol{c} \in \mathcal{P}^k_{F'',\langle \boldsymbol{z}s\rangle}$ and (8), there exist $\boldsymbol{x} \in \mathcal{S}^{L-|\boldsymbol{z}s|}$ and $\boldsymbol{y} \in \mathcal{S}^*$ such that

$$f''^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x}\boldsymbol{y}) \succeq \boldsymbol{c}. \tag{72}$$

By Lemma 3, we may assume

$$|f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y})| \geq k \geq 1. \tag{73}$$

We have

$$\begin{aligned} f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x})f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y}) &\overset{\text{(A)}}{=} f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x})\psi_{\boldsymbol{z}s\boldsymbol{x}}(f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y}))\\ &\overset{\text{(B)}}{=} \psi_{\boldsymbol{z}s}(f''^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x}\boldsymbol{y}))\\ &\overset{\text{(C)}}{\succeq} \psi_{\boldsymbol{z}s}(\boldsymbol{c}), \end{aligned}$$

where (A) follows from Lemma 19 (iii) and $|\boldsymbol{z}s\boldsymbol{x}| = L$, (B) follows from (73) and the second case of (61), and (C) follows from (72) and Lemma 19 (i).

Hence, we have

$$f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x})[f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \succeq [\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k. \tag{74}$$

Also, we have

$$[f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \in \mathcal{P}^k_{F'',\langle \boldsymbol{z}s\boldsymbol{x}\rangle} \overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F',\langle \boldsymbol{z}s\boldsymbol{x}\rangle}, \tag{75}$$

where (A) follows from Lemma 18 (iii) and $\langle \boldsymbol{z}s\boldsymbol{x}\rangle \in \mathcal{S}^L \subseteq \mathcal{J}$. Hence, there exists $\boldsymbol{y}' \in \mathcal{S}^*$ such that $f'^*_{\langle \boldsymbol{z}\boldsymbol{x}\rangle}(\boldsymbol{y}') \succeq [f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y})]_k$, which leads to

$$\begin{aligned} f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x}\boldsymbol{y}') &= f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x})f'^*_{\langle \boldsymbol{z}\boldsymbol{x}s\rangle}(\boldsymbol{y}')\\ &\succeq f'^*_{\langle \boldsymbol{z}s\rangle}(\boldsymbol{x})[f''^*_{\langle \boldsymbol{z}s\boldsymbol{x}\rangle}(\boldsymbol{y})]_k\\ &\overset{\text{(A)}}{\succeq} [\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k, \end{aligned} \tag{76}$$

where (A) follows from (74). This shows

$$[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k \in \mathcal{P}^k_{F',\langle \boldsymbol{z}s\rangle} \tag{77}$$

by (8). By (71) and (77), the code-tuple F' does not satisfy Definition 5 (i), which conflicts with $F' \in \mathscr{F}_{k\text{-dec}}$.

Consequently, F'' satisfies Definition 5 (i).

Next, we show that F'' satisfies Definition 5 (ii). Namely, we show that for any $i \in [F'']$ and $s, s' \in \mathcal{S}$ such that $s \neq s'$ and $f''_i(s) = f''_i(s')$, we have $\mathcal{P}^k_{F',\tau'_i(s)} \cap \mathcal{P}^k_{F',\tau'_i(s')} = \emptyset$. We prove for the following two cases: the case $i \in \mathcal{J}$ and the case $i \in [F''] \setminus \mathcal{J}$.

- The case $i \in \mathcal{J}$: Then for any $i \in \mathcal{J}$ and $s, s' \in \mathcal{S}$ such that $s \neq s'$ and $f''_i(s) = f''_i(s')$, we have

$$f'_i(s) = f'_i(s') \tag{78}$$

by Lemma 18 (ii), and we have

$$\begin{aligned} \mathcal{P}^k_{F'',\tau''_i(s)} \cap \mathcal{P}^k_{F'',\tau''_i(s')} &\overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F',\tau''_i(s)} \cap \mathcal{P}^k_{F',\tau''_i(s')}\\ &\overset{\text{(B)}}{=} \mathcal{P}^k_{F',\tau'_i(s)} \cap \mathcal{P}^k_{F',\tau'_i(s')} \overset{\text{(C)}}{=} \emptyset, \end{aligned}$$

where (A) follows from Lemma 18 (i) and (iii) since $\tau''_i(s), \tau''_i(s') \in [F]$, (B) follows from (52), and (C) follows from $F' \in \mathscr{F}_{k\text{-dec}}$ and (78).

- The case $i \in [F''] \setminus \mathcal{J}$: We prove by contradiction assuming that there exists $\boldsymbol{z} \in \mathcal{S}^{\leq L-1}, s, s' \in \mathcal{S}$, and $\boldsymbol{c} \in \mathcal{P}^k_{F'',\langle \boldsymbol{z}s\rangle} \cap \mathcal{P}^k_{F'',\langle \boldsymbol{z}s'\rangle}$ such that $s \neq s'$ and

$$f''_{\langle \boldsymbol{z}\rangle}(s) = f''_{\langle \boldsymbol{z}\rangle}(s'). \tag{79}$$

By the similar way to derive (77), we obtain

$$[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k \in \mathcal{P}^k_{F',\langle \boldsymbol{z}s\rangle} \tag{80}$$

from $\boldsymbol{c} \in \mathcal{P}^k_{F'',\langle \boldsymbol{z}s\rangle}$. By (79) and Lemma 19 (i), we have

$$\psi_{\langle \boldsymbol{z}\rangle}(f''_{\langle \boldsymbol{z}\rangle}(s)) = \psi_{\langle \boldsymbol{z}\rangle}(f''_{\langle \boldsymbol{z}\rangle}(s')). \tag{81}$$

By Lemma 19 (ii), exactly one of $f'_{\langle \boldsymbol{z}\rangle}(s) = f'_{\langle \boldsymbol{z}\rangle}(s')$, $f'_{\langle \boldsymbol{z}\rangle}(s) \prec f'_{\langle \boldsymbol{z}\rangle}(s')$, and $f'_{\langle \boldsymbol{z}\rangle}(s) \succ f'_{\langle \boldsymbol{z}\rangle}(s')$ holds. Therefore, $f'_{\langle \boldsymbol{z}\rangle}$ is not prefix-free, which conflicts with $F' \in \mathscr{F}_{k\text{-dec}}$ in the case $k = 0$ by Lemma 2. We consider the case $k \geq 1$, that is,

$$\boldsymbol{c} \neq \lambda. \tag{82}$$

We consider the following two cases separately: the case $f'_{\langle \boldsymbol{z}\rangle}(s) = f'_{\langle \boldsymbol{z}\rangle}(s')$ and the case $f'_{\langle \boldsymbol{z}\rangle}(s) \prec f'_{\langle \boldsymbol{z}\rangle}(s')$. Note that we may exclude the case $f'_{\langle \boldsymbol{z}\rangle}(s) \succ f'_{\langle \boldsymbol{z}\rangle}(s')$ by symmetry.

  - The case $f'_{\langle \boldsymbol{z}\rangle}(s) = f'_{\langle \boldsymbol{z}\rangle}(s')$: By (60), we have $\psi_{\boldsymbol{z}s}(\boldsymbol{c}) = \psi_{\boldsymbol{z}s'}(\boldsymbol{c})$ and thus

$$[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k = [\psi_{\boldsymbol{z}s'}(\boldsymbol{c})]_k \overset{\text{(A)}}{\in} \mathcal{P}^k_{F',\langle \boldsymbol{z}s'\rangle}, \tag{83}$$

  where (A) is obtained from $\boldsymbol{c} \in \mathcal{P}^k_{F'',\langle \boldsymbol{z}s'\rangle}$ by the similar way to derive (77).
  By (80), (83), and $f'_{\langle \boldsymbol{z}\rangle}(s) = f'_{\langle \boldsymbol{z}\rangle}(s')$, the code-tuple F' does not satisfy Definition 5 (ii), which conflicts with $F' \in \mathscr{F}_{k\text{-dec}}$.
  - The case $f'_{\langle \boldsymbol{z}\rangle}(s) \prec f'_{\langle \boldsymbol{z}\rangle}(s')$: Then by (81) and Lemma 19 (ii), it must hold that

$$f'^*_{\langle \lambda\rangle}(\boldsymbol{z}) \prec f'^*_{\langle \lambda\rangle}(\boldsymbol{z}s') = \boldsymbol{d} \tag{84}$$

  and

$$f'_{\langle \boldsymbol{z}\rangle}(s) = \mathrm{pref}(f'_{\langle \boldsymbol{z}\rangle}(s')). \tag{85}$$

  Thus, we have

$$f'_{\langle \boldsymbol{z}\rangle}(s)d_l \overset{\text{(A)}}{=} \mathrm{pref}(f'_{\langle \boldsymbol{z}\rangle}(s'))d_l$$

$$\begin{aligned}
&= f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})\mathrm{pref}(f'_{\langle\boldsymbol{z}\rangle}(s'))d_l \\
&\overset{\text{(B)}}{=} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s'))d_l \\
&\overset{\text{(C)}}{=} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\mathrm{pref}(\boldsymbol{d})d_l \\
&= f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\boldsymbol{d} \\
&\overset{\text{(D)}}{=} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s') \\
&\overset{\text{(E)}}{=} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z})^{-1} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}) f'_{\langle\boldsymbol{z}\rangle}(s') \\
&= f'_{\langle\boldsymbol{z}\rangle}(s'), \qquad (86)
\end{aligned}$$

where (A) follows from (85), (B) follows from Lemma 1 (i) and Lemma 16 (i), (C) follows from (84), (D) follows from (84), and (E) follows from Lemma 1 (i) and Lemma 16 (i).
Also, we have

$$\begin{aligned}
\mathrm{pref}(\boldsymbol{d}) &\overset{\text{(A)}}{=} \mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s')) \\
&= \mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}) f'_{\langle\boldsymbol{z}\rangle}(s')) \\
&\overset{\text{(B)}}{=} \mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}) f'_{\langle\boldsymbol{z}\rangle}(s)d_l) \\
&= f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}) f'_{\langle\boldsymbol{z}\rangle}(s) \\
&= f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s), \qquad (87)
\end{aligned}$$

where (A) follows from (84), and (B) follows from (86).
By $\boldsymbol{c} \in \mathcal{P}^k_{F'',\langle\boldsymbol{z}s'\rangle}$ and (8), there exist $\boldsymbol{x} \in \mathcal{S}^{L-|\boldsymbol{z}s'|}$ and $\boldsymbol{y} \in \mathcal{S}^*$ such that

$$f''^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x}\boldsymbol{y}) \succeq \boldsymbol{c}. \qquad (88)$$

By Lemma 3, we may assume

$$|f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y})| \geq k \geq 1. \qquad (89)$$

We have

$$\begin{aligned}
& f'_{\langle\boldsymbol{z}\rangle}(s') f'^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x}) f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y}) \\
&\overset{\text{(A)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s')\psi_{\boldsymbol{z}s'}(f''^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x}) f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y})) \\
&\overset{\text{(B)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s')\psi_{\boldsymbol{z}s'}(f''^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x}\boldsymbol{y})) \\
&\overset{\text{(C)}}{\succeq} f'_{\langle\boldsymbol{z}\rangle}(s')\psi_{\boldsymbol{z}s'}(\boldsymbol{c}) \\
&\overset{\text{(D)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s)d_l\psi_{\boldsymbol{z}s'}(\boldsymbol{c}) \\
&\overset{\text{(E)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s)d_l\boldsymbol{c} \\
&= f'_{\langle\boldsymbol{z}\rangle}(s)\mathrm{pref}(\boldsymbol{d})^{-1}\boldsymbol{d}\,\mathrm{pref}(\boldsymbol{d})^{-1}(\mathrm{pref}(\boldsymbol{d})\boldsymbol{c}) \\
&\overset{\text{(F)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s) f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s)^{-1}\boldsymbol{d}\,\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s)\boldsymbol{c}) \\
&\overset{\text{(G)}}{=} f'_{\langle\boldsymbol{z}\rangle}(s)\psi_{\boldsymbol{z}s}(\boldsymbol{c}),
\end{aligned}$$

where (A) follows from (89) and the second case of (61), (B) follows from Lemma 1 (i) and Lemma 16 (i), (C) follows from (88) and Lemma 19 (i), (D) follows from (86), (E) follows from the second case of (60) because $f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s') \preceq \mathrm{pref}(\boldsymbol{d})$ does not hold by (84), (F) follows from (87), and (G) follows from the first case of (60) because $f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s) = \mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s')) = \mathrm{pref}(\boldsymbol{d}) \prec f'^{*}_{\langle\lambda\rangle}(\boldsymbol{z}s)\boldsymbol{c}$ by (84), (85), and (82).
Hence, by $f'_{\langle\boldsymbol{z}\rangle}(s) \prec f'_{\langle\boldsymbol{z}\rangle}(s')$, we have

$$f'_{\langle\boldsymbol{z}\rangle}(s') f'^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x})[f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \succeq f'_{\langle\boldsymbol{z}\rangle}(s)[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k. \qquad (90)$$

Also, we have

$$[f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y})]_k \in \mathcal{P}^k_{F'',\langle\boldsymbol{z}s'\boldsymbol{x}\rangle} \overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F',\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}, \qquad (91)$$

where (A) follows from Lemma 18 (iii) and $\langle\boldsymbol{z}s'\boldsymbol{x}\rangle \in \mathcal{S}^L \subseteq \mathcal{J}$. Hence, there exists $\boldsymbol{y}' \in \mathcal{S}^*$ such that $f'^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y}') \succeq [f''^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y})]_k$, which leads to

$$\begin{aligned}
f'^{*}_{\langle\boldsymbol{z}\rangle}(s'\boldsymbol{x}\boldsymbol{y}') &= f'_{\langle\boldsymbol{z}\rangle}(s') f'^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x}) f'^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y}') \\
&\succeq f'_{\langle\boldsymbol{z}\rangle}(s') f'^{*}_{\langle\boldsymbol{z}s'\rangle}(\boldsymbol{x})[f'^{*}_{\langle\boldsymbol{z}s'\boldsymbol{x}\rangle}(\boldsymbol{y}')]_k \\
&\overset{\text{(A)}}{\succeq} f'_{\langle\boldsymbol{z}\rangle}(s)[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k, \qquad (92)
\end{aligned}$$

where (A) follows from (90). The assumption that $f'_{\langle\boldsymbol{z}\rangle}(s) \prec f'_{\langle\boldsymbol{z}\rangle}(s')$ and (92) shows that

$$[\psi_{\boldsymbol{z}s}(\boldsymbol{c})]_k \in \bar{\mathcal{P}}^k_{F',\langle\boldsymbol{z}\rangle}(f'_{\langle\boldsymbol{z}\rangle}(s)) \qquad (93)$$

by (5). By (80) and (93), the code-tuple F' does not satisfy Definition 5 (i), which conflicts with $F' \in \mathscr{F}_{k\text{-dec}}$.

Consequently, F'' satisfies Definition 5 (ii). □

4. Conclusion

This paper discussed the general properties of k-bit delay decodable code-tuples for $k \geq 0$ and proved two main theorems. Theorem 1 guarantees that it suffices to consider only irreducible code-tuples with at most $2^{(2^k)}$ code tables to achieve the optimal average codeword length. Theorem 2 is a generalization of the necessary condition of Huffman codes that every internal node in the code-tree has two child nodes. Both theorems enable us to limit the scope of code-tuples to be considered when discussing optimal k-bit delay decodable code-tuples.

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant Number JP18H01436, JP20K11674, and in part by KIOXIA. We would like to thank the editor and anonymous reviewers for their valuable comments and suggestions.

