Tight Upper Bounds on the Redundancy of Optimal Binary AIFV codes

11 July 2016

Weihua Hu, Hirosuke Yamamoto, Junya Honda
The University of Tokyo

Outline

1. Background

Binary Almost instantaneous FV (AIFV) codes

2. Main result:

Tight upper bounds on the redundancy of binary AIFV codes (worst-case redundancy)

3. Comparison with Huffman codes

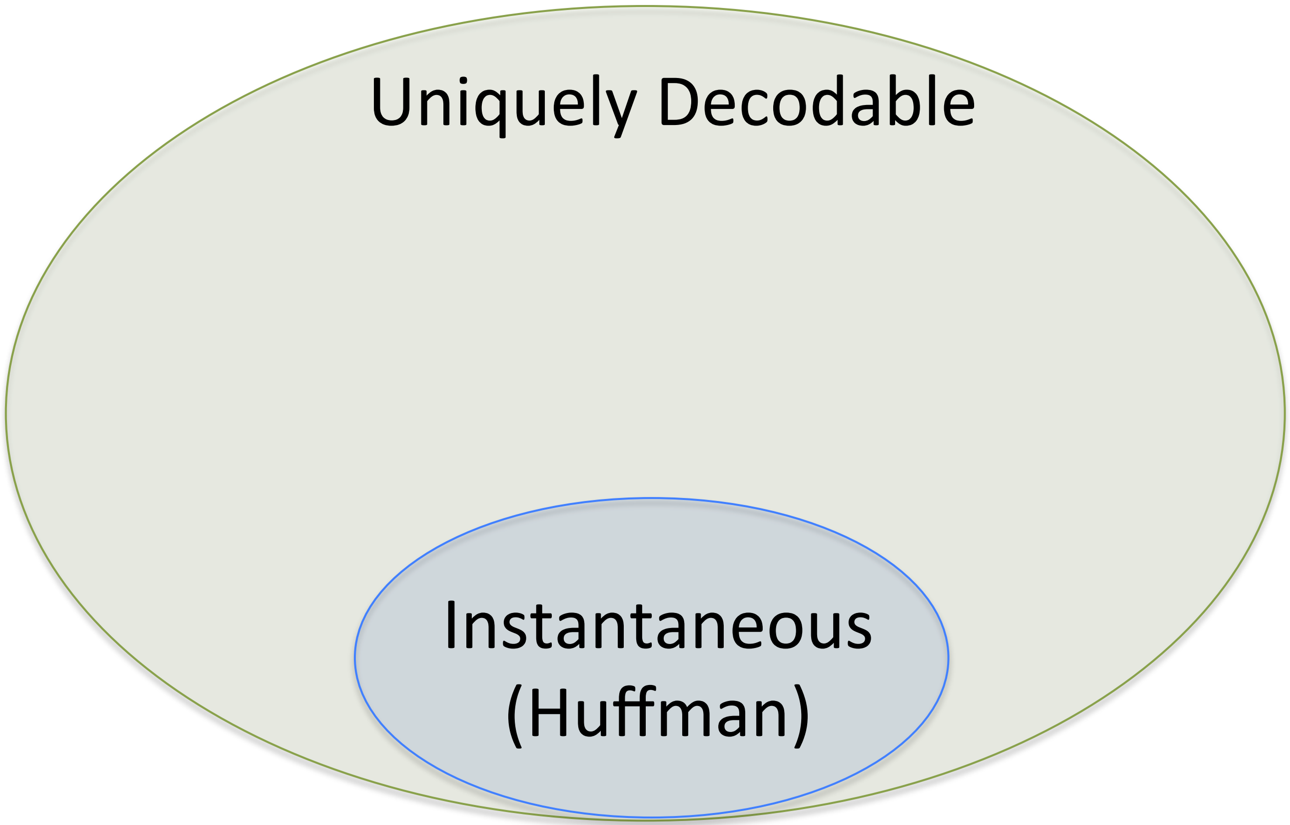
4. Idea & outline of proofs

5. Conclusion

Binary AIFV codes

Fixed-to-**Variable** length (FV) codes

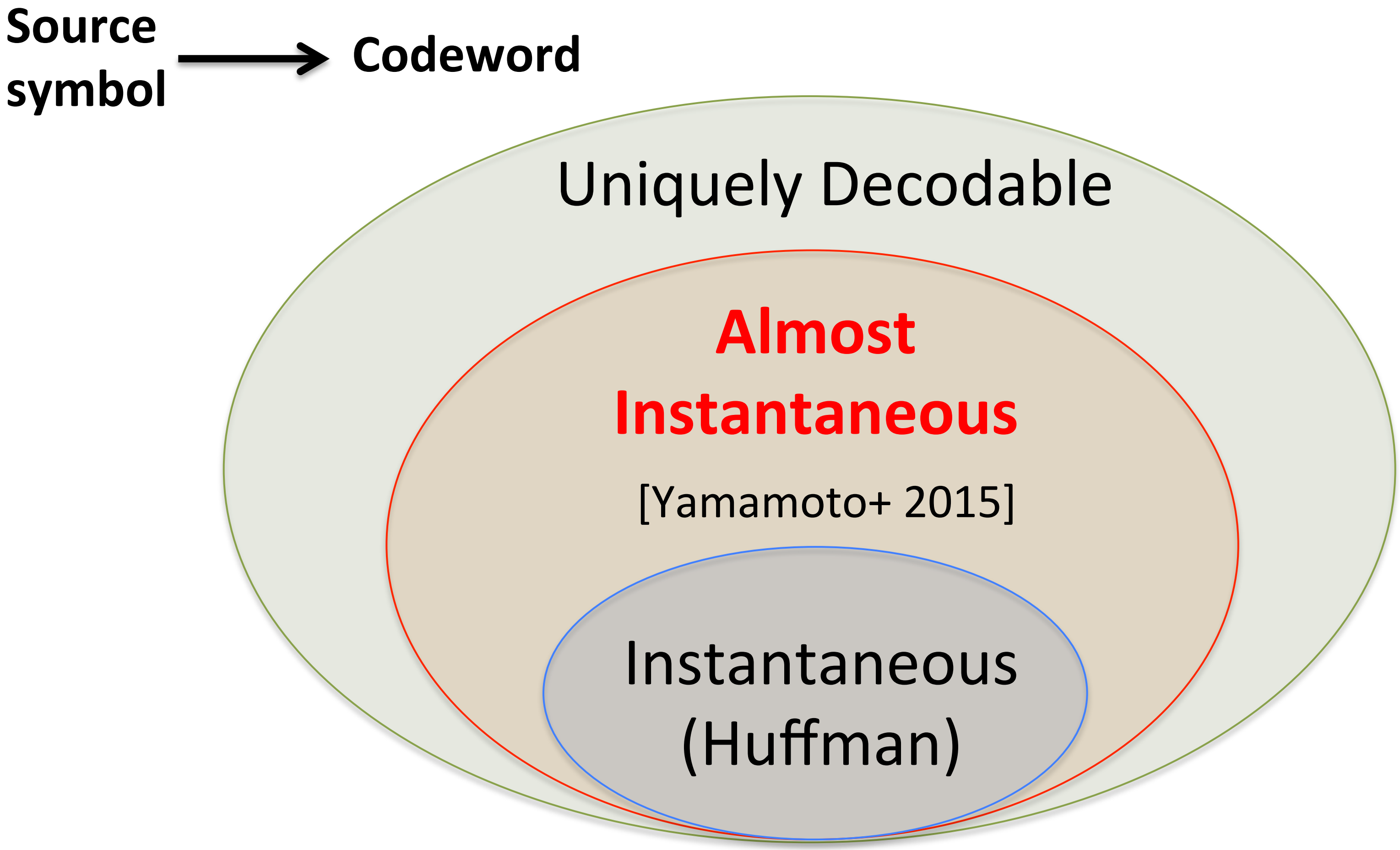