References

[1] D.A. Huffman, “A method for the construction of minimum-redundancy codes,” Proc. IRE, vol.40, no.9, pp.1098–1102, 1952.
[2] B. McMillan, “Two inequalities implied by unique decipherability,” IRE Trans. Inf. Theory, vol.2, no.4, pp.115–116, Dec. 1956.
[3] H. Yamamoto, M. Tsuchihashi, and J. Honda, “Almost instantaneous fixed-to-variable length codes,” IEEE Trans. Inf. Theory, vol.61, no.12, pp.6432–6443, Dec. 2015.
[4] K. Hashimoto and K. Iwata, “Optimality of Huffman code in the class of 1-bit delay decodable Codes,” IEEE J. Sel. Areas Inf. Theory, doi: 10.1109/JSAIT.2022.3230745, arXiv:2209.08874. [Online]. Available: http://arxiv.org/abs/2209.08874
[5] K. Hashimoto and K. Iwata, “On the optimality of binary AIFV codes with two code trees,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Melbourne, Victoria, Australia (Virtual Conference), pp.3173–3178, July 2021.
[6] K. Hashimoto and K. Iwata, “The optimality of AIFV codes in the class of 2-bit delay decodable codes,” arXiv:2306.09671. [Online]. Available: http://arxiv.org/abs/2306.09671
[7] W. Hu, H. Yamamoto, and J. Honda, “Worst-case redundancy of optimal binary AIFV codes and their extended codes,” IEEE Trans. Inf. Theory, vol.63, no.8, pp.5074–5086, Aug. 2017.
[8] R. Fujita K. Iwata, and H. Yamamoto, “Worst-case redundancy of optimal binary AIFV-m codes for $m = 3, 5$,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Los Angeles, CA, USA (Virtual Conference), pp.2355–2359, June 2020.
[9] K. Iwata and H. Yamamoto, “A dynamic programming algorithm to construct optimal code trees of AIFV codes,” Proc. Int. Symp. Inf. Theory and its Appl. (ISITA), Monterey, CA, USA, pp.641–645, Oct./Nov. 2016.
[10] K. Iwata and H. Yamamoto, “An iterative algorithm to construct optimal binary AIFV-m codes,” Proc. IEEE Inf. Theory Workshop (ITW), Kaohsiung, Taiwan, pp.519–523, Oct. 2017.
[11] K. Sumigawa and H. Yamamoto, “Coding of binary AIFV code trees,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Aachen, Germany, pp.1152–1156, June 2017.
[12] T. Hiraoka and H. Yamamoto, “Alphabetic AIFV codes constructed from Hu-tucker codes,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Vail, CO, USA, pp.2182–2186, June 2018.
[13] R. Fujita, K. Iwata, and H. Yamamoto, “An optimality proof of the iterative algorithm for AIFV-m codes,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Vail, CO, USA, pp.2187–2191, June 2018.
[14] T. Hiraoka and H. Yamamoto, “Dynamic AIFV coding,” Proc. Int. Symp. Inf. Theory and Its Appl. (ISITA), Singapore, pp.577–581, Oct. 2018.
[15] R. Fujita, K. Iwata, and H. Yamamoto, “An iterative algorithm to optimize the average performance of Markov chains with finite states,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Paris, France, pp.1902–1906, July 2019.
[16] K. Hashimoto, K. Iwata, and H. Yamamoto, “Enumeration and coding of compact code trees for binary AIFV codes,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Paris, France, pp.1527–1531, July 2019.
[17] M.J. Golin and E.Y. Harb, “Polynomial time algorithms for constructing optimal AIFV codes,” Proc. Data Compression Conference (DCC), Snowbird, UT, USA, pp.231–240, March 2019.
[18] M.J. Golin and E.Y. Harb, “Polynomial time algorithms for constructing optimal binary AIFV-2 codes,” 2020, arXiv:2001.11170. [Online]. Available: http://arxiv.org/abs/2001.11170
[19] H. Yamamoto, K. Imaeda, K. Hashimoto, and K. Iwata, “A universal data compression scheme based on the AIFV coding techniques,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Los Angeles, LA, USA (Virtual Conference), pp.2378–2382, June 2020.
[20] K. Iwata and H. Yamamoto, “An algorithm for construction the optimal code trees for binary alphabetic AIFV-m codes,” Proc. IEEE Inf. Theory Workshop (ITW), Riva de Garda, Italy (Virtual Conference), pp.261–265, April 2021 (postponed from 2020 to 2021).
[21] M.J. Golin and E.Y. Harb, “Speeding up the AIFV-2 dynamic programs by two orders of magnitude using Range Minimum Queries,” Theoretical Computer Science, vol.865, no.14, pp.99–118, April 2021.
[22] M.J. Golin and A.J.L. Parupat, “Speeding up AIFV-m dynamic programs by $m - 1$ orders of magnitude,” Proc. IEEE Int. Symp. on Inf. Theory (ISIT), Espoo, Finland, pp.282–287, June–July 2022.
[23] R. Sugiura, Y. Kamamoto, N. Harada, and T. Moriya, “Optimal Golomb-Rice code extension for lossless coding of low-entropy exponentially distributed sources,” IEEE Trans Inf. Theory, vol.64, no.4, pp.3153–3161, Jan. 2018.
[24] R. Sugiura, Y. Kamamoto, and T. Moriya, “General form of almost instantaneous fixed-to-variable-length codes and optimal code tree construction,” arXiv:2203.08437, 2022. [Online]. Available: http://arxiv.org/abs/2203.08437
[25] M.L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, John Wiley and Sons, Hoboken, NJ, USA, 2005.

Appendix A: Equivalency of the Definitions of a k-Bit Delay Decodable Code-Tuple

We confirm that Definition 5 in this paper is equivalent to the definition of a k-bit delay decodable code-tuple in [4]. We first introduce the following Definition 17.

Definition 17. *Let $F(f,\tau) \in \mathscr{F}$ and $i \in [F]$.*

(i) A pair $(\boldsymbol{x},\boldsymbol{c}) \in \mathcal{S}^* \times \mathcal{C}^*$ is said to be f_i^*-positive *if for any $\boldsymbol{x}' \in \mathcal{S}^*$, if $f_i^*(\boldsymbol{x})\boldsymbol{c} \preceq f_i^*(\boldsymbol{x}')$, then $\boldsymbol{x} \preceq \boldsymbol{x}'$.*
(ii) A pair $(\boldsymbol{x},\boldsymbol{c}) \in \mathcal{S}^* \times \mathcal{C}^*$ is said to be f_i^*-negative *if for any $\boldsymbol{x}' \in \mathcal{S}^*$, if $f_i^*(\boldsymbol{x})\boldsymbol{c} \preceq f_i^*(\boldsymbol{x}')$, then $\boldsymbol{x} \not\preceq \boldsymbol{x}'$.*

Then the definition of a k-bit delay decodable code-tuple in [4] is stated as the following Definition 18.

Definition 18. *Let $k \geq 0$ be an integer. A code-tuple F is said to be* k-bit delay decodable *if for any $i \in [F]$ and $(\boldsymbol{x},\boldsymbol{c}) \in \mathcal{S}^* \times \mathcal{C}^k$, the pair $(\boldsymbol{x},\boldsymbol{c})$ is f_i^*-positive or f_i^*-negative.*

We show that the following conditions (a) and (b) are equivalent for any $F(f,\tau) \in \mathscr{F}$.

(a) For any $i \in [F]$ and $(\boldsymbol{x},\boldsymbol{c}) \in \mathcal{S}^* \times \mathcal{C}^k$, the pair $(\boldsymbol{x},\boldsymbol{c})$ is f_i^*-positive or f_i^*-negative.
(b) The code-tuple F satisfies Definition 5 (i) and (ii).

((a) $\Longrightarrow$ (b)): We show the contraposition. Assume that (b) does not hold. We consider the following two cases separately: the case where Definition 5 (i) is false and the case where Definition 5 (ii) is false.

- The case where Definition 5 (i) is false: Then there exist $i \in [F]$, $s \in \mathcal{S}$, and $\boldsymbol{c} \in \mathcal{P}^k_{F,\tau_i(s)} \cap \bar{\mathcal{P}}^k_{F,i}(f_i(s))$. By (5) and (8), there exist $\boldsymbol{x} = x_1x_2\ldots,x_n \in \mathcal{S}^*$ and $\boldsymbol{x}' = x'_1x'_2\ldots x'_{n'} \in \mathcal{S}^+$ such that

$$f^*_{\tau_i(s)}(\boldsymbol{x}) \succeq \boldsymbol{c}, \tag{A$\cdot$1}$$
$$f_i^*(\boldsymbol{x}') \succeq f_i(s)\boldsymbol{c}, \tag{A$\cdot$2}$$
$$f_i(x'_1) \succ f_i(s). \tag{A$\cdot$3}$$

We have

$$f_i^*(s\boldsymbol{x}) \overset{\text{(A)}}{=} f_i(s) f^*_{\tau_i(s)}(\boldsymbol{x}) \overset{\text{(B)}}{\succeq} f_i(s)\boldsymbol{c}, \tag{A·4}$$

where (A) follows from (2), and (B) follows from (A· 1). By (A· 4) and $s \preceq s\boldsymbol{x}$, the pair $(s, \boldsymbol{c})$ is not f_i^*-negative. On the other hand, since $s \neq x'_1$ by (A· 3), we have $s \npreceq \boldsymbol{x}'$. Hence, by (A· 2), the pair $(s, \boldsymbol{c})$ is not f_i^*-positive. Since the pair $(s, \boldsymbol{c})$ is neither f_i^*-positive nor f_i^*-negative, the condition (a) does not hold.

- The case where Definition 5 (ii) is false: Then there exist $i \in [F]$, $s, s' \in \mathcal{S}$, and $\boldsymbol{c} \in \mathcal{P}^k_{F,\tau_i(s)} \cap \mathcal{P}^k_{F,\tau_i(s')}$ such that $s \neq s'$ and

$$f_i(s) = f_i(s'). \tag{A·5}$$

By (8), there exist $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{S}^*$ such that

$$f^*_{\tau_i(s)}(\boldsymbol{x}) \succeq \boldsymbol{c} \tag{A·6}$$

and

$$f^*_{\tau_i(s')}(\boldsymbol{x}') \succeq \boldsymbol{c}. \tag{A·7}$$

Thus, we have

$$f_i^*(s\boldsymbol{x}) \overset{\text{(A)}}{=} f_i(s) f^*_{\tau_i(s)}(\boldsymbol{x}) \overset{\text{(B)}}{\succeq} f_i(s)\boldsymbol{c} \tag{A·8}$$

and

$$\begin{aligned} f_i^*(s'\boldsymbol{x}') &\overset{\text{(C)}}{=} f_i(s') f^*_{\tau_i(s')}(\boldsymbol{x}') \\ &\overset{\text{(D)}}{=} f_i(s) f^*_{\tau_i(s')}(\boldsymbol{x}') \\ &\overset{\text{(E)}}{\succeq} f_i(s)\boldsymbol{c}, \end{aligned} \tag{A·9}$$

where (A) follows from (2), (B) follows from (A· 6), (C) follows from (2), (D) follows from (A· 5), and (E) follows from (A· 7). By (A· 8) and $s \preceq s\boldsymbol{x}$, the pair $(s, \boldsymbol{c})$ is not f_i^*-negative. On the other hand, by $s \npreceq s'\boldsymbol{x}$ and (A· 9), the pair $(s, \boldsymbol{c})$ is not f_i^*-positive. Since the pair $(s, \boldsymbol{c})$ is neither f_i^*-positive nor f_i^*-negative, the condition (a) does not hold.

((b) $\implies$ (a)): We show the contraposition. Assume that (a) does not hold. Then there exist $i \in [F]$ and $(\boldsymbol{x}, \boldsymbol{c}) \in \mathcal{S}^* \times \mathcal{C}^k$ such that $(\boldsymbol{x}, \boldsymbol{c})$ is neither f_i^*-positive nor f_i^*-negative. Thus, there exist $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{S}^*$ such that

$$f_i^*(\boldsymbol{x})\boldsymbol{c} \preceq f_i^*(\boldsymbol{x}'), \tag{A·10}$$
$$f_i^*(\boldsymbol{x})\boldsymbol{c} \preceq f_i^*(\boldsymbol{x}''), \tag{A·11}$$
$$\boldsymbol{x} \preceq \boldsymbol{x}', \tag{A·12}$$
$$\boldsymbol{x} \npreceq \boldsymbol{x}''. \tag{A·13}$$

We consider the following two cases separately: the case $\boldsymbol{x} \succeq \boldsymbol{x}''$ and the case $\boldsymbol{x} \nsucceq \boldsymbol{x}''$.

- The case $\boldsymbol{x} \succeq \boldsymbol{x}''$: By Lemma 1 (iii), we have

$$f_i^*(\boldsymbol{x}) \succeq f_i^*(\boldsymbol{x}''). \tag{A·14}$$

Hence, by (A· 11), it must hold that $\boldsymbol{c} = \lambda$. Namely, only $k = 0$ is possible now.
Since (A· 13) and $\boldsymbol{x} \succeq \boldsymbol{x}''$ lead to $\boldsymbol{x} \succ \boldsymbol{x}''$, there exists $\boldsymbol{u} = u_1 u_2 \ldots u_n \in \mathcal{S}^+$ such that $\boldsymbol{x} = \boldsymbol{x}''\boldsymbol{u}$. Defining $j := \tau_i^*(\boldsymbol{x}'')$, we have

$$\begin{aligned} f_i^*(\boldsymbol{x}) &\overset{\text{(A)}}{=} f_i^*(\boldsymbol{x}'') f_j^*(\boldsymbol{u}) \\ &\overset{\text{(B)}}{=} f_i^*(\boldsymbol{x}) f_j^*(\boldsymbol{u}) \\ &\overset{\text{(C)}}{=} f_i^*(\boldsymbol{x}) f_j(u_1) f^*_{\tau_j(u_1)}(\mathrm{suff}(\boldsymbol{u})), \end{aligned} \tag{A·15}$$

where (A) follows from Lemma 1 (i), (B) follows because we have $f_i^*(\boldsymbol{x}) \preceq f_i^*(\boldsymbol{x}'')$ by (A· 11) and we have $f_i^*(\boldsymbol{x}) \succeq f_i^*(\boldsymbol{x}'')$ by (A· 14), and (C) follows from (2). Comparing both sides of (A· 15), we obtain

$$f_j(u_1) = \lambda \tag{A·16}$$

and $f^*_{\tau_j(u_1)}(\mathrm{suff}(\boldsymbol{u})) = \lambda$.
We now show that (b) does not hold dividing into two cases by whether f_j is injective.

  - If f_j is not injective, then F does not satisfy Definition 5 (ii) by $k = 0$ and Lemma 2.
  - If f_j is injective, then there exists $s \in \mathcal{S}$ such that $f_j(s) \succ \lambda$ by $\sigma \geq 2$, which leads to

$$\bar{\mathcal{P}}^0_{F,j} \neq \emptyset \tag{A·17}$$

by (5). We see that F does not satisfy Definition 5 (i) because

$$\begin{aligned} &\mathcal{P}^0_{F,\tau_j(u_1)} \cap \bar{\mathcal{P}}^0_{F,j}(f_j(u_1)) \\ &\overset{\text{(A)}}{=} \mathcal{P}^0_{F,\tau_j(u_1)} \cap \bar{\mathcal{P}}^0_{F,j} \\ &\overset{\text{(B)}}{=} \{\lambda\} \cap \bar{\mathcal{P}}^0_{F,j} \\ &\overset{\text{(C)}}{=} \{\lambda\} \cap \{\lambda\} \\ &= \{\lambda\} \\ &\neq \emptyset, \end{aligned}$$

where (A) follows from (A· 16), (B) follows from (8), and (C) follows from (A· 17).

- The case $\boldsymbol{x} \nsucceq \boldsymbol{x}''$: By (A· 13) and $\boldsymbol{x} \nsucceq \boldsymbol{x}''$, there exist $\boldsymbol{z} = z_1 z_2 \ldots z_n \in \mathcal{S}^+$ and $\boldsymbol{z}'' = z''_1 z''_2 \ldots z''_{n''} \in \mathcal{S}^+$ such that

$$\boldsymbol{x} = \boldsymbol{y}\boldsymbol{z}, \tag{A·18}$$
$$\boldsymbol{x}'' = \boldsymbol{y}\boldsymbol{z}'', \tag{A·19}$$
$$z_1 \neq z''_1, \tag{A·20}$$

where $\boldsymbol{y}$ is the longest common prefix of $\boldsymbol{x}$ and $\boldsymbol{x}''$. Also, by (A· 12), there exists $\boldsymbol{w} \in \mathcal{S}^*$ such that

$$\boldsymbol{x}' = \boldsymbol{x}\boldsymbol{w}. \tag{A·21}$$

Defining $\boldsymbol{z}' = z'_1 z'_2 \ldots z'_{n'} := \boldsymbol{z}\boldsymbol{w}$, we have

$$\boldsymbol{x}' = \boldsymbol{x}\boldsymbol{w} = \boldsymbol{y}\boldsymbol{z}\boldsymbol{w} = \boldsymbol{y}\boldsymbol{z}', \tag{A$\cdot$22}$$
$$z_1 = z_1'. \tag{A$\cdot$23}$$

Then defining $j \coloneqq \tau_i^*(\boldsymbol{y})$, we have

$$\begin{aligned} &f_i^*(\boldsymbol{y})f_j(z_1')f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}')) \\ &\overset{\text{(A)}}{=} f_i^*(\boldsymbol{y})f_j^*(\boldsymbol{z}') \overset{\text{(B)}}{=} f_i^*(\boldsymbol{y}\boldsymbol{z}') \overset{\text{(C)}}{=} f_i^*(\boldsymbol{x}') \overset{\text{(D)}}{\succeq} f_i^*(\boldsymbol{x})\boldsymbol{c}, \end{aligned} \tag{A$\cdot$24}$$

where (A) follows from (2), (B) follows from Lemma 1 (i), (C) follows from (A$\cdot$22), and (D) follows from (A$\cdot$10). Similarly, by (A$\cdot$11) and (A$\cdot$19), we have

$$\begin{aligned} &f_i^*(\boldsymbol{y})f_j(z_1'')f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) \\ &= f_i^*(\boldsymbol{y})f_j^*(\boldsymbol{z}'') = f_i^*(\boldsymbol{y}\boldsymbol{z}'') = f_i^*(\boldsymbol{x}'') \succeq f_i^*(\boldsymbol{x})\boldsymbol{c}. \end{aligned} \tag{A$\cdot$25}$$

Also, we have

$$\begin{aligned} f_i^*(\boldsymbol{x})\boldsymbol{c} &\overset{\text{(A)}}{=} f_i^*(\boldsymbol{y}\boldsymbol{z})\boldsymbol{c} \\ &\overset{\text{(B)}}{=} f_i^*(\boldsymbol{y})f_j^*(\boldsymbol{z})\boldsymbol{c} \\ &\overset{\text{(C)}}{=} f_i^*(\boldsymbol{y})f_j(z_1)f^*_{\tau_j(z_1)}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c}, \end{aligned} \tag{A$\cdot$26}$$

where (A) follows from (A$\cdot$18), (B) follows from Lemma 1 (i), and (C) follows from (2).
Thus, we have

$$\begin{aligned} f_j(z_1')f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}')) &\overset{\text{(A)}}{\succeq} f_j(z_1)f^*_{\tau_j(z_1)}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c} \\ &\overset{\text{(B)}}{=} f_j(z_1')f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c} \\ &\succeq f_j(z_1')\boldsymbol{c}', \end{aligned} \tag{A$\cdot$27}$$

where $\boldsymbol{c}' \in \mathcal{C}^k$ is defined as the prefix of length k of $f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c}$, and (A) follows from (A$\cdot$24) and (A$\cdot$26), and (B) follows from (A$\cdot$23). Similarly, we have

$$\begin{aligned} f_j(z_1'')f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) &\overset{\text{(A)}}{\succeq} f_j(z_1)f^*_{\tau_j(z_1)}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c} \\ &\overset{\text{(B)}}{=} f_j(z_1')f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}))\boldsymbol{c} \\ &\succeq f_j(z_1')\boldsymbol{c}', \end{aligned} \tag{A$\cdot$28}$$

where (A) follows from (A$\cdot$25) and (A$\cdot$26), and (B) follows from (A$\cdot$23). By (A$\cdot$27), we have $f^*_{\tau_j(z_1')}(\mathrm{suff}(\boldsymbol{z}')) \succeq \boldsymbol{c}'$, which leads to

$$\boldsymbol{c}' \in \mathcal{P}^k_{F,\tau_j(z_1')} \tag{A$\cdot$29}$$

by (8).
By (A$\cdot$28), at least one of $f_j(z_1') \preceq f_j(z_1'')$ and $f_j(z_1') \succeq f_j(z_1'')$ holds. We may assume $f_j(z_1') \preceq f_j(z_1'')$ by symmetry. We consider the following two cases separately: the case $f_j(z_1') \prec f_j(z_1'')$ and the case $f_j(z_1') = f_j(z_1'')$.

- The case $f_j(z_1') \prec f_j(z_1'')$: We have

$$f_j^*(\boldsymbol{z}'') \overset{\text{(A)}}{=} f_j(z_1'')f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) \overset{\text{(B)}}{\succeq} f_j(z_1')\boldsymbol{c}', \tag{A$\cdot$30}$$

where (A) follows from (2), and (B) follows from (A$\cdot$28). By (A$\cdot$30) and $f_j(z_1') \prec f_j(z_1'')$, we obtain

$$\boldsymbol{c}' \in \bar{\mathcal{P}}^k_{F,j}(f_j(z_1')) \tag{A$\cdot$31}$$

by (5). By (A$\cdot$29) and (A$\cdot$31), the code-tuple F does not satisfy Definition 5 (i).

- The case $f_j(z_1') = f_j(z_1'')$: We have

$$\begin{aligned} &f_j(z_1')f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) \\ &\overset{\text{(A)}}{=} f_j(z_1'')f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) \\ &\overset{\text{(B)}}{\succeq} f_j(z_1')\boldsymbol{c}', \end{aligned}$$

where (A) follows from $f_j(z_1') = f_j(z_1'')$, and (B) follows from (A$\cdot$28). This shows $f^*_{\tau_j(z_1'')}(\mathrm{suff}(\boldsymbol{z}'')) \succeq \boldsymbol{c}'$, which leads to

$$\boldsymbol{c}' \in \mathcal{P}^k_{F,\tau_j(z_1'')} \tag{A$\cdot$32}$$

by (8). By $f_j(z_1') = f_j(z_1'')$, (A$\cdot$20), (A$\cdot$29), and (A$\cdot$32), the code-tuple F does not satisfy Definition 5 (ii).

Appendix B: Proof of the Existence of an Optimal Code-Tuple

For $m \in \{1, 2, \ldots, M \coloneqq 2^{(2^k)}\}$, the number of possible tuples $(\tau_0, \tau_1, \ldots, \tau_{m-1})$ (i.e., a tuple of m mappings from $\mathcal{S}$ to $[m]$) is $m^{\sigma m}$, in particular, finite. Hence, the number of possible vectors $\boldsymbol{\pi}(F') = (\pi_0(F'), \pi_1(F'), \ldots, \pi_{m-1}(F'))$ of a code-tuple $F' \in \mathscr{F}'$ is also finite (cf. Remark 2), where

$$\mathscr{F}' \coloneqq \{F' \in \mathscr{F}_{\mathrm{irr}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}} : |F'| \leq M\}. \tag{A$\cdot$33}$$

Therefore, $\mathcal{D} \coloneqq \{\pi_i(F') : F' \in \mathscr{F}', i \in [F']\}$ is a finite set and has the minimum value $\delta \coloneqq \min \mathcal{D}$. Note that $\delta > 0$ holds since $\pi_i(F') > 0$ for any $F' \in \mathscr{F}'$ and $i \in [F']$ by $\mathscr{F}' \subseteq \mathscr{F}_{\mathrm{irr}}$ and Lemma 8 (ii).

Now, we define

$$\mathscr{F}'' \coloneqq \{F'(f', \tau') \in \mathscr{F}' : \sum_{i \in [F'], s \in \mathcal{S}} |f_i'(s)| \leq \frac{l}{\delta\nu}\}, \tag{A$\cdot$34}$$

where $l \coloneqq \lceil \log_2 \sigma \rceil$ and $\nu \coloneqq \min_{s \in \mathcal{S}} \mu(s)$. Note that

$$0 < \nu \leq 1/\sigma. \tag{A$\cdot$35}$$

Then $\mathscr{F}''$ is not empty because $\tilde{F}(\tilde{f}_0, \tilde{\tau}_0) \in \mathscr{F}^{(1)}$ defined as the following (A$\cdot$36) is in $\mathscr{F}''$:

$$\tilde{f}_0(s_r) = b(r), \quad \tilde{\tau}_0(s_r) = 0 \tag{A$\cdot$36}$$

for $r = 0, 1, 2, \ldots, \sigma - 1$, where $\mathcal{S} = \{s_0, s_1, \ldots, s_{\sigma-1}\}$ and $b(r)$ denotes the binary representation of length l of the integer r. In fact, we obtain $\tilde{F} \in \mathscr{F}''$ by

$$\sum_{i \in [\tilde{F}], s \in \mathcal{S}} |\tilde{f}_i(s)| = \sum_{s \in \mathcal{S}} |\tilde{f}_0(s)| = \sigma l \overset{\text{(A)}}{\leq} \frac{l}{\nu} \overset{\text{(B)}}{\leq} \frac{l}{\delta\nu}, \tag{A$\cdot$37}$$

where (A) follows from (A· 35), and (B) follows from $0 < \delta \leq 1$. Since $\mathscr{F}''$ is a non-empty and finite set, there exists $F^* \in \mathscr{F}''$ such that

$$L(F^*) = \min_{F'' \in \mathscr{F}''} L(F''). \tag{A$\cdot$38}$$

To complete the proof, it suffices to show that $L(F^*) \leq L(F)$ for any $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$.

First, we can see that $L(F^*) \leq L(F')$ for any $F' \in \mathscr{F}'$ because for any $F'(f', \tau') \in \mathscr{F}' \setminus \mathscr{F}''$, we have

$$\begin{aligned}
L(F') &= \sum_{i \in [F']} \pi_i(F') L_i(F') \\
&= \sum_{i \in [F']} \pi_i(F') \sum_{s \in \mathcal{S}} \mu(s) |f'_i(s)| \\
&\overset{\text{(A)}}{\geq} \delta\nu \sum_{i \in [F'], s \in \mathcal{S}} |f'_i(s)| \\
&\overset{\text{(B)}}{>} \delta\nu \cdot \frac{l}{\delta\nu} \\
&= l \\
&= L(\tilde{F}) \\
&\overset{\text{(C)}}{\geq} L(F^*),
\end{aligned}$$

where (A) follows from the definitions of δ and ν, (B) follows from $F' \notin \mathscr{F}''$, and (C) follows from (A· 38). Hence, we have

$$L(F^*) = \min_{F' \in \mathscr{F}'} L(F'). \tag{A$\cdot$39}$$

By Theorem 1, for any $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$, there exists $F' \in \mathscr{F}_{\mathrm{irr}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$ such that $L(F') \leq L(F)$ and $|\mathscr{P}^k_{F'}| = |F'|$. Then we have $F' \in \mathscr{F}'$ because

$$|F'| = |\mathscr{P}^k_{F'}| \leq |\mathcal{P}(C^k)| = 2^{(2^k)} = M, \tag{A$\cdot$40}$$

where $\mathcal{P}(C^k)$ denotes the power set of C^k. Therefore, for any $F \in \mathscr{F}_{\mathrm{reg}} \cap \mathscr{F}_{\mathrm{ext}} \cap \mathscr{F}_{k\text{-dec}}$, we have

$$L(F) \geq L(F') \overset{\text{(A)}}{\geq} L(F^*) \tag{A$\cdot$41}$$

as desired, where (A) follows from (A· 39).

Appendix C: Proofs of Lemmas

C.1 Proof of Lemma 5

To prove Lemma 5, we first show the following Lemma 20.

Lemma 20. *For any integer $k \geq 0$, $F(f, \tau) \in \mathscr{F}$ and $i \in [F]$, the following statements (i) and (ii) hold.*

(i) $\mathcal{P}^k_{F,i} \supseteq \bar{\mathcal{P}}^k_{F,i}$.

(ii) For any $s \in \mathcal{S}$ such that $f_i(s) = \lambda$, we have $\mathcal{P}^k_{F,i} \supseteq \mathcal{P}^k_{F,\tau_i(s)}$.

Proof of Lemma 20. (Proof of (i)): Directly from (4) and (5).

(Proof of (ii)): Choose $\boldsymbol{c} \in \mathcal{P}^k_{F,\tau_i(s)}$ arbitrarily. Then there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that

$$f^*_{\tau_i(s)}(\boldsymbol{x}) \geq \boldsymbol{c} \tag{A$\cdot$42}$$

by (8). We have

$$f^*_i(s\boldsymbol{x}) \overset{\text{(A)}}{=} f_i(s) f^*_{\tau_i(s)}(\boldsymbol{x}) \overset{\text{(B)}}{=} f^*_{\tau_i(s)}(\boldsymbol{x}) \overset{\text{(C)}}{\geq} \boldsymbol{c}, \tag{A$\cdot$43}$$

where (A) follows from Lemma 1 (i), (B) follows from the assumption, and (C) follows from (A· 42). This leads to $\boldsymbol{c} \in \mathcal{P}^k_{F,i}$. □

Proof of Lemma 5. It suffices to show that $|f^*_i(\boldsymbol{x})| \geq 1$ holds for any $i \in [F]$ and $\boldsymbol{x} \in \mathcal{S}^{|F|}$. We prove by contradiction assuming that there exist $i \in [F]$ and $\boldsymbol{x} = x_1 x_2 \ldots x_{|F|} \in \mathcal{S}^{|F|}$ such that $f^*_i(\boldsymbol{x}) = \lambda$. Then by pigeonhole principle, we can choose integers p, q such that $0 \leq p < q \leq |F|$ and

$$\tau^*_i(x_1 x_2 \ldots x_p) = \tau^*_i(x_1 x_2 \ldots x_q) =: j. \tag{A$\cdot$44}$$

We have

$$\begin{aligned}
\tau^*_j(x_{p+1} x_{p+2} \ldots x_q) &\overset{\text{(A)}}{=} \tau^*_{\tau^*_i(x_1 x_2 \ldots x_p)}(x_{p+1} x_{p+2} \ldots x_q) \\
&\overset{\text{(B)}}{=} \tau^*_i(x_1 x_2 \ldots x_q) \overset{\text{(C)}}{=} j,
\end{aligned} \tag{A$\cdot$45}$$

where (A) follows from (A· 44), (B) follows from Lemma 1 (ii), and (C) follows from (A· 44). Thus, we obtain

$$\begin{aligned}
\mathcal{P}^k_{F,\tau_j(x_{p+1})} &\overset{\text{(A)}}{\supseteq} \mathcal{P}^k_{F,\tau^*_j(x_{p+1} x_{p+2})} \\
&\overset{\text{(A)}}{\supseteq} \cdots \\
&\overset{\text{(A)}}{\supseteq} \mathcal{P}^k_{F,\tau^*_j(x_{p+1} x_{p+2} \ldots x_q)} \\
&\overset{\text{(B)}}{=} \mathcal{P}^k_{F,j},
\end{aligned} \tag{A$\cdot$46}$$

where (A)s follow from Lemma 20 (ii) and $f^*_i(\boldsymbol{x}) = \lambda$, and (B) follows from (A· 45).

We consider the following two cases separately: the case $\bar{\mathcal{P}}^k_{F,j} \neq \emptyset$ and the case $\bar{\mathcal{P}}^k_{F,j} = \emptyset$.

- The case $\bar{\mathcal{P}}^k_{F,j} \neq \emptyset$: We have

$$\begin{aligned}
\mathcal{P}^k_{F,\tau_j(x_{p+1})} \cap \bar{\mathcal{P}}^k_{F,j} &\overset{\text{(A)}}{\supseteq} \mathcal{P}^k_{F,j} \cap \bar{\mathcal{P}}^k_{F,j} \\
&\overset{\text{(B)}}{\supseteq} \bar{\mathcal{P}}^k_{F,j} \cap \bar{\mathcal{P}}^k_{F,j} \\
&= \bar{\mathcal{P}}^k_{F,j} \\
&\overset{\text{(C)}}{\neq} \emptyset,
\end{aligned}$$

where (A) follows from (A· 46), (B) follows from Lemma 20 (i), and (C) follows from the assumption. Therefore, F does not satisfy Definition 5 (i), which conflicts with $F \in \mathscr{F}_{k\text{-dec}}$.

- The case $\bar{\mathcal{P}}^k_{F,j} = \emptyset$: By Corollary 1 (ii), we have $\bar{\mathcal{P}}^0_{F,j} = \emptyset$. Hence, by (5), there is no symbol $s' \in \mathcal{S}$ such that $f_j(s') > \lambda$. Therefore, by $\sigma \geq 2$, there exists $s \in \mathcal{S}$ such that $s \neq x_{p+1}$ and $f_j(s) = \lambda = f_j(x_{p+1})$. We have

$$\begin{aligned}\mathcal{P}^k_{F,\tau_j(x_{p+1})} \cap \mathcal{P}^k_{F,\tau_j(s)} &\overset{\text{(A)}}{\supseteq} \mathcal{P}^k_{F,j} \cap \mathcal{P}^k_{F,\tau_j(s)}\\ &\overset{\text{(B)}}{\supseteq} \mathcal{P}^k_{F,\tau_j(s)} \cap \mathcal{P}^k_{F,\tau_j(s)}\\ &= \mathcal{P}^k_{F,\tau_j(s)} \overset{\text{(C)}}{\neq} \emptyset,\end{aligned}$$

where (A) follows from (A· 46), (B) follows from Lemma 20 (ii), and (C) follows from $F \in \mathscr{F}_{\text{ext}}$ and Corollary 1 (i). Therefore, F does not satisfy Definition 5 (ii), which conflicts with $F \in \mathscr{F}_{k\text{-dec}}$.

□

C.2 Proof of Lemma 6

In preparation for the proof, we introduce the following Definition 19 and Lemma 21.

Definition 19. *Let $F(f,\tau) \in \mathscr{F}$. A set $\mathcal{I} \subseteq [F]$ is said to be* closed *if for any $i \in \mathcal{I}$ and $s \in \mathcal{S}$, it holds that $\tau_i(s) \in \mathcal{I}$.*

Lemma 21. *For any $F \in \mathscr{F}$ and $\boldsymbol{x} = (x_0, x_1, \ldots, x_{|F|-1}) \in \mathbb{R}^{|F|}$, if*

$$\boldsymbol{x}Q(F) = \boldsymbol{x}, \tag{A· 47}$$

then both of $\mathcal{I}_+ := \{i \in [F] : x_i > 0\}$ and $\mathcal{I}_- := \{i \in [F] : x_i < 0\}$ are closed.

Proof of Lemma 21. By symmetry, it suffices to prove only that $\mathcal{I}_+$ is closed. We have

$$\begin{aligned}&\sum_{i\in\mathcal{I}_+}\sum_{j\in\mathcal{I}_+} x_jQ_{j,i}(F) + \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_jQ_{j,i}(F)\\ &= \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]} x_jQ_{j,i}(F)\\ &\overset{\text{(A)}}{=} \sum_{i\in\mathcal{I}_+} x_i\\ &\overset{\text{(B)}}{=} \sum_{i\in\mathcal{I}_+} x_i \sum_{j\in[F]} Q_{i,j}(F)\\ &= \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]} x_iQ_{i,j}(F)\\ &= \sum_{i\in\mathcal{I}_+}\sum_{j\in\mathcal{I}_+} x_iQ_{i,j}(F) + \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_iQ_{i,j}(F)\\ &\overset{\text{(C)}}{=} \sum_{i\in\mathcal{I}_+}\sum_{j\in\mathcal{I}_+} x_jQ_{j,i}(F) + \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_iQ_{i,j}(F),\end{aligned} \tag{A· 48}$$

where (A) follows from (A· 47), (B) follows from $\sum_{j\in[F]} Q_{i,j}(F) = 1$ for any $i \in [F]$, and (C) is obtained by exchanging the roles of i and j in the first term. Therefore, we have

$$\begin{aligned}0 &\overset{\text{(A)}}{\geq} \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_jQ_{j,i}(F)\\ &\overset{\text{(B)}}{=} \sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_iQ_{i,j}(F)\\ &\overset{\text{(C)}}{\geq} 0,\end{aligned}$$

where (A) follows since $x_j \leq 0$ for any $j \in [F] \setminus \mathcal{I}_+$, (B) is obtained by eliminating the first terms from the leftmost and rightmost sides of (A· 48), and (C) follows since $x_i > 0$ for any $i \in \mathcal{I}_+$. This shows

$$\sum_{i\in\mathcal{I}_+}\sum_{j\in[F]\setminus\mathcal{I}_+} x_iQ_{i,j}(F) = 0.$$

Since $x_i > 0$ holds for any $i \in \mathcal{I}_+$, it must hold that $Q_{i,j}(F) = 0$ for any $i \in \mathcal{I}_+$ and $j \in [F] \setminus \mathcal{I}_+$. This implies that for any $i \in \mathcal{I}_+$ and $s \in \mathcal{S}$, we have $\tau_i(s) \in \mathcal{I}_+$; that is, $\mathcal{I}_+$ is closed as desired. □