Source symbol → **Codeword**



Class of FV Codes

Binary AIFV codes

Fixed-to-**Variable** length (FV) codes



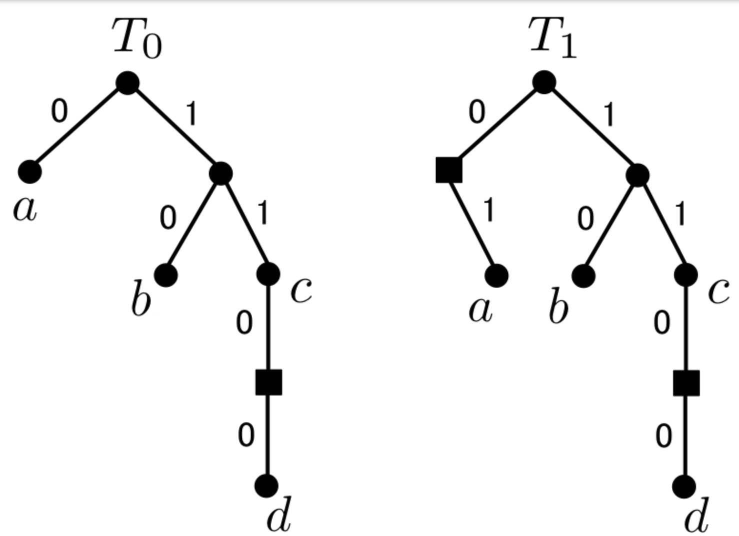
Class of FV Codes

Almost Instantaneous (AI) FV codes

Generalization of instantaneous binary FV codes

[Yamamoto, Tsuchihashi, Honda, 2015]

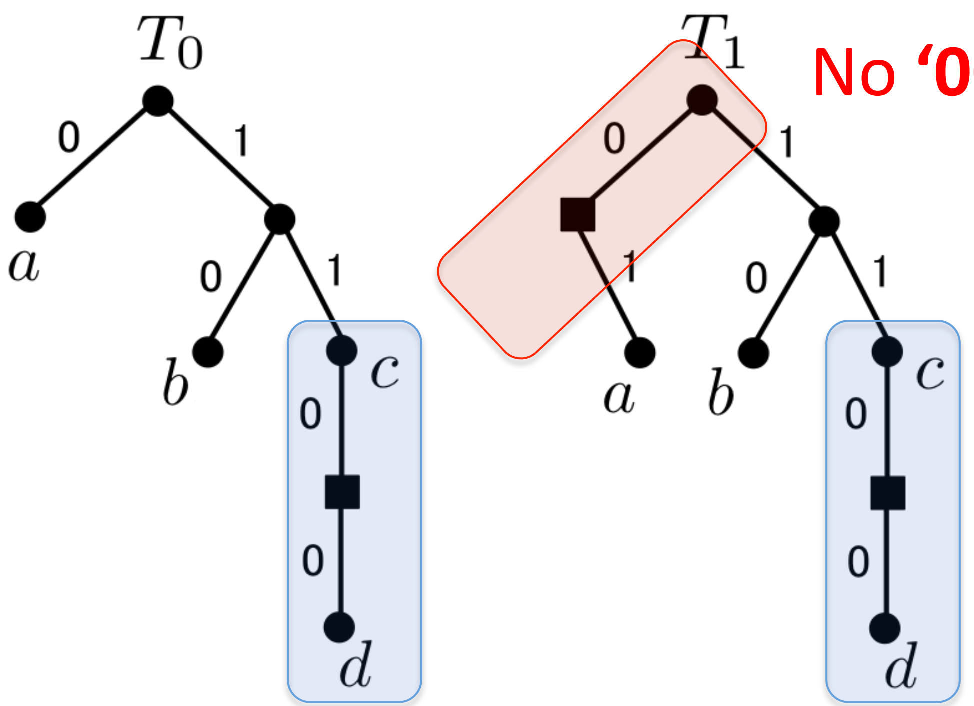
	Instantaneous	Almost Instantaneous
No.(Code Trees)	One	**Two**
Source Symbols	Leaves	**Leaves + incomplete node (master node)**
Decoding Delay	None	**At most 2 bits.**



Example of binary AIFV code trees.

Almost Instantaneous (AI) FV codes

Encoding (decoding) procedure use T_0 and T_1 iteratively.
After using a **master node**, use **T₁** for the next.



No **'00'** from the root.

Only **'00'** from the **master node.**

→The codes are **uniquely decodable.**

Worst-case Redundancy of AIFV codes

	Huffman code $\subset$	AIFV code
Redundancy	< 1	**< 1** [Yamamoto+ 2015]

AIFV codes have **good empirical performance**.
Even beat Huffman code for $\mathcal{X}^2$ for some sources.
[Yamamoto+ 2015]