Proof of Lemma 6. Equation (19) can be rewritten as

$$\boldsymbol{\pi}A = \boldsymbol{0}, \tag{A· 49}$$

where $A = (A_{i,j}) := Q(F) - E$ and E is the identity matrix. We have $\det A = 0$ because the sum of each row of A equals 0: for any $i \in [F]$, we have

$$\begin{aligned}\sum_{j\in[F]} A_{i,j} &= \sum_{j\in[F]} (Q_{i,j}(F) - \delta_{ij})\\ &= \sum_{j\in[F]} Q_{i,j}(F) - \sum_{j\in[F]} \delta_{ij}\\ &= \sum_{j\in[F]} Q_{i,j}(F) - 1\\ &= \sum_{j\in[F]}\sum_{s\in\mathcal{S},\tau_i(s)=j} \mu(s) - 1\\ &= \sum_{s\in\mathcal{S}} \mu(s) - 1\\ &= 0,\end{aligned}$$

where δ_{ij} denotes Kronecker delta. Thus, the dimension of the null space of A is greater than or equal to 1. In particular, Equation (A· 49), which is equivalent to (19), has a non-trivial solution $\boldsymbol{\pi} \neq \boldsymbol{0}$. We choose such $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1}) \neq \boldsymbol{0}$. Then both of $\mathcal{I}_+ := \{i \in [F] : \pi_i > 0\}$ and $\mathcal{I}_- := \{i \in [F] : \pi_i < 0\}$ are closed by Lemma 21. Hence, we have

$$^\forall i \in \mathcal{I}_+; {}^\forall j \in [F] \setminus \mathcal{I}_+; Q_{i,j}(F) = 0, \tag{A· 50}$$

$$^\forall i \in \mathcal{I}_-; {}^\forall j \in [F] \setminus \mathcal{I}_-; Q_{i,j}(F) = 0. \tag{A· 51}$$

Since $\boldsymbol{\pi} \neq \boldsymbol{0}$, we have $\sum_{i\in[F]} |\pi_i| > 0$ and thus we can define $\boldsymbol{\pi}' = (\pi'_0, \pi'_1, \ldots, \pi'_{|F|-1}) \in \mathbb{R}^{|F|}$ as

$$\pi'_i = \frac{|\pi_i|}{\sum_{i\in[F]} |\pi_i|} \tag{A$\cdot$ 52}$$

for $i \in [F]$. This vector $\boldsymbol{\pi}'$ is a desired stationary distribution of F. In fact, by the definition, $\boldsymbol{\pi}'$ clearly satisfies (20) and $\pi'_i \geq 0$ for any $i \in [F]$. Also, we can see that $\boldsymbol{\pi}'$ satisfies (19) because for any $j \in [F]$, we have

$$\begin{aligned}
&\Big(\sum_{i\in[F]} |\pi_i|\Big)\Big(\sum_{i\in[F]} \pi'_i Q_{i,j}(F)\Big)\\
&\overset{\text{(A)}}{=} \sum_{i\in[F]} |\pi_i| Q_{i,j}(F)\\
&= \sum_{i\in\mathcal{I}_+} \pi_i Q_{i,j}(F) - \sum_{i\in\mathcal{I}_-} \pi_i Q_{i,j}(F)\\
&\overset{\text{(B)}}{=} \begin{cases} \sum_{i\in\mathcal{I}_+} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_+,\\ -\sum_{i\in\mathcal{I}_-} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_-,\\ 0 & \text{otherwise,}\end{cases}\\
&\overset{\text{(C)}}{=} \begin{cases} \sum_{i\in\mathcal{I}_+} \pi_i Q_{i,j}(F) + \sum_{i\in\mathcal{I}_-} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_+,\\ -\sum_{i\in\mathcal{I}_+} \pi_i Q_{i,j}(F) - \sum_{i\in\mathcal{I}_-} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_-,\\ 0 & \text{otherwise,}\end{cases}\\
&= \begin{cases} \sum_{i\in[F]} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_+,\\ -\sum_{i\in[F]} \pi_i Q_{i,j}(F) & \text{if } j \in \mathcal{I}_-,\\ 0 & \text{otherwise,}\end{cases}\\
&\overset{\text{(D)}}{=} \begin{cases} \pi_j & \text{if } j \in \mathcal{I}_+,\\ -\pi_j & \text{if } j \in \mathcal{I}_-,\\ 0 & \text{otherwise,}\end{cases}\\
&= |\pi_j|\\
&\overset{\text{(E)}}{=} \Big(\sum_{i\in[F]} |\pi_i|\Big)\pi'_j,
\end{aligned}$$

where (A) follows from (A$\cdot$ 52), (B) follows from (A$\cdot$ 50) and (A$\cdot$ 51), (C) follows from (A$\cdot$ 50) and (A$\cdot$ 51), (D) follows since $\boldsymbol{\pi}$ is a stationary distribution of F, and (E) follows from (A$\cdot$ 52). □

C.3 Proof of Lemma 7

Proof of Lemma 7. (Proof of (i)): We first show that $f'^*_i(\boldsymbol{x}) = f^*_{\varphi(i)}(\boldsymbol{x})$ for any $i \in [F']$ and $\boldsymbol{x} \in \mathcal{S}^*$ by induction for $|\boldsymbol{x}|$. For the base case $|\boldsymbol{x}| = 0$, we have $f'^*_i(\lambda) = \lambda = f^*_{\varphi(i)}(\lambda)$ by (2). We consider the induction step for $|\boldsymbol{x}| \geq 1$. We have

$$\begin{aligned}
f'^*_i(\boldsymbol{x}) &\overset{\text{(A)}}{=} f'_i(x_1) f'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x}))\\
&\overset{\text{(B)}}{=} f_{\varphi(i)}(x_1) f'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x}))\\
&\overset{\text{(C)}}{=} f_{\varphi(i)}(x_1) f^*_{\varphi(\tau'_i(x_1))}(\mathrm{suff}(\boldsymbol{x}))\\
&\overset{\text{(D)}}{=} f_{\varphi(i)}(x_1) f^*_{\tau_{\varphi(i)}(x_1)}(\mathrm{suff}(\boldsymbol{x}))\\
&\overset{\text{(E)}}{=} f^*_{\varphi(i)}(\boldsymbol{x})
\end{aligned}$$

as desired, where (A) follows from (2), (B) follows from (23), (C) follows from the induction hypothesis, (D) follows from (24), and (E) follows from (2).

Next, we show that $\varphi(\tau'^*_i(\boldsymbol{x})) = \tau^*_{\varphi(i)}(\boldsymbol{x})$ for any $i \in [F']$ and $\boldsymbol{x} \in \mathcal{S}^*$ by induction for $|\boldsymbol{x}|$. For the base case $|\boldsymbol{x}| = 0$, we have $\varphi(\tau'^*_i(\lambda)) = \varphi(i) = \tau^*_{\varphi(i)}(\lambda)$ by (3). We consider the induction step for $|\boldsymbol{x}| \geq 1$. We have

$$\begin{aligned}
\varphi(\tau'^*_i(\boldsymbol{x})) &\overset{\text{(A)}}{=} \varphi(\tau'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x}))) \overset{\text{(B)}}{=} \tau^*_{\varphi(\tau'_i(x_1))}(\mathrm{suff}(\boldsymbol{x}))\\
&\overset{\text{(C)}}{=} \tau^*_{\tau_{\varphi(i)}(x_1)}(\mathrm{suff}(\boldsymbol{x})) \overset{\text{(D)}}{=} \tau^*_{\varphi(i)}(\boldsymbol{x})
\end{aligned}$$

as desired, where (A) follows from (3), (B) follows from the induction hypothesis, (C) follows from (24), and (D) follows from (3).

(Proof of (ii)): For any $\boldsymbol{c} \in C^*$, we have

$$\begin{aligned}
&\boldsymbol{c} \in \mathcal{P}^*_{F',i}(\boldsymbol{b})\\
&\overset{\text{(A)}}{\iff} {}^\exists \boldsymbol{x} \in \mathcal{S}^+; (f'^*_i(\boldsymbol{x}) \geq \boldsymbol{bc}, f'_i(x_1) \geq \boldsymbol{b})\\
&\overset{\text{(B)}}{\iff} {}^\exists \boldsymbol{x} \in \mathcal{S}^+; (f^*_{\varphi(i)}(\boldsymbol{x}) \geq \boldsymbol{bc}, f_{\varphi(i)}(x_1) \geq \boldsymbol{b})\\
&\overset{\text{(C)}}{\iff} \boldsymbol{c} \in \mathcal{P}^*_{F,\varphi(i)}(\boldsymbol{b}),
\end{aligned}$$

where (A) follows from (4), (B) follows from (i) of this lemma, and (C) follows from (4). This shows that $\mathcal{P}^*_{F',i}(\boldsymbol{b}) = \mathcal{P}^*_{F,\varphi(i)}(\boldsymbol{b})$. We can prove $\bar{\mathcal{P}}^*_{F',i}(\boldsymbol{b}) = \bar{\mathcal{P}}^*_{F,\varphi(i)}(\boldsymbol{b})$ by the same way using (5).

(Proof of (iii)): For any $i' \in [F']$ and $j \in [F]$, we have

$$\begin{aligned}
\sum_{j'\in\mathcal{A}_j} Q_{i',j'}(F') &\overset{\text{(A)}}{=} \sum_{j'\in\mathcal{A}_j} \sum_{\substack{s\in\mathcal{S}\\ \tau'_{i'}(s)=j'}} \mu(s) = \sum_{\substack{s\in\mathcal{S}\\ \tau'_{i'}(s)\in\mathcal{A}_j}} \mu(s)\\
&\overset{\text{(B)}}{=} \sum_{\substack{s\in\mathcal{S}\\ \varphi(\tau'_{i'}(s))=j}} \mu(s) \overset{\text{(C)}}{=} \sum_{\substack{s\in\mathcal{S}\\ \tau_{\varphi(i')}(s)=j}} \mu(s)\\
&\overset{\text{(D)}}{=} Q_{\varphi(i'),j}(F) = Q_{i,j}(F),
\end{aligned} \tag{A$\cdot$ 53}$$

where $i \coloneqq \varphi(i')$ and (A) follows from (17), (B) follows from (26), (C) follows from (24), and (D) follows from (17).

Thus, for any $j \in [F]$, we have

$$\begin{aligned}
\pi_j &= \sum_{j'\in\mathcal{A}_j} \pi'_{j'}\\
&\overset{\text{(A)}}{=} \sum_{j'\in\mathcal{A}_j} \sum_{i'\in[F']} \pi'_{i'} Q_{i',j'}(F')\\
&= \sum_{j'\in\mathcal{A}_j} \sum_{i\in[F]} \sum_{i'\in\mathcal{A}_i} \pi'_{i'} Q_{i',j'}(F')\\
&= \sum_{i\in[F]} \sum_{i'\in\mathcal{A}_i} \pi'_{i'} \sum_{j'\in\mathcal{A}_j} Q_{i',j'}(F')\\
&\overset{\text{(B)}}{=} \sum_{i\in[F]} \sum_{i'\in\mathcal{A}_i} \pi'_{i'} Q_{i,j}(F)
\end{aligned}$$

$$= \sum_{i \in [F]} Q_{i,j}(F) \sum_{i' \in \mathcal{A}_i} \pi'_{i'}$$
$$= \sum_{i \in [F]} Q_{i,j}(F)\pi_i, \tag{A·54}$$

where (A) follows since $\boldsymbol{\pi}'$ satisfies (19), and (B) follows from (A·53) and $i' \in \mathcal{A}_i$.

Also, we have

$$\sum_{i \in [F]} \pi_i = \sum_{i \in [F]} \sum_{i' \in \mathcal{A}_i} \pi'_{i'} = \sum_{i' \in [F']} \pi'_{i'} \stackrel{(A)}{=} 1, \tag{A·55}$$

where (A) follows since $\boldsymbol{\pi}'$ satisfies (20). By (A·54) and (A·55), $\boldsymbol{\pi}$ is a stationary distribution of F.

(Proof of (iv)): We have

$$\begin{aligned} F \in \mathscr{F}_{\mathrm{ext}} &\iff {}^{\forall} i \in [F]; \mathcal{P}^1_{F,i} \neq \emptyset \\ &\implies {}^{\forall} i' \in [F']; \mathcal{P}^1_{F,\varphi(i')} \neq \emptyset \\ &\stackrel{(A)}{\iff} {}^{\forall} i' \in [F']; \mathcal{P}^1_{F',i'} \neq \emptyset \\ &\iff F' \in \mathscr{F}_{\mathrm{ext}}, \end{aligned}$$

where (A) follows from (ii) of this lemma.

(Proof of (v)): By $F, F' \in \mathscr{F}_{\mathrm{reg}}$, the code-tuples F and F' have the unique stationary distributions $\boldsymbol{\pi}(F)$ and $\boldsymbol{\pi}(F')$, respectively. By (iii) of this lemma, we have

$${}^{\forall} j \in [F]; \pi_j(F) = \sum_{j' \in \mathcal{A}_j} \pi_{j'}(F'), \tag{A·56}$$

where

$$\mathcal{A}_i \coloneqq \{i' \in [F'] : \varphi(i') = i\} \tag{A·57}$$

for $i \in [F]$. Therefore, we have

$$\begin{aligned} L(F') &= \sum_{i' \in [F']} \pi_{i'}(F') L_{i'}(F') \\ &= \sum_{i \in [F]} \sum_{i' \in \mathcal{A}_i} \pi_{i'}(F') L_{i'}(F') \\ &\stackrel{(A)}{=} \sum_{i \in [F]} \sum_{i' \in \mathcal{A}_i} \pi_{i'}(F') L_{\varphi(i')}(F) \\ &\stackrel{(B)}{=} \sum_{i \in [F]} \sum_{i' \in \mathcal{A}_i} \pi_{i'}(F') L_i(F) \\ &= \sum_{i \in [F]} L_i(F) \sum_{i' \in \mathcal{A}_i} \pi_{i'}(F') \\ &\stackrel{(C)}{=} \sum_{i \in [F]} \pi_i(F) L_i(F) \\ &= L(F) \end{aligned}$$

as desired, where (A) follows from (23) (cf. Remark 2), (B) follows from (A·57) and $i' \in \mathcal{A}_i$, and (C) follows from (A·56).

(Proof of (vi)): For any $i \in [F']$ and $s \in \mathcal{S}$, we have

$$\begin{aligned} \mathcal{P}^k_{F',\tau'_i(s)} \cap \bar{\mathcal{P}}^k_{F',i}(f'_i(s)) &\stackrel{(A)}{=} \mathcal{P}^k_{F,\varphi(\tau'_i(s))} \cap \bar{\mathcal{P}}^k_{F,\varphi(i)}(f'_i(s)) \\ &\stackrel{(B)}{=} \mathcal{P}^k_{F,\tau_{\varphi(i)}(s)} \cap \bar{\mathcal{P}}^k_{F,\varphi(i)}(f_{\varphi(i)}(s)) \\ &\stackrel{(C)}{=} \emptyset, \end{aligned}$$

where (A) follows from (ii) of this lemma, (B) follows from (23) and (24), and (C) follows from $F \in \mathscr{F}_{k\text{-dec}}$. Namely, F' satisfies Definition 5 (i).

Choose $i \in [F']$ and $s, s' \in \mathcal{S}$ such that $s \neq s'$ and $f'_i(s) = f'_i(s')$ arbitrarily. Then by (23), we have

$$f_{\varphi(i)}(s) = f'_i(s) = f'_i(s') = f_{\varphi(i)}(s'). \tag{A·58}$$

Thus, we obtain

$$\begin{aligned} \mathcal{P}^k_{F',\tau'_i(s)} \cap \mathcal{P}^k_{F',\tau'_i(s')} &\stackrel{(A)}{=} \mathcal{P}^k_{F,\varphi(\tau'_i(s))} \cap \mathcal{P}^k_{F,\varphi(\tau'_i(s'))} \\ &\stackrel{(B)}{=} \mathcal{P}^k_{F,\tau_{\varphi(i)}(s)} \cap \mathcal{P}^k_{F,\tau_{\varphi(i)}(s')} \stackrel{(C)}{=} \emptyset, \end{aligned}$$

where (A) follows from (ii) of this lemma, (B) follows from (24), (C) follows from (A·58) and $F \in \mathscr{F}_{k\text{-dec}}$. Namely, F' satisfies Definition 5 (ii). □

C.4 Proof of Lemma 8

To prove Lemma 8, we first prove the following Lemmas 22–24. Lemmas 22 and 23 relate to closed sets defined in Appendix C.2.

Lemma 22. *For any $F \in \mathscr{F}$, the following statements (i) and (ii) hold.*

(i) $\mathcal{R}_F$ is closed.
(ii) For any non-empty closed set $\mathcal{I} \subseteq [F]$, we have $\mathcal{R}_F \subseteq \mathcal{I}$.

Proof of Lemma 22. (Proof of (i)): Choose $i \in \mathcal{R}_F$ and $s \in \mathcal{S}$ arbitrarily. For any $j \in [F]$, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau^*_j(\boldsymbol{x}) = i$, which leads to

$$\tau^*_j(\boldsymbol{x}s) \stackrel{(A)}{=} \tau_{\tau^*_j(\boldsymbol{x})}(s) = \tau_i(s), \tag{A·59}$$

where (A) follows from Lemma 1 (ii). This shows $\tau_i(s) \in \mathcal{R}_F$.