Worst-case Redundancy of AIFV codes

	Huffman code $\subset$	AIFV code
Redundancy	< 1	**< 1** [Yamamoto+ 2015] **< 1/2 (Our result)**

$$p_{\max} \equiv \max_{x \in \mathcal{X}} p_X(x).$$

Worst-case redundancy in terms of $p_{\max}$ **(Our result)**

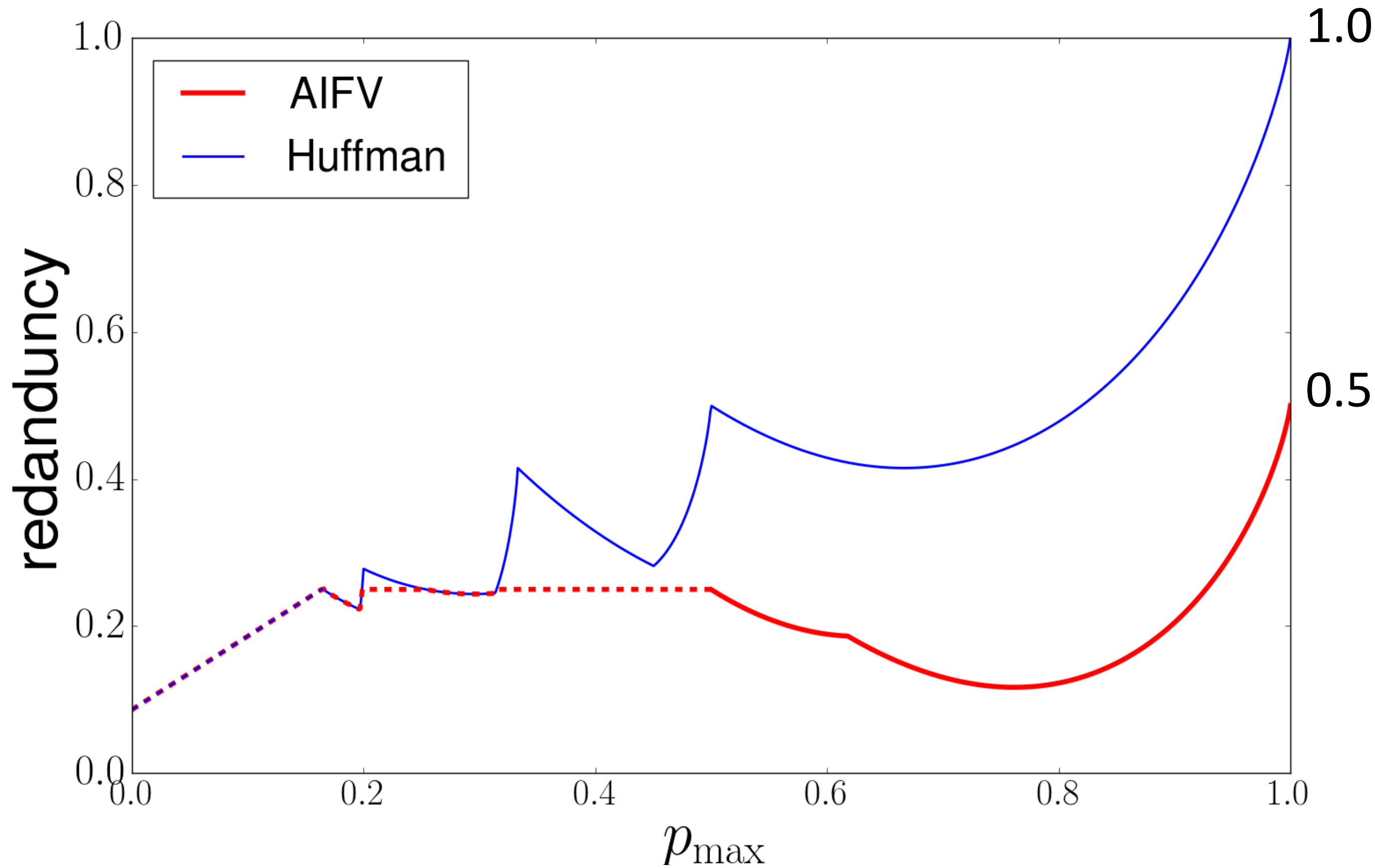
Worst-case Redundancy of AIFV codes

Theorem (Worst-case redundancy of AIFV codes)
For $p_{\max} = p \geq 1/2$, the worst-case redundancy of AIFV codes is

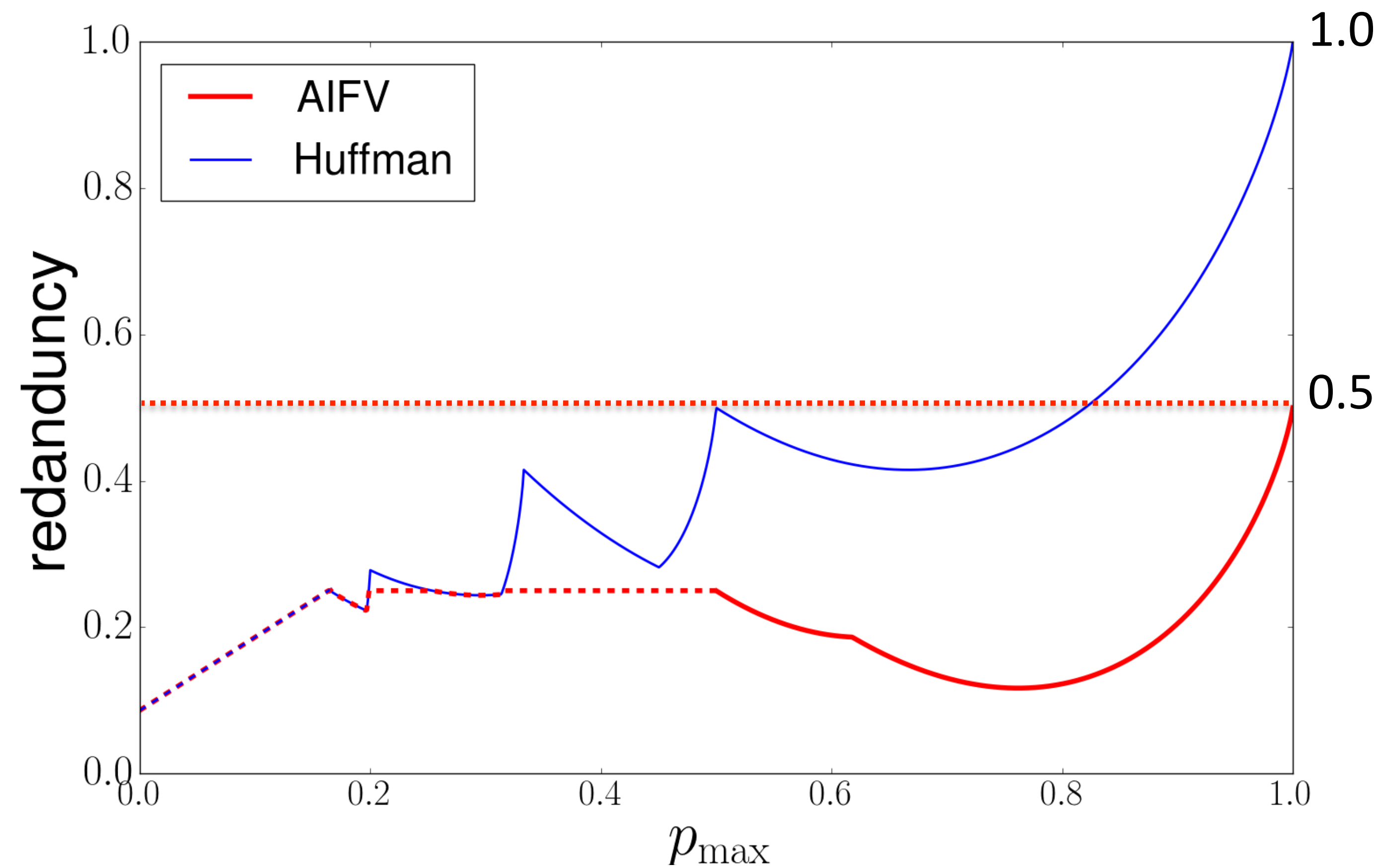
$$f(p) = \begin{cases} p^2 - 2p + 2 - h(p) & \text{if } \frac{1}{2} \leq p \leq \frac{-1+\sqrt{5}}{2}, \\ \frac{-2p^2+p+2}{1+p} - h(p) & \text{if } \frac{-1+\sqrt{5}}{2} \leq p < 1. \end{cases}$$

Theorem (Redundancy upper bound of AIFV codes)
For $p_{\max} < 1/2$, the worst-case redundancy is at most $\mathbf{1/4}$.

Comparison with Huffman codes



Comparison with Huffman codes



Corollary (Worst-case Redundancy)
Worst-case redundancy of binary AIFV codes is $\mathbf{1/2}$.

Comparison with Huffman codes

	Huffman	**AIFV**	Huffman for $\mathcal{X}^2$
Redundancy	< 1	**< 1/2**	< 1/2
Storage for Code trees	$\mathcal{O}(\|\mathcal{X}\|)$	$\mathcal{O}(\|\mathcal{X}\|)$	$\mathcal{O}(\|\mathcal{X}\|^2)$

$\mathcal{X}$: Source alphabet

More memory efficient than Huffman codes for $\mathcal{X}^2$.