(Proof of (ii)): Choose $i \in \mathcal{R}_F$ arbitrarily. We prove $i \in \mathcal{I}$ by contradiction assuming the contrary $i \notin \mathcal{I}$. Since $\mathcal{I} \neq \emptyset$, we can choose $j \in \mathcal{I}$. By $i \in \mathcal{R}_F$, there exists $\boldsymbol{x} = x_1 x_2 \ldots x_n \in \mathcal{S}^*$ such that $\tau^*_j(\boldsymbol{x}) = i$. We define $i_l \coloneqq \tau^*_j(x_1 x_2 \ldots x_l)$ for $l = 0, 1, 2, \ldots, n$. Since $i_0 = \tau^*_j(\lambda) = j \in \mathcal{I}$ and $i_n = \tau^*_j(\boldsymbol{x}) = i \notin \mathcal{I}$, there exists an integer $0 \leq l < n$ such that $i_l \in \mathcal{I}$ and $i_{l+1} = \tau_{i_l}(x_{l+1}) \notin \mathcal{I}$. This conflicts with that $\mathcal{I}$ is closed. □

Lemma 23. *For any $F \in \mathscr{F}$ and non-empty closed set $\mathcal{I} \subseteq [F]$, the following statements (i) and (ii) hold.*

(i) There exist $F' \in \mathscr{F}^{(|\mathcal{I}|)}$ and an injective homomorphism $\varphi : [F'] \to [F]$ from F' to F such that $\mathcal{I} = \varphi([F']) \coloneqq \{\varphi(i) : i \in [F']\}$.
(ii) There exists a stationary distribution $\boldsymbol{\pi} = (\pi_0,$

$\pi_1, \ldots, \pi_{|F|-1})$ *of* F *such that* $\pi_i = 0$ *for any* $i \in [F] \setminus \mathcal{I}$.

Proof of Lemma 23. (Proof of (i)): Suppose $\mathcal{I} = \{i_0, i_1, \ldots, i_{m-1}\}$, where $i_0 < i_1 < \cdots < i_{m-1}$ and $m = |\mathcal{I}|$. We define a mapping $\varphi : [m] \to [F]$ as $\varphi(j) = i_j$ for $j \in [m]$. Since φ is injective and $\varphi([m]) = \mathcal{I}$, we can consider the inverse mapping $\varphi^{-1} : \mathcal{I} \to [m]$, which maps $\varphi(i)$ to i for any $i \in [m]$. Also, we define $F'(f', \tau') \in \mathscr{F}^{(m)}$ as

$$f'_i(s) = f_{\varphi(i)}(s), \tag{A· 60}$$

$$\tau'_i(s) = \varphi^{-1}(\tau_{\varphi(i)}(s)) \tag{A· 61}$$

for $i \in [F']$ and $s \in \mathcal{S}$. Since $\mathcal{I}$ is closed, we have $\tau_{\varphi(i)}(s) \in \mathcal{I}$ and thus $\tau'_i(s) = \varphi^{-1}(\tau_{\varphi(i)}(s)) \in [m] = [F']$; that is, F' is indeed well-defined. We can see that φ is a homomorphism from F' to F directly from (A· 60) and (A· 61).

(Proof of (ii)): By (i) of this lemma, there exist $F' \in \mathscr{F}$ and an injective homomorphism $\varphi : [F'] \to [F]$ from F' to F such that

$$\varphi([F']) = \mathcal{I}. \tag{A· 62}$$

By Lemma 6, we can choose a stationary distribution $\boldsymbol{\pi}'$ of F'. By Lemma 7 (iii), the vector $\boldsymbol{\pi} \in \mathbb{R}^{|F|}$ defined as (25) is a stationary distribution of F. This vector $\boldsymbol{\pi}$ is a desired stationary distribution because $\mathcal{A}_i = \{i' \in [F'] : \varphi(i') = i\} = \emptyset$ holds for any $i \in [F] \setminus \mathcal{I}$ by (A· 62). □

Lemma 24. *For any* $F \in \mathscr{F}$*, If* $\mathcal{R}_F = \emptyset$*, then there exist* $p, q \in [F]$ *such that* $\mathcal{I}_p \cap \mathcal{I}_q = \emptyset$*, where* $\mathcal{I}_i := \{\tau^*_i(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}^*\}$ *for* $i \in [F]$.

Proof of Lemma 24. We first show that for any $i, j \in [F]$, we have

$$j \in \mathcal{I}_i \implies \mathcal{I}_j \subseteq \mathcal{I}_i. \tag{A· 63}$$

Assume $j \in \mathcal{I}_i$ and choose $p \in \mathcal{I}_j$ arbitrarily. Then there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau^*_j(\boldsymbol{x}) = p$. Also, by $j \in \mathcal{I}_i$, there exists $\boldsymbol{y} \in \mathcal{S}^*$ such that $\tau^*_i(\boldsymbol{y}) = j$. Therefore, we have

$$\tau^*_i(\boldsymbol{y}\boldsymbol{x}) \overset{\text{(A)}}{=} \tau^*_{\tau^*_i(\boldsymbol{y})}(\boldsymbol{x}) = \tau^*_j(\boldsymbol{x}) = p,$$

where (A) follows from Lemma 1 (ii). This leads to $p \in \mathcal{I}_i$ and thus we obtain (A· 63).

Now, we prove Lemma 24 by proving its contraposition. Namely, we show $\mathcal{R}_F \neq \emptyset$ assuming that

$$^\forall i, j \in [F]; \mathcal{I}_i \cap \mathcal{I}_j \neq \emptyset. \tag{A· 64}$$

We can see that

$$\mathcal{R}_F = \bigcap_{i \in [F]} \mathcal{I}_i$$

because for any $j \in [F]$, it holds that

$$\begin{aligned} j \in \bigcap_{i \in [F]} \mathcal{I}_i &\iff {}^\forall i \in [F]; j \in \mathcal{I}_i \\ &\iff {}^\forall i \in [F]; {}^\exists \boldsymbol{x} \in \mathcal{S}^*; \tau^*_i(\boldsymbol{x}) = j \\ &\iff j \in \mathcal{R}_F. \end{aligned}$$

Thus, to show $\mathcal{R}_F \neq \emptyset$, it suffices to show that

$$\bigcap_{i \in [r]} \mathcal{I}_i \neq \emptyset \tag{A· 65}$$

for any $r = 1, 2, \ldots, |F|$ since the case $r = |F|$ gives the desired result.

We prove (A· 65) by induction for r. The base case $r = 1$ is trivial since $\mathcal{I}_0 \ni 0$. We consider the induction step for $r \geq 2$. By the induction hypothesis, we have $\bigcap_{i \in [r-1]} \mathcal{I}_i \neq \emptyset$. Therefore, we can choose $j \in [F]$ such that $j \in \mathcal{I}_i$ for any $i \in [r-1]$. By (A· 63), we have $\mathcal{I}_j \subseteq \mathcal{I}_i$ for any $i \in [r-1]$ and thus

$$\mathcal{I}_j \subseteq \bigcap_{i \in [r-1]} \mathcal{I}_i. \tag{A· 66}$$

Hence, we obtain

$$\bigcap_{i \in [r]} \mathcal{I}_i = \Big(\bigcap_{i \in [r-1]} \mathcal{I}_i \Big) \cap \mathcal{I}_{r-1} \overset{\text{(A)}}{\supseteq} \mathcal{I}_j \cap \mathcal{I}_{r-1} \overset{\text{(B)}}{\neq} \emptyset$$

as desired, where (A) follows from (A· 66), and (B) follows from (A· 64). □

Proof of Lemma 8. (Proof of (i)): (Necessity) We assume $\mathcal{R}_F = \emptyset$ and show that F has two distinct stationary distributions. By Lemma 24, we can choose $p, q \in [F]$ such that

$$\mathcal{I}_p \cap \mathcal{I}_q = \emptyset. \tag{A· 67}$$

We can see that $\mathcal{I}_p$ is not empty since $\mathcal{I}_p \ni p$ and also see that $\mathcal{I}_p$ is closed because for any $i \in \mathcal{I}_p$, we have

$$\{\tau_i(s) : s \in \mathcal{S}\} \subseteq \{\tau^*_i(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}^*\} = \mathcal{I}_i \overset{\text{(A)}}{\subseteq} \mathcal{I}_p, \tag{A· 68}$$

where (A) follows from (A· 63). By the same argument, also $\mathcal{I}_q$ is a non-empty closed set. Therefore, by Lemma 23 (ii), there exist stationary distributions $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1})$ and $\boldsymbol{\pi}' = (\pi'_0, \pi'_1, \ldots, \pi'_{|F|-1})$ of F such that

$$^\forall i \in [F] \setminus \mathcal{I}_p; \pi_i = 0 \tag{A· 69}$$

and

$$^\forall i \in [F] \setminus \mathcal{I}_q; \pi'_i = 0. \tag{A· 70}$$

Since $\boldsymbol{\pi}$ satisfies (20), we have $\pi_j > 0$ for some $j \in [F]$. By (A· 69) and (A· 67), it must hold that $j \in \mathcal{I}_p \subseteq [F] \setminus \mathcal{I}_q$. Hence, we obtain $\pi'_j = 0 < \pi_j$ by (A· 70). This shows $\boldsymbol{\pi} \neq \boldsymbol{\pi}'$. Therefore, we conclude that F has two distinct stationary distributions as desired.

(Sufficiency) We prove $\mathcal{R}_F = \emptyset$ assuming that there exist two distinct stationary distributions $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{|F|-1})$ and $\boldsymbol{\pi}' = (\pi'_0, \pi'_1, \ldots, \pi'_{|F|-1})$ of F. Then $\boldsymbol{x} = (x_0, x_1, \ldots, x_{|F|-1}) := \boldsymbol{\pi} - \boldsymbol{\pi}' \neq \boldsymbol{0}$ satisfies

$$\boldsymbol{x}Q(F) = \boldsymbol{\pi}Q(F) - \boldsymbol{\pi}'Q(F) \overset{\text{(A)}}{=} \boldsymbol{\pi} - \boldsymbol{\pi}' = \boldsymbol{x}, \tag{A· 71}$$

$$\sum_{i\in[F]} x_i = \sum_{i\in[F]} \pi_i - \sum_{i\in[F]} \pi'_i \overset{\text{(B)}}{=} 1 - 1 = 0, \tag{A· 72}$$

where (A) follows from (19), and (B) follows from (20). Thus, by $\boldsymbol{x} \neq \boldsymbol{0}$ and (A· 72), both of $\mathcal{I}_+ := \{i \in [F] : x_i > 0\}$ and $\mathcal{I}_- := \{i \in [F] : x_i < 0\}$ are non-empty sets. Also, both of $\mathcal{I}_+$ and $\mathcal{I}_-$ are closed by (A· 71) and Lemma 21 stated in Appendix C.2. Therefore, by Lemma 22 (ii), we obtain $\mathcal{R}_F \subseteq \mathcal{I}_+$ and $\mathcal{R}_F \subseteq \mathcal{I}_-$, which conclude $\mathcal{R}_F \subseteq \mathcal{I}_+ \cap \mathcal{I}_- = \emptyset$ as desired.

(Proof of (ii)): We show $\mathcal{R}_F = \mathcal{I}_+ := \{i \in [F] : \pi_i(F) > 0\}$.

($\mathcal{R}_F \subseteq \mathcal{I}_+$) By (20), the set $\mathcal{I}_+$ is not empty. Also, by (19) and Lemma 21 stated in Appendix C.2, the set $\mathcal{I}_+$ is closed. Hence, we obtain $\mathcal{R}_F \subseteq \mathcal{I}_+$ by Lemma 22 (ii).

($\mathcal{R}_F \supseteq \mathcal{I}_+$) Since $\mathcal{R}_F$ is closed by Lemma 22 (i), we see from Lemma 23 (ii) that the unique stationary distribution $\boldsymbol{\pi}(F)$ satisfies $\pi_i(F) = 0$ for any $i \in [F] \setminus \mathcal{R}_F$. Therefore, we obtain $\mathcal{R}_F \supseteq \mathcal{I}_+$. □

C.5 Proof of Lemma 10

The proof of Lemma 10 relies on Lemmas 22 and 23 stated in Appendix C.4.

Proof of Lemma 10. Since $\mathcal{R}_F$ is closed by Lemma 22 (i), we see from Lemma 23 (i) that there exist $\bar{F}(\bar{f}, \bar{\tau}) \in \mathscr{F}$ and an injective homomorphism $\varphi : [\bar{F}] \to [F]$ from F' to F such that $\varphi([\bar{F}]) = \mathcal{R}_F$. Now, it suffices to show $\bar{F} \in \mathscr{F}_{\mathrm{irr}}$.

For any $i, j \in [\bar{F}]$, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that

$$\tau^*_{\varphi(i)}(\boldsymbol{x}) = \varphi(j) \tag{A· 73}$$

by $\varphi(j) \in \varphi([\bar{F}]) = \mathcal{R}_F$. Thus, for any $i, j \in [F]$, we have

$$\begin{aligned}\bar{\tau}^*_i(\boldsymbol{x}) = \varphi^{-1}(\varphi(\bar{\tau}^*_i(\boldsymbol{x}))) &\overset{\text{(A)}}{=} \varphi^{-1}(\tau^*_{\varphi(i)}(\boldsymbol{x}))\\ &\overset{\text{(B)}}{=} \varphi^{-1}(\varphi(j)) = j,\end{aligned}$$

where (A) follows from Lemma 7 (i), and (B) follows from (A· 73). Therefore, $\bar{F} \in \mathscr{F}_{\mathrm{irr}}$ holds. □

C.6 Proof of Lemma 12

Proof of Lemma 12. (Proof of (i)): We prove only $\mathcal{P}^k_{F,i}(\boldsymbol{b}) \supseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$ for any $i \in [F]$ and $\boldsymbol{b} \in C^*$ because we can prove $\mathcal{P}^k_{F,i}(\boldsymbol{b}) \subseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$, $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) \supseteq \bar{\mathcal{P}}^k_{F',i}(\boldsymbol{b})$, and $\bar{\mathcal{P}}^k_{F,i}(\boldsymbol{b}) \subseteq \bar{\mathcal{P}}^k_{F',i}(\boldsymbol{b})$ in the similar way. To prove $\mathcal{P}^k_{F,i}(\boldsymbol{b}) \supseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$, it suffices to prove that for any $(i, \boldsymbol{x}, \boldsymbol{b}, \boldsymbol{c}) \in [F] \times \mathcal{S}^+ \times C^* \times C^{\leq k}$, we have

$$\begin{aligned}&(f'^*_i(\boldsymbol{x}) \geq \boldsymbol{bc}, f'_i(x_1) \geq \boldsymbol{b})\\ &\quad\Longrightarrow\ ^\exists\boldsymbol{x}' \in \mathcal{S}^+; (f^*_i(\boldsymbol{x}') \geq \boldsymbol{bc}, f_i(x'_1) \geq \boldsymbol{b})\end{aligned} \tag{A· 74}$$

because this shows that for any $i \in [F']$, $\boldsymbol{b} \in C^*$, and $\boldsymbol{c} \in C^k$, we have

$$\begin{aligned}\boldsymbol{c} \in \mathcal{P}^k_{F',i}(\boldsymbol{b}) &\overset{\text{(A)}}{\iff} {}^\exists\boldsymbol{x} \in \mathcal{S}^+; (f'^*_i(\boldsymbol{x}) \geq \boldsymbol{bc}, f_i(x_1) \geq \boldsymbol{b})\\ &\overset{\text{(B)}}{\implies} {}^\exists\boldsymbol{x}' \in \mathcal{S}^+; (f^*_i(\boldsymbol{x}') \geq \boldsymbol{bc}, f_i(x'_1) \geq \boldsymbol{b})\\ &\overset{\text{(C)}}{\iff} \boldsymbol{c} \in \mathcal{P}^k_{F,i}(\boldsymbol{b})\end{aligned}$$

as desired, where (A) follows from (4), (B) follows from (A· 74), and (C) follows from (4).

Choose $(i, \boldsymbol{x}, \boldsymbol{b}, \boldsymbol{c}) \in [F] \times \mathcal{S}^+ \times C^* \times C^{\leq k}$ arbitrarily and assume

$$f'^*_i(\boldsymbol{x}) \geq \boldsymbol{bc} \tag{A· 75}$$

and

$$f'_i(x_1) \geq \boldsymbol{b}. \tag{A· 76}$$

Then we have

$$f_i(x_1) \overset{\text{(A)}}{=} f'_i(x_1) \overset{\text{(B)}}{\geq} \boldsymbol{b}, \tag{A· 77}$$

where (A) follows from the assumption (a) of this lemma, and (B) follows from (A· 76).

We prove (A· 74) by induction for $|\boldsymbol{x}|$. For the base case $|\boldsymbol{x}| = 1$, we have

$$f^*_i(\boldsymbol{x}) = f_i(x_1) \overset{\text{(A)}}{=} f'_i(x_1) = f'^*_i(\boldsymbol{x}) \overset{\text{(B)}}{\geq} \boldsymbol{bc}, \tag{A· 78}$$

where (A) follows from the assumption (a) of this lemma, and (B) follows from (A· 75). By (A· 78) and (A· 77), the claim (A· 74) holds for the base case $|\boldsymbol{x}| = 1$.