Proof idea

Goal:

Prove bounds of optimal binary AIFV codes

Challenge:

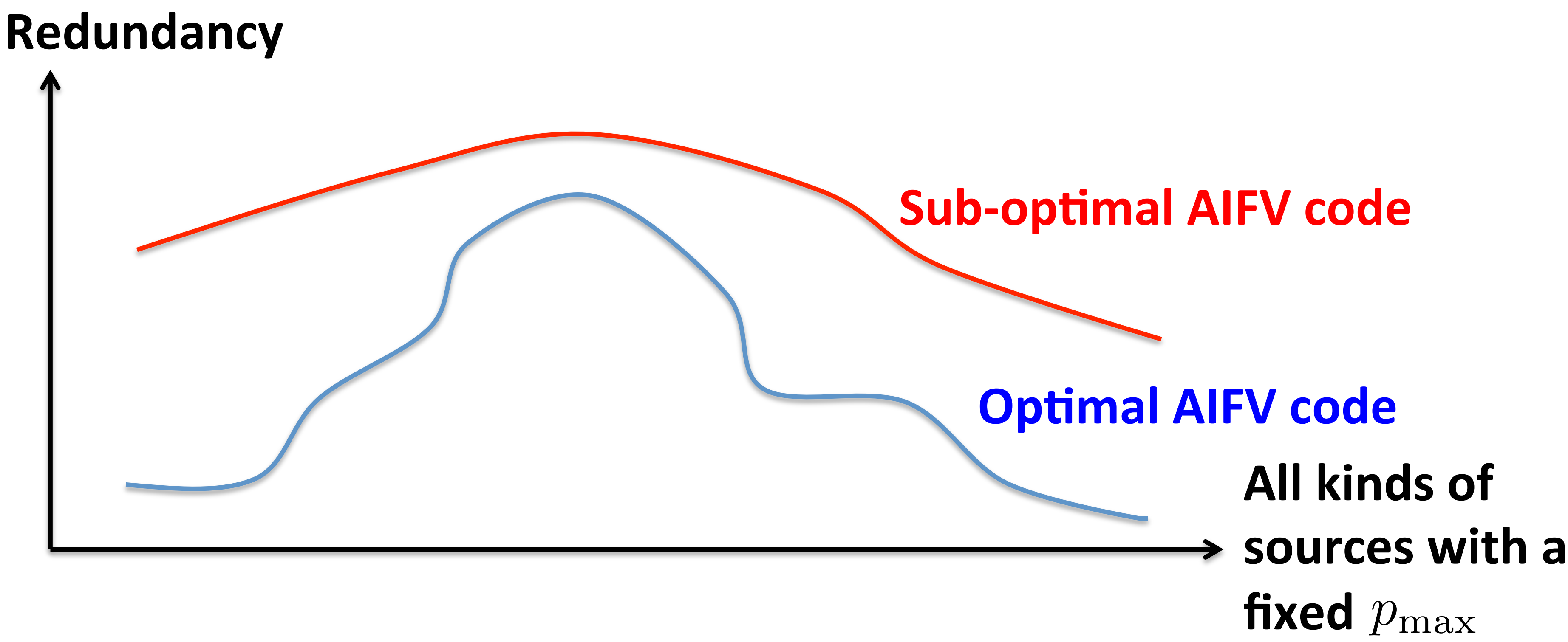
No simple algorithm known to construct the optimal AIFV code.

→ **Difficult to analyze** optimal code directly…

Proof idea

Our approach:

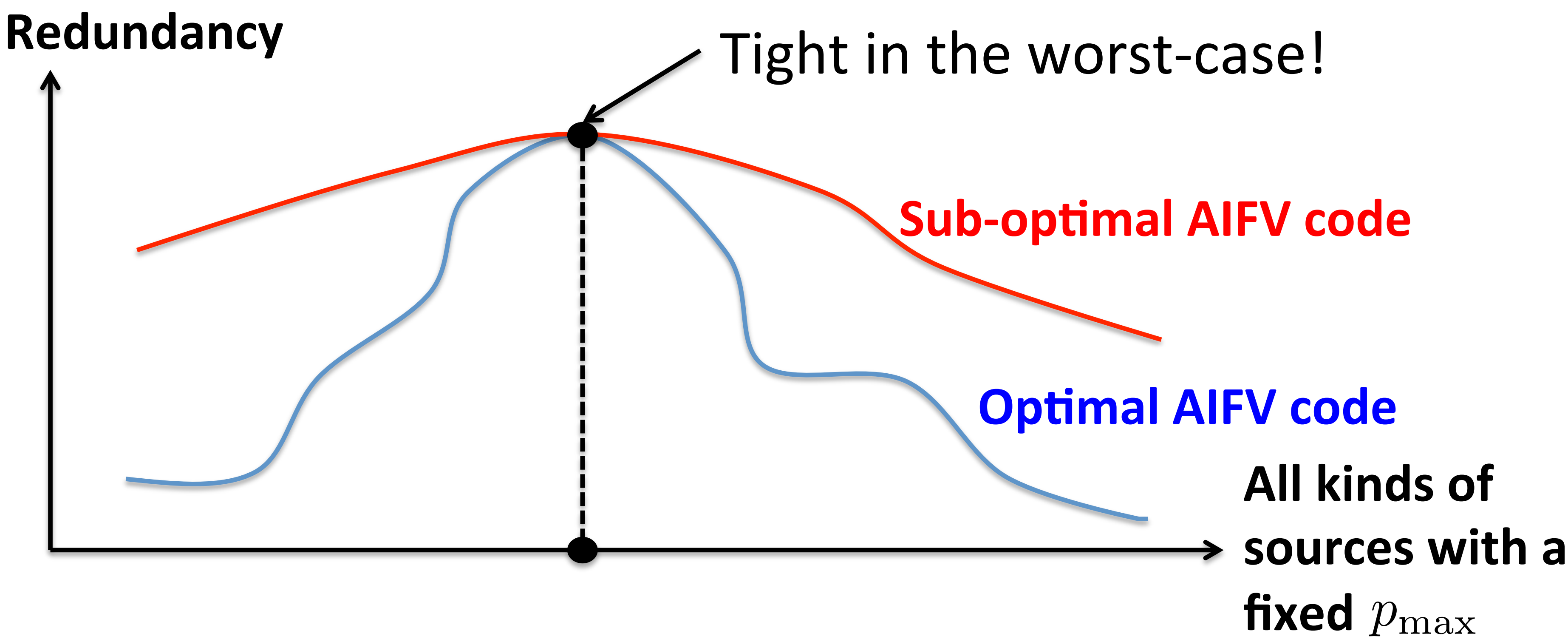
Simple Construction of **sub-optimal AIFV** codes from Huffman codes.



Proof idea

Our approach:

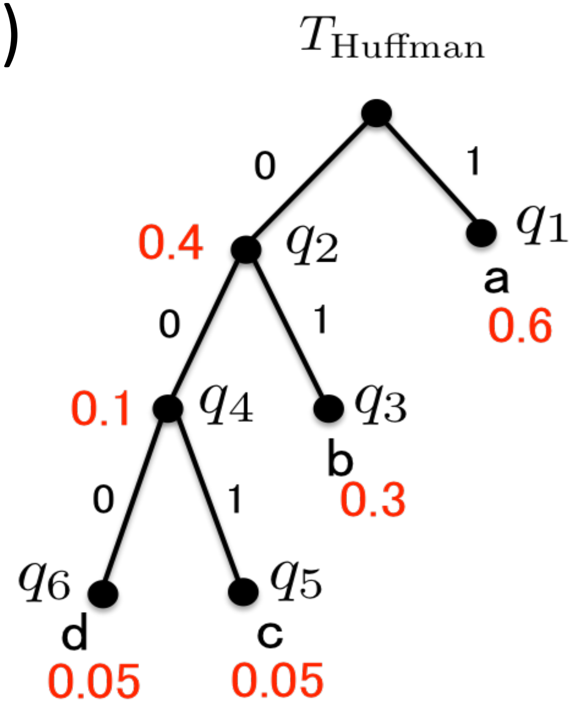
Simple Construction of **sub-optimal AIFV** codes from Huffman codes.



Proof outline (1/6)

- Simple **two-stage construction** of sub-optimal AIFV code trees from Huffman tree

Ex.)