We consider the induction step for $|\boldsymbol{x}| \geq 2$. We have

$$\begin{aligned}f_i(x_1)f'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x})) &\overset{\text{(A)}}{=} f'_i(x_1)f'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(B)}}{=} f'^*_i(\boldsymbol{x}) \overset{\text{(C)}}{\geq} \boldsymbol{bc},\end{aligned} \tag{A· 79}$$

where (A) follows from the assumption (a) of this lemma, (B) follows from (2), and (C) follows from (A· 75). Therefore, $f_i(x_1) \geq \boldsymbol{bc}$ or $f_i(x_1) \prec \boldsymbol{bc}$ holds. In the case $f_i(x_1) \geq \boldsymbol{bc}$, clearly $\boldsymbol{x}' := x_1$ satisfies $f^*_i(\boldsymbol{x}') \geq \boldsymbol{bc}$ and $f_i(x'_1) = f_i(x_1) \geq \boldsymbol{b}$ by (A· 77) as desired. Thus, now we assume $f_i(x_1) \prec \boldsymbol{bc}$. Then we have

$$\begin{aligned}|f_i(x_1)^{-1}\boldsymbol{bc}| &= -|f_i(x_1)| + |\boldsymbol{b}| + |\boldsymbol{c}|\\ &\overset{\text{(A)}}{=} -|f'_i(x_1)| + |\boldsymbol{b}| + |\boldsymbol{c}|\\ &\overset{\text{(B)}}{\leq} |\boldsymbol{c}| \leq k,\end{aligned} \tag{A· 80}$$

where (A) follows from the assumption (a) of this lemma, and (B) follows from (A· 76). By (A· 79), we have

$$f'^*_{\tau'_i(x_1)}(\mathrm{suff}(\boldsymbol{x})) \geq f_i(x_1)^{-1}\boldsymbol{bc}. \tag{A· 81}$$

By (A· 80) and (A· 81), we can apply the induction hypothesis to $(\tau'_i(x_1), \mathrm{suff}(\boldsymbol{x}), \lambda, f_i(x_1)^{-1}\boldsymbol{bc})$. Hence, there exists $\boldsymbol{x}' \in \mathcal{S}^*$ such that $f^*_{\tau'_i(x_1)}(\boldsymbol{x}') \geq f_i(x_1)^{-1}\boldsymbol{bc}$, which leads to

$f_i(x_1)^{-1}\boldsymbol{bc} \in \mathcal{P}^{k'}_{F,\tau'_i(x_1)}$ by (8), where $k' := |f_i(x_1)^{-1}\boldsymbol{bc}|$. By Lemma 4 (i), there exists $\boldsymbol{c}' \in C^{k-k'}$ such that

$$f_i(x_1)^{-1}\boldsymbol{bcc}' \in \mathcal{P}^{k}_{F,\tau'_i(x_1)} \overset{\text{(A)}}{=} \mathcal{P}^{k}_{F,\tau_i(x_1)}, \tag{A·82}$$

where (A) follows from the assumption (b) of this lemma. By (8), there exists $\boldsymbol{x}'' \in \mathcal{S}^*$ such that

$$f^*_{\tau_i(x_1)}(\boldsymbol{x}'') \geq f_i(x_1)^{-1}\boldsymbol{bcc}' \geq f_i(x_1)^{-1}\boldsymbol{bc}. \tag{A·83}$$

Thus, we have

$$\begin{aligned} f^*_i(x_1\boldsymbol{x}'') &\overset{\text{(A)}}{=} f_i(x_1)f^*_{\tau_i(x_1)}(\boldsymbol{x}'') \\ &\overset{\text{(B)}}{\geq} f_i(x_1)f_i(x_1)^{-1}\boldsymbol{bc} \\ &= \boldsymbol{bc}, \end{aligned} \tag{A·84}$$

where (A) follows from (2), and (B) follows from (A·83). The induction is completed by (A·77) and (A·84).

(Proof of (ii)): We have

$$\begin{aligned} F \in \mathscr{F}_{\text{ext}} &\iff {}^\forall i \in [F]; \mathcal{P}^1_{F,i} \neq \emptyset \\ &\overset{\text{(A)}}{\iff} {}^\forall i \in [F']; \mathcal{P}^1_{F',i} \neq \emptyset \\ &\iff F' \in \mathscr{F}_{\text{ext}}, \end{aligned}$$

where (A) follows from (i) of this lemma.

(Proof of (iii)): For any $i \in [F']$ and $s \in \mathcal{S}$, we have

$$\begin{aligned} \mathcal{P}^k_{F',\tau'_i(s)} \cap \bar{\mathcal{P}}^k_{F',i}(f'_i(s)) &\overset{\text{(A)}}{=} \mathcal{P}^k_{F,\tau'_i(s)} \cap \bar{\mathcal{P}}^k_{F,i}(f'_i(s)) \\ &\overset{\text{(B)}}{=} \mathcal{P}^k_{F,\tau_i(s)} \cap \bar{\mathcal{P}}^k_{F,i}(f_i(s)) \overset{\text{(C)}}{=} \emptyset, \end{aligned}$$

where (A) follows from (i) of this lemma, (B) follows from the assumptions (a) and (b), and (C) follows from $F \in \mathscr{F}_{k\text{-dec}}$. Namely, F' satisfies Definition 5 (i).

For any $i \in [F']$ and $s, s' \in \mathcal{S}$ such that $s \neq s'$ and $f'_i(s) = f'_i(s')$, we have

$$f_i(s) = f_i(s') \tag{A·85}$$

by the assumption (a), and we have

$$\begin{aligned} \mathcal{P}^k_{F',\tau'_i(s)} \cap \mathcal{P}^k_{F',\tau'_i(s')} &\overset{\text{(A)}}{\subseteq} \mathcal{P}^k_{F,\tau'_i(s)} \cap \mathcal{P}^k_{F,\tau'_i(s')} \\ &\overset{\text{(B)}}{=} \mathcal{P}^k_{F,\tau_i(s)} \cap \mathcal{P}^k_{F,\tau_i(s')} \overset{\text{(C)}}{=} \emptyset, \end{aligned}$$

where (A) follows from (i) of this lemma, (B) follows from the assumptions (b), and (C) follows from $F \in \mathscr{F}_{k\text{-dec}}$ and (A·85). Namely, F' satisfies Definition 5 (ii). □

C.7 Proof of Lemma 13

Proof of Lemma 13. We show $\mathcal{R}_{F'} \ni p$ since this implies $F' \in \mathscr{F}_{\text{reg}}$ by Lemma 8 (i). Namely, we show that for any $j \in [F']$, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau'^*_j(\boldsymbol{x}) = p$.

For $j = p$, the sequence $\boldsymbol{x} := \lambda$ satisfies $\tau'^*_j(\boldsymbol{x}) = p$ by (3). Thus, we now consider the case $j \neq p$. Choose $j \in [F'] \setminus \{p\}$ arbitrarily. Since $p \in \mathcal{R}_F$ by $F \in \mathscr{F}_{\text{irr}}$, there exists $\boldsymbol{x} = x_1x_2\ldots x_n \in \mathcal{S}^+$ such that $\tau^*_j(\boldsymbol{x}) = p$. Let $r \geq 1$ be the minimum positive integer such that

$$\tau^*_j(x_1x_2\ldots x_r) \in \mathcal{I}. \tag{A·86}$$

Note that there exists such an integer $r \leq n$ since $\tau^*_j(\boldsymbol{x}) = \tau^*_j(x_1x_2\ldots x_n) = p \in \mathcal{I}$. We show that

$$\tau'^*_j(x_1x_2\ldots x_{r'}) = \tau^*_j(x_1x_2\ldots x_{r'}) \tag{A·87}$$

for any $r' = 1, 2, \ldots, r-1$ by induction for r'. For the base case $r' = 0$, we have $\tau'^*_j(\lambda) = j = \tau^*_j(\lambda)$ by (3). We consider the induction step for $r' \geq 1$. We have

$$\begin{aligned} \tau'^*_j(x_1x_2\ldots x_{r'}) &\overset{\text{(A)}}{=} \tau'_{\tau'^*_j(x_1x_2\ldots x_{r'-1})}(x_{r'}) \\ &\overset{\text{(B)}}{=} \tau'_{\tau^*_j(x_1x_2\ldots x_{r'-1})}(x_{r'}) \\ &\overset{\text{(C)}}{=} \tau_{\tau^*_j(x_1x_2\ldots x_{r'-1})}(x_{r'}) \\ &\overset{\text{(D)}}{=} \tau^*_j(x_1x_2\ldots x_{r'}) \end{aligned}$$

as desired, where (A) follows from Lemma 1 (ii), (B) follows from the induction hypothesis, (C) is obtained by applying the second case of (33) since $\tau_{\tau^*_j(x_1x_2\ldots x_{r'-1})}(x_{r'}) = \tau^*_j(x_1x_2\ldots x_{r'}) \notin \mathcal{I}$ by $r' \leq r-1$ and the minimality of r, and (D) follows from Lemma 1 (ii).

Thus, we obtain

$$\begin{aligned} \tau'^*_j(x_1x_2\ldots x_r) &\overset{\text{(A)}}{=} \tau'_{\tau'^*_j(x_1x_2\ldots x_{r-1})}(x_r) \\ &\overset{\text{(B)}}{=} \tau'_{\tau^*_j(x_1x_2\ldots x_{r-1})}(x_r) \overset{\text{(C)}}{=} p \end{aligned}$$

as desired, where (A) follows from Lemma 1 (ii), (B) follows from (A·87), and (C) follows from (A·86) and the first case of (33). □

C.8 Proof of Lemma 15

Proof of Lemma 15. Let $p \in \arg\min_{i \in [F]}(h_i(F) - h_i(F'))$. Then it holds that

$${}^\forall i \in [F]; h_i(F') - h_p(F') \leq h_i(F) - h_p(F). \tag{A·88}$$

We have

$$\begin{aligned} &\sum_{i \in [F]} (h_i(F) - h_p(F))Q_{p,i}(F) \\ &= \sum_{i \in [F]} (h_i(F) - h_p(F)) \sum_{\substack{s \in \mathcal{S} \\ \tau_p(s) = i}} \mu(s) \\ &= \sum_{i \in [F]} \sum_{\substack{s \in \mathcal{S} \\ \tau_p(s) = i}} (h_i(F) - h_p(F))\mu(s) \\ &= \sum_{i \in [F]} \sum_{\substack{s \in \mathcal{S} \\ \tau_p(s) = i}} (h_{\tau_p(s)}(F) - h_p(F))\mu(s) \end{aligned}$$

$$= \sum_{s \in \mathcal{S}} (h_{\tau_p(s)}(F) - h_p(F))\mu(s). \tag{A· 89}$$

Similarly, we have

$$\sum_{i \in [F]} (h_i(F) - h_p(F)) Q_{p,i}(F') = \sum_{s \in \mathcal{S}} (h_{\tau'_p(s)}(F) - h_p(F))\mu(s). \tag{A· 90}$$

Hence, we obtain

$$\begin{aligned} L(F') &\overset{\text{(A)}}{=} L_p(F') + \sum_{i \in [F]} (h_i(F') - h_p(F')) Q_{p,i}(F') \\ &\overset{\text{(B)}}{\leq} L_p(F') + \sum_{i \in [F]} (h_i(F) - h_p(F)) Q_{p,i}(F') \\ &\overset{\text{(C)}}{=} L_p(F') + \sum_{s \in \mathcal{S}} (h_{\tau'_p(s)}(F) - h_p(F))\mu(s) \\ &\overset{\text{(D)}}{\leq} L_p(F) + \sum_{s \in \mathcal{S}} (h_{\tau_p(s)}(F) - h_p(F))\mu(s) \\ &\overset{\text{(E)}}{=} L_p(F) + \sum_{i \in [F]} (h_i(F) - h_p(F)) Q_{p,i}(F) \\ &\overset{\text{(F)}}{=} L(F) \end{aligned}$$

as desired, where (A) follows from (34), (B) follows from (A· 88), (C) follows from (A· 90), (D) follows from the assumptions (a) and (b) of this lemma, (E) follows from (A· 89), and (F) follows from (34). □

C.9 Proof of Lemma 16

Proof of Lemma 16. (Proof of (i)): We prove by induction for $|\boldsymbol{z}|$. For the base case $|\boldsymbol{z}| = 0$, we have $\tau'^*_{\langle\lambda\rangle}(\lambda) = \langle\lambda\rangle$ by (3). We consider the induction step for $|\boldsymbol{z}| \geq 1$. We have

$$\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \overset{\text{(A)}}{=} \tau'_{\tau'^*_{\langle\lambda\rangle}(\mathrm{pref}(\boldsymbol{z}))}(z_n) \overset{\text{(B)}}{=} \tau'_{\langle \mathrm{pref}(\boldsymbol{z})\rangle}(z_n) \overset{\text{(C)}}{=} \langle\boldsymbol{z}\rangle, \tag{A· 91}$$

where $\boldsymbol{z} = z_1 z_2 \ldots z_n$ and (A) follows from Lemma 1 (ii), (B) follows from the induction hypothesis, and (C) follows from the first case of (49).

(Proof of (ii)): It suffices to show that $\langle\lambda\rangle \in \mathcal{R}_{F'}$ because it guarantees that for any $j \in [F']$, there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau'^*_j(\boldsymbol{x}) = \langle\lambda\rangle$, which leads to that for any $\boldsymbol{z} \in \mathcal{S}^{\leq L}$, we have

$$\tau'^*_j(\boldsymbol{x}\boldsymbol{z}) \overset{\text{(A)}}{=} \tau'^*_{\tau'^*_j(\boldsymbol{x})}(\boldsymbol{z}) = \tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \overset{\text{(B)}}{=} \langle\boldsymbol{z}\rangle \tag{A· 92}$$

as desired, where (A) follows from Lemma 1 (ii), and (B) follows from (i) of this lemma.

To prove $\langle\lambda\rangle \in \mathcal{R}_{F'}$, we show that there exists $\boldsymbol{x} \in \mathcal{S}^*$ such that $\tau'^*_j(\boldsymbol{x}) = \langle\lambda\rangle$ for the following two cases separately: (I) the case $j \in [F]$ and (II) the case $j = [F'] \setminus [F]$.

(I) The case $j \in [F]$: By the assumption that $p = \langle\lambda\rangle \in \mathcal{R}_F$, there exists $\boldsymbol{x} = x_1 x_2 \ldots x_{n'} \in \mathcal{S}^*$ such that $\tau^*_j(\boldsymbol{x}) = \langle\lambda\rangle$. We choose the shortest $\boldsymbol{x}$ among such sequences. Then we can see $\tau'^*_j(x_1 x_2 \ldots x_r) = \tau^*_j(x_1 x_2 \ldots x_r)$ for any $r = 0, 1, 2, \ldots, n$ by induction for r. For the base case $r = 0$, we have $\tau'^*_j(\lambda) = j = \tau^*_j(\lambda)$ by (3). We consider the induction step for $r \geq 1$. We have

$$\begin{aligned} \tau'^*_j(x_1 x_2 \ldots x_r) &\overset{\text{(A)}}{=} \tau'_{\tau'^*_j(x_1 x_2 \ldots x_{r-1})}(x_r) \\ &\overset{\text{(B)}}{=} \tau'_{\tau^*_j(x_1 x_2 \ldots x_{r-1})}(x_r) \\ &\overset{\text{(C)}}{=} \tau_{\tau^*_j(x_1 x_2 \ldots x_{r-1})}(x_r) \\ &\overset{\text{(D)}}{=} \tau^*_j(x_1 x_2 \ldots x_r) \end{aligned}$$

as desired, where (A) follows from (3), (B) follows from the induction hypothesis, (C) follows from the third case of (49) since $\tau^*_j(x_1 x_2 \ldots x_{r-1}) \in [F] \setminus \{\langle\lambda\rangle\}$ by the definition of $\boldsymbol{x}$, and (D) follows from Lemma 1 (ii). Therefore, we obtain $\tau'^*_j(\boldsymbol{x}) = \tau^*_j(\boldsymbol{x}) = \langle\lambda\rangle$ as desired.

(II) The case where $j = [F'] \setminus [F]$: Then we have $j = \langle\boldsymbol{z}\rangle$ for some $\boldsymbol{z} \in \mathcal{S}^{\leq L}$. Choose $\boldsymbol{z'} = z'_1 z'_2 \ldots z'_{n'} \in \mathcal{S}^{L-|\boldsymbol{z}|+1}$ arbitrarily. We have

$$\begin{aligned} \tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{z'}) &\overset{\text{(A)}}{=} \tau'_{\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}\mathrm{pref}(\boldsymbol{z'}))}(z'_{n'}) \\ &\overset{\text{(B)}}{=} \tau'_{\langle\boldsymbol{z}\mathrm{pref}(\boldsymbol{z'})\rangle}(z'_{n'}) \overset{\text{(C)}}{=} \tau^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{z'}) \\ &= \tau^*_{|F|-1}(\boldsymbol{z}\boldsymbol{z'}) \in [F], \end{aligned}$$

where (A) follows from Lemma 1 (ii), (B) follows from (i) of this lemma and $\boldsymbol{z}\mathrm{pref}(\boldsymbol{z'}) \in \mathcal{S}^{\leq L}$, and (C) follows from the second case of (49) and $\boldsymbol{z}\mathrm{pref}(\boldsymbol{z'}) \in \mathcal{S}^{L}$. Hence, by the discussion for the case (I) above, there exists $\boldsymbol{x'} \in \mathcal{S}^*$ such that $\tau'^*_{\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{z'})}(\boldsymbol{x'}) = \langle\lambda\rangle$. Thus, $\boldsymbol{x} := \boldsymbol{z'}\boldsymbol{x'}$ satisfies

$$\begin{aligned} \tau'^*_{\langle\boldsymbol{z}\rangle}(\boldsymbol{x}) = \tau'^*_{\langle\boldsymbol{z}\rangle}(\boldsymbol{z'}\boldsymbol{x'}) &\overset{\text{(A)}}{=} \tau'^*_{\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z})}(\boldsymbol{z'}\boldsymbol{x'}) \overset{\text{(B)}}{=} \tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{z'}\boldsymbol{x'}) \\ &\overset{\text{(C)}}{=} \tau'^*_{\tau'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{z'})}(\boldsymbol{x'}) = \langle\lambda\rangle, \end{aligned}$$

where (A) follows from (i) of this lemma, (B) follows from Lemma 1 (ii), and (C) follows from Lemma 1 (ii).

□

C.10 Proof of Lemma 17

Proof of Lemma 17. (Proof of (i)): We prove by the induction for $|\boldsymbol{x}|$. For the base case $|\boldsymbol{x}| = 0$, we have $f''^*_{\langle\boldsymbol{z}\rangle}(\lambda) = \lambda = f'^*_{\langle\boldsymbol{z}\rangle}(\lambda)$ by (2). We consider the induction step for $|\boldsymbol{x}| \geq 1$ choosing $\boldsymbol{z} \in \mathcal{S}^{\leq L}$ arbitrarily and dividing into the following two cases: the case $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{x})$ and the other case.

- The case $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{z}\boldsymbol{x})$: We consider the following two cases separately: the case $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zx_1)$ and the case $f'^*_{\langle\lambda\rangle}(zx_1) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$.

– The case $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zx_1)$: We have

$$\begin{aligned}&f''^*_{\langle z\rangle}(\boldsymbol{x})\\ &\overset{\text{(A)}}{=} f''_{\langle z\rangle}(x_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(B)}}{=} f'^*_{\langle z\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(zx_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(C)}}{=} f'^*_{\langle z\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(zx_1)f'^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(D)}}{=} f'^*_{\langle z\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}),\end{aligned}$$

where (A) follows from (2) and Lemma 16 (i), (B) follows from the first case of (51) and $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zx_1)$, (C) follows from the second case of (53) by the induction hypothesis and $f'^*_{\langle\lambda\rangle}(zx_1) \not\prec \boldsymbol{d}$, and (D) follows from (2).

– The case $f'^*_{\langle\lambda\rangle}(zx_1) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$: We have

$$\begin{aligned}&f''^*_{\langle z\rangle}(\boldsymbol{x})\\ &\overset{\text{(A)}}{=} f''_{\langle z\rangle}(x_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(B)}}{=} f'_{\langle z\rangle}(x_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(C)}}{=} f'_{\langle z\rangle}(x_1)f'^*_{\langle\lambda\rangle}(zx_1)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}))\\ &\overset{\text{(D)}}{=} f'^*_{\langle z\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}),\end{aligned}$$

where (A) follows from (2) and Lemma 16 (i), (B) follows from the second case of (51) since $\boldsymbol{d} \npreceq f'^*_{\langle\lambda\rangle}(zx_1)$, (C) follows from the first case of (53) by the induction hypothesis and $f'^*_{\langle\lambda\rangle}(zx_1) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$, and (D) follows from (2).

• The other case: We have

$$\begin{aligned}f''^*_{\langle z\rangle}(\boldsymbol{x}) &\overset{\text{(A)}}{=} f''_{\langle z\rangle}(x_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(B)}}{=} f'_{\langle z\rangle}(x_1)f''^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(C)}}{=} f'_{\langle z\rangle}(x_1)f'^*_{\langle zx_1\rangle}(\mathrm{suff}(\boldsymbol{x}))\\ &\overset{\text{(D)}}{=} f'^*_{\langle z\rangle}(\boldsymbol{x}),\end{aligned}$$

where (A) follows from (2) and Lemma 16 (i), (B) follows from the second case of (51) since $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zx_1)$ does not hold, (C) follows from the second case of (53) by the induction hypothesis and that $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zx_1)$ does not hold, and (D) follows from (2).

(Proof of (ii)): Assume that

$$f''_{\langle z\rangle}(s) \prec f''_{\langle z\rangle}(s'). \tag{A·93}$$

In the case $f'^*_{\langle\lambda\rangle}(z) \not\prec \boldsymbol{d}$, we have

$$f'_{\langle z\rangle}(s) \overset{\text{(A)}}{=} f''_{\langle z\rangle}(s) \overset{\text{(B)}}{\prec} f''_{\langle z\rangle}(s') \overset{\text{(C)}}{=} f'_{\langle z\rangle}(s') \tag{A·94}$$

as desired, where (A) follows from the second case of (51) and $f'^*_{\langle\lambda\rangle}(z) \not\prec \boldsymbol{d}$, (B) follows from (A·93), and (C) follows from the second case of (51) and $f'^*_{\langle\lambda\rangle}(z) \not\prec \boldsymbol{d}$.

We consider the case $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d}$ dividing into four cases by whether $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs)$ and whether $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs')$.

• The case $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs), \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs')$: We have

$$\begin{aligned}&f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(z)f'^*_{\langle z\rangle}(s))\\ &\overset{\text{(A)}}{=} f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(zs)\\ &\overset{\text{(B)}}{=} f''_{\langle z\rangle}(s)\\ &\overset{\text{(C)}}{\prec} f''_{\langle z\rangle}(s')\\ &\overset{\text{(D)}}{=} f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^*_{\langle\lambda\rangle}(zs')\\ &\overset{\text{(E)}}{=} f'^*_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^*_{\langle\lambda\rangle}(z)f'_{\langle z\rangle}(s')),\end{aligned} \tag{A·95}$$

where (A) follows from Lemma 1 (i) and Lemma 16 (i), (B) follows from the first case of (51) and $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs)$, (C) follows from (A·93), (D) follows from the first case of (51) and $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs')$, and (E) follows from Lemma 1 (i) and Lemma 16 (i). Comparing both sides of (A·95), we obtain $f'_{\langle z\rangle}(s) \prec f'_{\langle z\rangle}(s')$ as desired.

• The case $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs), \boldsymbol{d} \npreceq f'^*_{\langle\lambda\rangle}(zs')$: We show that this case is impossible. We have

$$\begin{aligned}&f'^*_{\langle\lambda\rangle}(zs')\\ &\overset{\text{(A)}}{=} f'^*_{\langle\lambda\rangle}(z)f'_{\langle z\rangle}(s')\\ &\overset{\text{(B)}}{=} f'^*_{\langle\lambda\rangle}(z)f''_{\langle z\rangle}(s')\\ &\overset{\text{(C)}}{\succ} f'^*_{\langle\lambda\rangle}(z)f''_{\langle z\rangle}(s)\\ &\overset{\text{(D)}}{=} f'^*_{\langle\lambda\rangle}(z)f'^*_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}f'^*_{\langle\lambda\rangle}(zs)\\ &= \boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}\boldsymbol{d}\\ &\succeq \boldsymbol{d},\end{aligned}$$

where (A) follows from Lemma 1 (i) and Lemma 16 (i), (B) follows from the second case of (51) and $\boldsymbol{d} \npreceq f'^*_{\langle\lambda\rangle}(zs')$, (C) follows from (A·93), and (D) follows from the first case of (51) and $\boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs)$. This conflicts with $\boldsymbol{d} \npreceq f'^*_{\langle\lambda\rangle}(zs')$.

• The case $\boldsymbol{d} \npreceq f'^*_{\langle\lambda\rangle}(zs), \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(zs')$: We have

$$\begin{aligned}&f'^*_{\langle\lambda\rangle}(zs)\\ &\overset{\text{(A)}}{=} f'^*_{\langle\lambda\rangle}(z)f'_{\langle z\rangle}(s)\\ &\overset{\text{(B)}}{=} f'^*_{\langle\lambda\rangle}(z)f''_{\langle z\rangle}(s)\\ &\overset{\text{(C)}}{\prec} f'^*_{\langle\lambda\rangle}(z)f''_{\langle z\rangle}(s')\\ &\overset{\text{(D)}}{=} f'^*_{\langle\lambda\rangle}(z)f'^*_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}f'^*_{\langle\lambda\rangle}(zs')\end{aligned}$$

$$= \boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1} f'^{*}_{\langle\lambda\rangle}(zs'),$$

where (A) follows from Lemma 1 (i) and Lemma 16 (i), (B) follows from the second case of (51) and $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs)$, (C) follows from (A· 93), and (D) follows from the first case of (51) and $\boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(zs')$.
Therefore, we have at least one of $f'^{*}_{\langle\lambda\rangle}(zs) \prec \boldsymbol{d}$ and $f'^{*}_{\langle\lambda\rangle}(zs) \succeq \boldsymbol{d}$. Since $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs)$, we have $f'^{*}_{\langle\lambda\rangle}(zs) \prec \boldsymbol{d}$. Thus, we have $f'^{*}_{\langle\lambda\rangle}(zs) \prec \boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(zs')$, which leads to $f'_{\langle z\rangle}(s) \prec f'_{\langle z\rangle}(s')$ as desired.

- The case $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs), \boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs')$: We have

$$f'_{\langle z\rangle}(s) \overset{\text{(A)}}{=} f''_{\langle z\rangle}(s) \overset{\text{(B)}}{\prec} f''_{\langle z\rangle}(s') \overset{\text{(C)}}{=} f'_{\langle z\rangle}(s') \tag{A· 96}$$

as desired, where (A) follows from the second case of (51) and $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs)$, (B) follows from (A· 93), and (C) follows from the second case of (51) and $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(zs')$.

(Proof of (iii)): Choose $\boldsymbol{x} \in \mathcal{S}^{\geq L}$ arbitrarily. We have

$$|f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})| \overset{\text{(A)}}{=} |f^{*}_{\langle\lambda\rangle}(\boldsymbol{x})| \overset{\text{(B)}}{\geq} \left\lfloor \frac{|\boldsymbol{x}|}{|F|} \right\rfloor \geq \left\lfloor \frac{L}{|F|} \right\rfloor \overset{\text{(C)}}{=} \left\lfloor \frac{|F|(|\boldsymbol{d}|+1)}{|F|} \right\rfloor = |\boldsymbol{d}| + 1, \tag{A· 97}$$

where (A) follows from Lemma 7 (i) since φ defined in (50) is a homomorphism from F' to F, (B) follows from Lemma 5, and (C) follows from the definition of L. Also, we have

$$\begin{aligned} &|f''^{*}_{\langle\lambda\rangle}(\boldsymbol{x})| \\ &\overset{\text{(A)}}{\geq} \min\{|f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})|, |f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}))|\} \\ &= \min\{|f'^{*}_{\langle z\rangle}(\boldsymbol{x})|, |f'^{*}_{\langle z\rangle}(\boldsymbol{x})| - 1\} \\ &\overset{\text{(B)}}{\geq} |\boldsymbol{d}|, \end{aligned}$$

where (A) follows from (i) of this lemma, and (B) follows from (A· 97). □

C.11 Proof of Lemma 18

Proof of Lemma 18. (Proof of (i)): Assume

$$f''^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \boldsymbol{c}. \tag{A· 98}$$

We consider the following two cases separately: the case $\boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$ and the case $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$.

- The case $\boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$: We have

$$f''^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \overset{\text{(A)}}{=} \mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1} f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \mathrm{pref}(\boldsymbol{d}), \tag{A· 99}$$

where (A) follows from the first case of (53) and $\boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$. Comparing (A· 98) and (A· 99), we have $\mathrm{pref}(\boldsymbol{d}) \succeq \boldsymbol{c}$ since $|\mathrm{pref}(\boldsymbol{d})| \geq k \geq |\boldsymbol{c}|$. Therefore, by $\boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$, we obtain $f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \succeq \boldsymbol{d} \succeq \mathrm{pref}(\boldsymbol{d}) \succeq \boldsymbol{c}$ as desired.

- The case $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$: We have

$$f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \overset{\text{(A)}}{=} f''^{*}_{\langle\lambda\rangle}(\boldsymbol{x}) \overset{\text{(B)}}{\succeq} \boldsymbol{c}, \tag{A· 100}$$

where (A) follows from the second case of (53) and $\boldsymbol{d} \npreceq f'^{*}_{\langle\lambda\rangle}(\boldsymbol{x})$, and (B) follows from (A· 98).

(Proof of (ii)): For $i \in [F] \setminus \{\langle\lambda\rangle\}$, we have $f''_i(s) = f'_i(s)$ directly from the second case of (51). We consider the case where $i = \langle z\rangle$ for some $z \in \mathcal{S}^L$. Then we have $f'^{*}_{\langle\lambda\rangle}(z) \not\prec \boldsymbol{d}$ because $|f'^{*}_{\langle\lambda\rangle}(z)| \geq |\boldsymbol{d}| + 1$ by Lemma 17 (iii). Therefore, by the second case of (51), we obtain $f''_i(s) = f'_i(s)$.

(Proof of (iii)): We prove only that $\mathcal{P}^k_{F'',i}(\boldsymbol{b}) \subseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$ for any $i \in \mathcal{J}$ and $\boldsymbol{b} \in C^*$ because we can prove $\bar{\mathcal{P}}^k_{F'',i}(\boldsymbol{b}) \subseteq \bar{\mathcal{P}}^k_{F',i}(\boldsymbol{b})$ in the similar way. To prove $\mathcal{P}^k_{F'',i}(\boldsymbol{b}) \subseteq \mathcal{P}^k_{F',i}(\boldsymbol{b})$, it suffices to prove that for any $(i, \boldsymbol{x}, \boldsymbol{b}, \boldsymbol{c}) \in \mathcal{J} \times \mathcal{S}^+ \times C^* \times C^{\leq k}$, we have

$$\begin{aligned} &(f''^{*}_i(\boldsymbol{x}) \succeq \boldsymbol{bc}, f''_i(x_1) \succeq \boldsymbol{b}) \\ &\quad \Longrightarrow \ {}^{\exists}\boldsymbol{x}' \in \mathcal{S}^+; (f'^{*}_i(\boldsymbol{x}') \succeq \boldsymbol{bc}, f'_i(x'_1) \succeq \boldsymbol{b}) \end{aligned} \tag{A· 101}$$

because this shows that for any $i \in \mathcal{J}$, $\boldsymbol{b} \in C^*$, and $\boldsymbol{c} \in C^k$, we have

$$\begin{aligned} &\boldsymbol{c} \in \mathcal{P}^k_{F'',i}(\boldsymbol{b}) \\ &\overset{\text{(A)}}{\iff} {}^{\exists}\boldsymbol{x} \in \mathcal{S}^+; (f''^{*}_i(\boldsymbol{x}) \succeq \boldsymbol{bc}, f''_i(x_1) \succeq \boldsymbol{b}) \\ &\overset{\text{(B)}}{\Longrightarrow} {}^{\exists}\boldsymbol{x}' \in \mathcal{S}^+; (f'^{*}_i(\boldsymbol{x}') \succeq \boldsymbol{bc}, f'_i(x'_1) \succeq \boldsymbol{b}) \\ &\overset{\text{(C)}}{\iff} \boldsymbol{c} \in \mathcal{P}^k_{F',i}(\boldsymbol{b}) \end{aligned}$$

as desired, where (A) follows from (4), (B) follows from (A· 101), and (C) follows from (4).

Choose $(i, \boldsymbol{x}, \boldsymbol{b}, \boldsymbol{c}) \in [F] \times \mathcal{S}^+ \times C^* \times C^{\leq k}$ arbitrarily and assume

$$f''^{*}_i(\boldsymbol{x}) \succeq \boldsymbol{bc} \tag{A· 102}$$

and

$$f''_i(x_1) \succeq \boldsymbol{b}. \tag{A· 103}$$

Then we have

$$f'_i(x_1) \overset{\text{(A)}}{=} f''_i(x_1) \overset{\text{(B)}}{\succeq} \boldsymbol{b}, \tag{A· 104}$$

where (A) follows from (ii) of this lemma, and (B) follows from (A· 103).

We prove (A· 101) by induction for $|\boldsymbol{x}|$. For the base case $|\boldsymbol{x}| = 1$, we have

$$f'^{*}_i(\boldsymbol{x}) = f'_i(x_1) \overset{\text{(A)}}{=} f''_i(x_1) = f''^{*}_i(\boldsymbol{x}) \overset{\text{(B)}}{\succeq} \boldsymbol{bc} \tag{A· 105}$$

as desired, where (A) follows from (ii) of this lemma, and (B) follows from (A· 102). By (A· 105) and (A· 104), the claim (A· 101) holds for the base case $|\boldsymbol{x}| = 1$.

We consider the induction step for $|\boldsymbol{x}| \geq 2$. We have

$$f_i'(x_1)f''^*_{\tau_i''(x_1)}(\mathrm{suff}(\boldsymbol{x})) \overset{\text{(A)}}{=} f_i''(x_1)f''^*_{\tau_i''(x_1)}(\mathrm{suff}(\boldsymbol{x})) \overset{\text{(B)}}{=} f_i''^*(\boldsymbol{x}) \overset{\text{(C)}}{\succeq} \boldsymbol{bc}, \quad \text{(A· 106)}$$

where (A) follows from (ii) of this lemma, (B) follows from (2), and (C) follows from (A· 102).

Therefore, $f_i'(x_1) \succeq \boldsymbol{bc}$ or $f_i'(x_1) \prec \boldsymbol{bc}$ holds. In the case $f_i'(x_1) \succeq \boldsymbol{bc}$, the sequence $\boldsymbol{x}' := x_1$ satisfies $f_i'^*(\boldsymbol{x}') \succeq \boldsymbol{bc}$ and $f_i'(x_1') = f_i'(x_1) \succeq \boldsymbol{b}$ by (A· 104) as desired. Thus, now we assume $f_i'(x_1) \prec \boldsymbol{bc}$. Then we have

$$\begin{aligned} |f_i'(x_1)^{-1}\boldsymbol{bc}| &= -|f_i'(x_1)| + |\boldsymbol{b}| + |\boldsymbol{c}| \\ &\overset{\text{(A)}}{=} -|f_i''(x_1)| + |\boldsymbol{b}| + |\boldsymbol{c}| \\ &\overset{\text{(B)}}{\leq} |\boldsymbol{c}| \\ &\leq k, \end{aligned} \quad \text{(A· 107)}$$

where (A) follows from (ii) of this lemma, and (B) follows from (A· 103). By (A· 106), we have

$$f''^*_{\tau_i''(x_1)}(\mathrm{suff}(\boldsymbol{x})) \succeq f_i'(x_1)^{-1}\boldsymbol{bc}. \quad \text{(A· 108)}$$

We can see that there exists $\boldsymbol{x}' \in \mathcal{S}^+$ such that

$$f'^*_{\tau_i''(x_1)}(\boldsymbol{x}') \succeq f_i'(x_1)^{-1}\boldsymbol{bc} \quad \text{(A· 109)}$$

as follows.

- The case $\tau_i''(x_1) = \langle\lambda\rangle$: By (A· 107), we can apply (i) of this lemma to obtain that $\boldsymbol{x}' := \mathrm{suff}(\boldsymbol{x})$ satisfies (A· 109) from (A· 108).
- The case $\tau_i''(x_1) \in \mathcal{J}$: By (A· 107) and (A· 108), we can apply the induction hypothesis to $(\tau_i''(x_1), \mathrm{suff}(\boldsymbol{x}), \lambda, f_i'(x_1)^{-1}\boldsymbol{bc})$.

Therefore, we have

$$\begin{aligned} f_i'^*(x_1\boldsymbol{x}') &\overset{\text{(A)}}{=} f_i'(x_1)f'^*_{\tau_i'(x_1)}(\boldsymbol{x}') \\ &\overset{\text{(B)}}{=} f_i'(x_1)f'^*_{\tau_i''(x_1)}(\boldsymbol{x}') \\ &\overset{\text{(C)}}{\succeq} f_i'(x_1)f_i'(x_1)^{-1}\boldsymbol{bc} \\ &= \boldsymbol{bc}, \end{aligned} \quad \text{(A· 110)}$$

where (A) follows from (2), (B) follows from (52), and (C) follows from (A· 109). The induction is completed by (A· 104) and (A· 110). □

C.12 Proof of Lemma 19

Proof of Lemma 19. (Proof of (i)): Assume that

$$\boldsymbol{b} \preceq \boldsymbol{b}'. \quad \text{(A· 111)}$$

In the case $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \not\preceq \mathrm{pref}(\boldsymbol{d})$, we have

$$\psi_{\boldsymbol{z}}(\boldsymbol{b}) \overset{\text{(A)}}{=} \boldsymbol{b} \overset{\text{(B)}}{\preceq} \boldsymbol{b}' \overset{\text{(C)}}{=} \psi_{\boldsymbol{z}}(\boldsymbol{b}'), \quad \text{(A· 112)}$$

where (A) follows from the second case of (60) and $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \not\preceq \mathrm{pref}(\boldsymbol{d})$, (B) follows from (A· 111), and (C) follows from the second case of (60) and $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \not\preceq \mathrm{pref}(\boldsymbol{d})$.

We consider the case $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \preceq \mathrm{pref}(\boldsymbol{d})$ dividing into four cases by whether $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$ and whether $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$.

- The case $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}, \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$: We have
$$\begin{aligned} \psi_{\boldsymbol{z}}(\boldsymbol{b}) &\overset{\text{(A)}}{=} f'^*_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}) \\ &\overset{\text{(B)}}{\preceq} f'^*_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}') \overset{\text{(C)}}{=} \psi_{\boldsymbol{z}}(\boldsymbol{b}') \end{aligned}$$
as desired, where (A) follows from the first case of (60) and $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$, (B) follows from (A· 111), and (C) follows from the first case of (60) and $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$.
- The case $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}, \mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$: This case is impossible because (A· 111) leads to $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b} \preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$, which conflicts with $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$.
- The case $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}, \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$: By (A· 111), we have
$$f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b} \preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'. \quad \text{(A· 113)}$$
By (A· 113) and $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$, exactly one of $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$ and $\mathrm{pref}(\boldsymbol{d}) \succeq f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$ holds. Since the former does not hold by $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$, the latter holds:
$$f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b} \preceq \mathrm{pref}(\boldsymbol{d}). \quad \text{(A· 114)}$$
Thus, we have
$$\begin{aligned} \psi_{\boldsymbol{z}}(\boldsymbol{b}) &\overset{\text{(A)}}{=} \boldsymbol{b} = f'^*_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b} \overset{\text{(B)}}{\preceq} f'^*_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\mathrm{pref}(\boldsymbol{d}) \\ &\preceq f'^*_{\langle\lambda\rangle}(\boldsymbol{z})^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}') \overset{\text{(C)}}{=} \psi_{\boldsymbol{z}}(\boldsymbol{b}'), \end{aligned}$$
where (A) follows from the second case of (60) and $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$, (B) follows from (A· 114), and (C) follows from the first case of (60) and $\mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$.
- The case $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}, \mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$: We have
$$\psi_{\boldsymbol{z}}(\boldsymbol{b}) \overset{\text{(A)}}{=} \boldsymbol{b} \overset{\text{(B)}}{\preceq} \boldsymbol{b}' \overset{\text{(C)}}{=} \psi_{\boldsymbol{z}}(\boldsymbol{b}') \quad \text{(A· 115)}$$
as desired, where (A) follows from the second case of (60) and $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}$, (B) follows from (A· 111), and (C) follows from the second case of (60) and $\mathrm{pref}(\boldsymbol{d}) \not\prec f'^*_{\langle\lambda\rangle}(\boldsymbol{z})\boldsymbol{b}'$.

(Proof of (ii)): We consider the following three cases separately: (I) the case $f'^*_{\langle\lambda\rangle}(\boldsymbol{z}) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{zx})$, (II) the case $f'^*_{\langle\lambda\rangle}(\boldsymbol{zx}) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(\boldsymbol{zx})\boldsymbol{c}$, and (III) the other case:

(I) The case $f'^{*}_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})$: We have

$$f'^{*}_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) \tag{A· 116}$$

since

$$\mathrm{pref}(\boldsymbol{d})\bar{d}_l \overset{(A)}{\notin} \mathcal{P}^{*}_{F,\langle\lambda\rangle} \overset{(B)}{=} \mathcal{P}^{*}_{F',\langle\lambda\rangle}, \tag{A· 117}$$

where (A) follows from (47), and (B) follows from Lemma 7 (ii) since φ defined in (50) is a homomorphism from F' to F. Therefore, by the second case of (60), we obtain

$$f''^{*}_{\langle z\rangle}(\boldsymbol{x}) = f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})). \tag{A· 118}$$

We consider the following two cases separately: (I-A) the case $f'^{*}_{\langle\lambda\rangle}(z) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}, \boldsymbol{c} = \lambda$ and (I-B) the other case.

(I-A) The case $f'^{*}_{\langle\lambda\rangle}(z) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}, \boldsymbol{c} = \lambda$: We have

$$\begin{aligned} & f'^{*}_{\langle\lambda\rangle}(z)f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c} \\ & \overset{(A)}{=} f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}))\boldsymbol{c} \\ & \overset{(B)}{=} f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}\boldsymbol{d}\boldsymbol{c} \\ & \overset{(C)}{=} \mathrm{pref}(\boldsymbol{d}) \\ & \not\succ \mathrm{pref}(\boldsymbol{d}), \end{aligned} \tag{A· 119}$$

where (A) follows from (A· 118), (B) follows from $f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}$, and (C) follows from $\boldsymbol{c} = \lambda$.
Hence, we have

$$\begin{aligned} & \psi_{z}(f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(A)}{=} f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c} \\ & \overset{(B)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c} \\ & \overset{(C)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}))\boldsymbol{d}^{-1}\boldsymbol{d} \\ & \overset{(D)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}f'^{*}_{\langle\lambda\rangle}(z)\mathrm{pref}(f'^{*}_{\langle z\rangle}(\boldsymbol{x})))\boldsymbol{d}^{-1}\boldsymbol{d} \\ & = \mathrm{pref}(f'^{*}_{\langle z\rangle}(\boldsymbol{x})) \end{aligned}$$

as desired, where (A) follows from the second case of (60) and (A· 119), (B) follows from (A· 118), (C) follows from $f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}$ and $\boldsymbol{c} = \lambda$, and (D) follows from Lemma 1 (i), Lemma 16 (i), and $f'^{*}_{\langle\lambda\rangle}(z) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})$.

(I-B) The other case: Then by (A· 116), we have

$$\boldsymbol{d} \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}, \tag{A· 120}$$

since it does not hold that $f'^{*}_{\langle\lambda\rangle}(z) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) = \boldsymbol{d}, \boldsymbol{c} = \lambda$ by the assumption of the case (I-B).
We have

$$\begin{aligned} & f'^{*}_{\langle\lambda\rangle}(z)f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c} \\ & \overset{(A)}{=} f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}))\boldsymbol{c} \\ & \overset{(B)}{\succ} f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle\lambda\rangle}(z)^{-1}\mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}\boldsymbol{d} \\ & = \mathrm{pref}(\boldsymbol{d}) \end{aligned} \tag{A· 121}$$

$$\overset{(C)}{\succeq} f'^{*}_{\langle\lambda\rangle}(z) \tag{A· 122}$$

as desired, where (A) follows from (A· 118), (B) follows from (A· 120), and (C) follows from the assumption of the case (I).
Hence, we have

$$\begin{aligned} & \psi_{z}(f''_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(A)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z)f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(B)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle\lambda\rangle}(z)^{-1} \\ & \quad \mathrm{pref}(\boldsymbol{d})\boldsymbol{d}^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c})) \\ & = f'^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c} \\ & \overset{(C)}{=} f'^{*}_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(\boldsymbol{c}), \end{aligned}$$

where (A) follows from the first case of (60), (A· 121), and (A· 122), (B) follows from (A· 118), and (C) follows from the second case of (60) and the assumption of the case (I).

(II) The case $f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}$: Then since $\boldsymbol{d} \not\preceq f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})$, we have

$$f''^{*}_{\langle z\rangle}(\boldsymbol{x}) = f'^{*}_{\langle z\rangle}(\boldsymbol{x}) \tag{A· 123}$$

applying the second case of (53). Therefore, we have

$$\begin{aligned} & f'^{*}_{\langle\lambda\rangle}(z) \preceq f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x}) \\ & \qquad \overset{(A)}{\preceq} \mathrm{pref}(\boldsymbol{d}) \end{aligned} \tag{A· 124}$$

$$\begin{aligned} & \qquad \overset{(A)}{\prec} f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c} \\ & \qquad \overset{(B)}{=} f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c} \\ & \qquad \overset{(C)}{=} f'^{*}_{\langle\lambda\rangle}(z)f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}, \end{aligned} \tag{A· 125}$$

where (A)s follow from the assumption of the case (II), (B) follows from Lemma 1 (i) and Lemma 16 (i), and (C) follows from (A· 123).
Hence, we have

$$\begin{aligned} & \psi_{z}(f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(A)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z)f''^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(B)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z)f'^{*}_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) \\ & \overset{(C)}{=} f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}) \\ & = f'^{*}_{\langle z\rangle}(\boldsymbol{x})f'^{*}_{\langle z\rangle}(\boldsymbol{x})^{-1}f'^{*}_{\langle\lambda\rangle}(z)^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^{*}_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}) \end{aligned}$$

$$\stackrel{\text{(D)}}{=} f'^*_{\langle z\rangle}(\boldsymbol{x}) f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})^{-1}\boldsymbol{d}\mathrm{pref}(\boldsymbol{d})^{-1}(f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c})$$
$$\stackrel{\text{(E)}}{=} f'^*_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(\boldsymbol{c})$$

as desired, where (A) follows from the first case of (60), (A· 124), and (A· 125), (B) follows from (A· 123), (C) follows from Lemma 1 (i) and Lemma 16 (i), (D) follows from Lemma 1 (i) and Lemma 16 (i), and (E) follows from the first case of (60) and the assumption of the case (II).

(III) The other case: The following implication holds:

$$f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}) \implies f'^*_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}). \quad \text{(A· 126)}$$

Now, it does not hold that $f'^*_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$ by the assumption of the case (III). Hence, by the contraposition of (A· 126), we see that $f'^*_{\langle\lambda\rangle}(z) \prec \boldsymbol{d} \preceq f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$ does not hold. Therefore, we obtain

$$f''^*_{\langle z\rangle}(\boldsymbol{x}) = f'^*_{\langle z\rangle}(\boldsymbol{x}) \quad \text{(A· 127)}$$

applying the second case of (53).

By the assumption of the case (III), neither $f'^*_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})$ nor $f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}$ hold. Hence, the following condition does not hold:

$$f'^*_{\langle\lambda\rangle}(z) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c} \stackrel{\text{(A)}}{=} f'^*_{\langle\lambda\rangle}(z) f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}, \quad \text{(A· 128)}$$

where (A) follows from (A· 127). Therefore, by the second case of (60), we have

$$\psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}) = f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}. \quad \text{(A· 129)}$$

Thus, we have

$$f'^*_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(\boldsymbol{c}) \stackrel{\text{(A)}}{=} f''^*_{\langle z\rangle}(\boldsymbol{x})\psi_{z\boldsymbol{x}}(\boldsymbol{c})$$
$$\stackrel{\text{(B)}}{=} f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c}$$
$$\stackrel{\text{(C)}}{=} \psi_z(f''^*_{\langle z\rangle}(\boldsymbol{x})\boldsymbol{c})$$

as desired, where (A) follows from (A· 127), (B) follows from the second case of (60) since $f'^*_{\langle\lambda\rangle}(z\boldsymbol{x}) \preceq \mathrm{pref}(\boldsymbol{d}) \prec f'^*_{\langle\lambda\rangle}(z\boldsymbol{x})\boldsymbol{c}$ does not hold by the assumption of the case (III), and (C) follows from (A· 129).

(Proof of (iii)): We have $f'^*_{\langle\lambda\rangle}(z) \not\preceq \mathrm{pref}(\boldsymbol{d})$ because $|f'^*_{\langle\lambda\rangle}(z)| > |\boldsymbol{d}|$ by Lemma 17 (iii). Hence, by the second case of (60), we obtain $\psi_{\langle z\rangle}(\boldsymbol{b}) = \boldsymbol{b}$ as desired. □

Appendix D: List of Notations

$\mathcal{A}\times\mathcal{B}$	the Cartesian product of sets $\mathcal{A}$ and $\mathcal{B}$, that is, $\{(a,b) : a \in \mathcal{A}, b \in \mathcal{B}\}$, defined at the beginning of Sect. 2.
$\lvert\mathcal{A}\rvert$	the cardinality of a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{A}^k$	the set of all sequences of length k over a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{A}^{\geq k}$	the set of all sequences of length greater than or equal to k over a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{A}^{\leq k}$	the set of all sequences of length less than or equal to k over a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{A}^*$	the set of all sequences of finite length over a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{A}^+$	the set of all sequences of finite positive length over a set $\mathcal{A}$, defined at the beginning of Sect. 2.
$\mathcal{C}$	the coding alphabet $\mathcal{C} = \{0,1\}$, at the beginning of Sect. 2.
$\bar{c}$	the negation of $c \in \mathcal{C}$, that is, $\bar{0} = 1, \bar{1} = 0$ defined at the beginning of the proof of Theorem 2.
f_i^*	defined in Definition 2.
F	simplified notation of a code-tuple $F(f_0, f_1, \ldots, f_{m-1}, \tau_0, \tau_1, \ldots, \tau_{m-1})$, also written as $F(f,\tau)$, defined below Definition 1.
$\bar{F}$	an irreducible part of F, defined in Definition 14.
$\lvert F\rvert$	the number of code tables of F, defined below Definition 1.
$[F]$	simplified notation of $[\lvert F\rvert] = \{0,1,2,\ldots,\lvert F\rvert-1\}$, defined below Definition 1.
$\mathscr{F}^{(m)}$	the set of all m-code-tuples, defined after Definition 1.
$\mathscr{F}$	the set of all code-tuples, defined after Definition 1.
$\mathscr{F}_{\mathrm{ext}}$	the set of all extendable code-tuples, defined in Definition 6.
$\mathscr{F}_{k\text{-opt}}$	defined in Definition 16.
$\mathscr{F}_{\mathrm{reg}}$	the set of all regular code-tuples, defined in Definition 9.
$h(F)$	defined after Lemma 14.
$L(F)$	the average codeword length of a code-tuple F, defined in Definition 10.
$L_i(F)$	the average codeword length of the i-th code table of F, defined in Definition 10.
$[m]$	$\{0,1,2,\ldots,m-1\}$, defined at the beginning of Sect. 2.
$\mathcal{P}^k_{F,i}$	defined in Definition 3.
$\bar{\mathcal{P}}^k_{F,i}$	defined in Definition 3.
$\mathcal{P}^*_{F,i}$	defined in Definition 4.
$\bar{\mathcal{P}}^*_{F,i}$	defined in Definition 4.
$\mathscr{P}^k_F$	defined in Definition 15.
$\mathrm{pref}(\boldsymbol{x})$	the sequence obtained by deleting the last letter of $\boldsymbol{x}$, defined at the beginning of Sect. 2.
$Q(F)$	the transition probability matrix, defined in Definition 7.
$Q_{i,j}(F)$	the transition probability, defined in Definition 7.
$\mathbb{R}$	the set of all real numbers.
$\mathbb{R}^m$	the set of all m-dimensional real row vectors for an integer $m \geq 1$.
$\mathcal{S}$	the source alphabet, defined at the beginning of Sect. 2.

$\mathrm{suff}(\boldsymbol{x})$	the sequence obtained by deleting the first letter of $\boldsymbol{x}$, defined at the beginning of Sect. 2.
$\boldsymbol{x} \preceq \boldsymbol{y}$	$\mathbf{x}$ is a prefix of $\mathbf{y}$, defined at the beginning of Sect. 2.
$\boldsymbol{x} \prec \boldsymbol{y}$	$\boldsymbol{x} \preceq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$, defined at the beginning of Sect. 2.
$\lvert\boldsymbol{x}\rvert$	the length of a sequence $\boldsymbol{x}$, defined at the beginning of Sect. 2.
$\boldsymbol{x}^{-1}\boldsymbol{y}$	the sequence $\boldsymbol{z}$ such that $\boldsymbol{x}\boldsymbol{z} = \boldsymbol{y}$ defined at the beginning of Sect. 2.
λ	the empty sequence, defined at the beginning of Sect. 2.
$\mu(s)$	the probability of occurrence of symbol s, defined at the beginning of Sect. 2.4.
$\boldsymbol{\pi}(F)$	defined in Definition 9.
σ	the alphabet size, defined at the beginning of Sect. 2.
τ_i^*	defined in Definition 2.

Kengo Hashimoto was born in Japan, in 1994. He received the B.E. and M.E. degrees from the Department of Information Science, University of Fukui, in 2019 and 2021, respectively. He is currently pursuing his Ph.D. degree in University of Fukui. His research interests include information theory and coding theory.

Ken-ichi Iwata received his B.Ed. degree from Wakayama University in 1993, the M.Sc. degree from Information Science from Japan Advanced Institute of Science and Technology in 1995, and the D.E. degree from the University of Electro-Communications in 2006. Since 2008, he has been with the University of Fukui, as an Associate Professor in the Graduate School of Engineering.