Sibling property

[Gallager 1978]

$$q_1 \geq q_2, \cdots \geq q_{2K-2}$$

K : size of source alphabet

Sibling pair: $(q_{2k-1},\ q_{2k})$

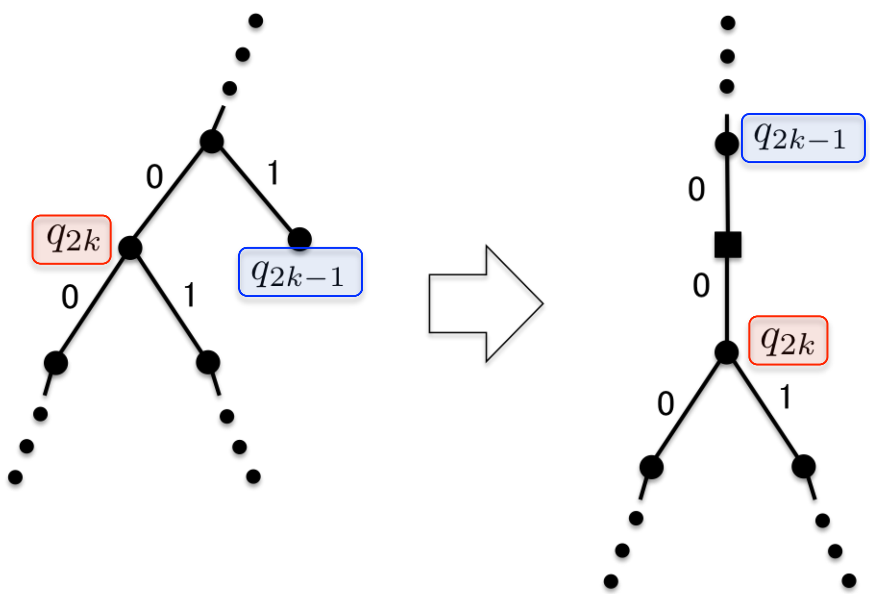
Proof outline (2/6)

- **Two-stage construction**

1. From T_{Huffman} to T_{base}

for sibling index $k = 2, \ldots, K-1$ **do**

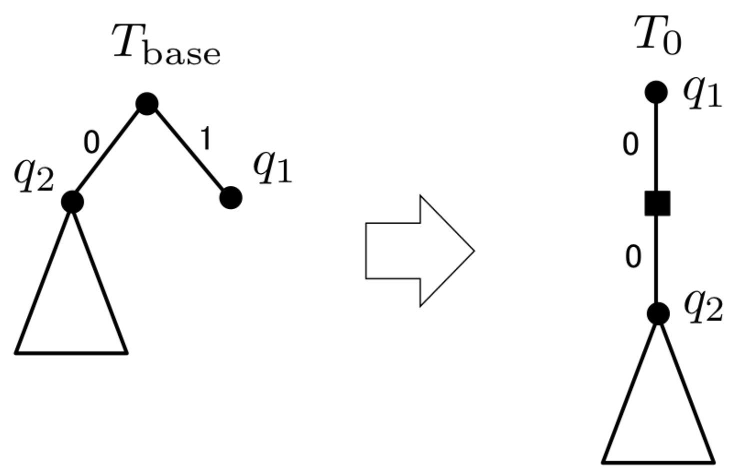
if $\boxed{q_{2k-1}}$ is a leaf and $\boxed{2q_{2k}} < \boxed{q_{2k-1}}$ **then**



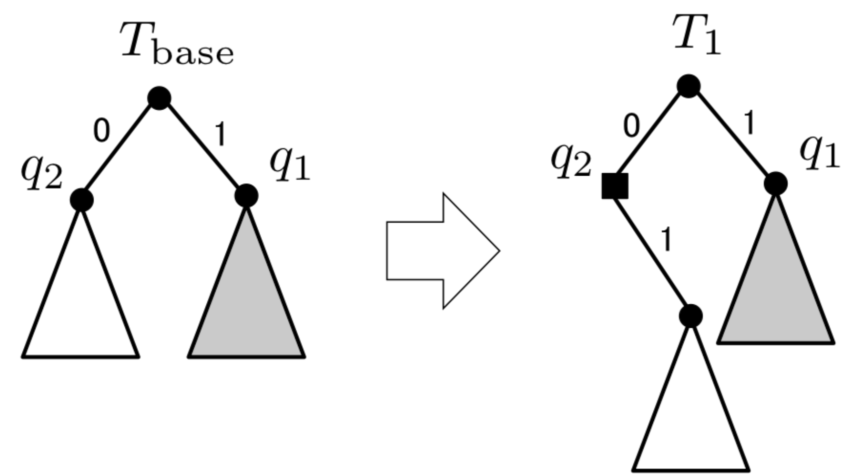
Proof outline (3/6)

- **Two-stage** construction

2. From T_{base} to T_0 and T_1



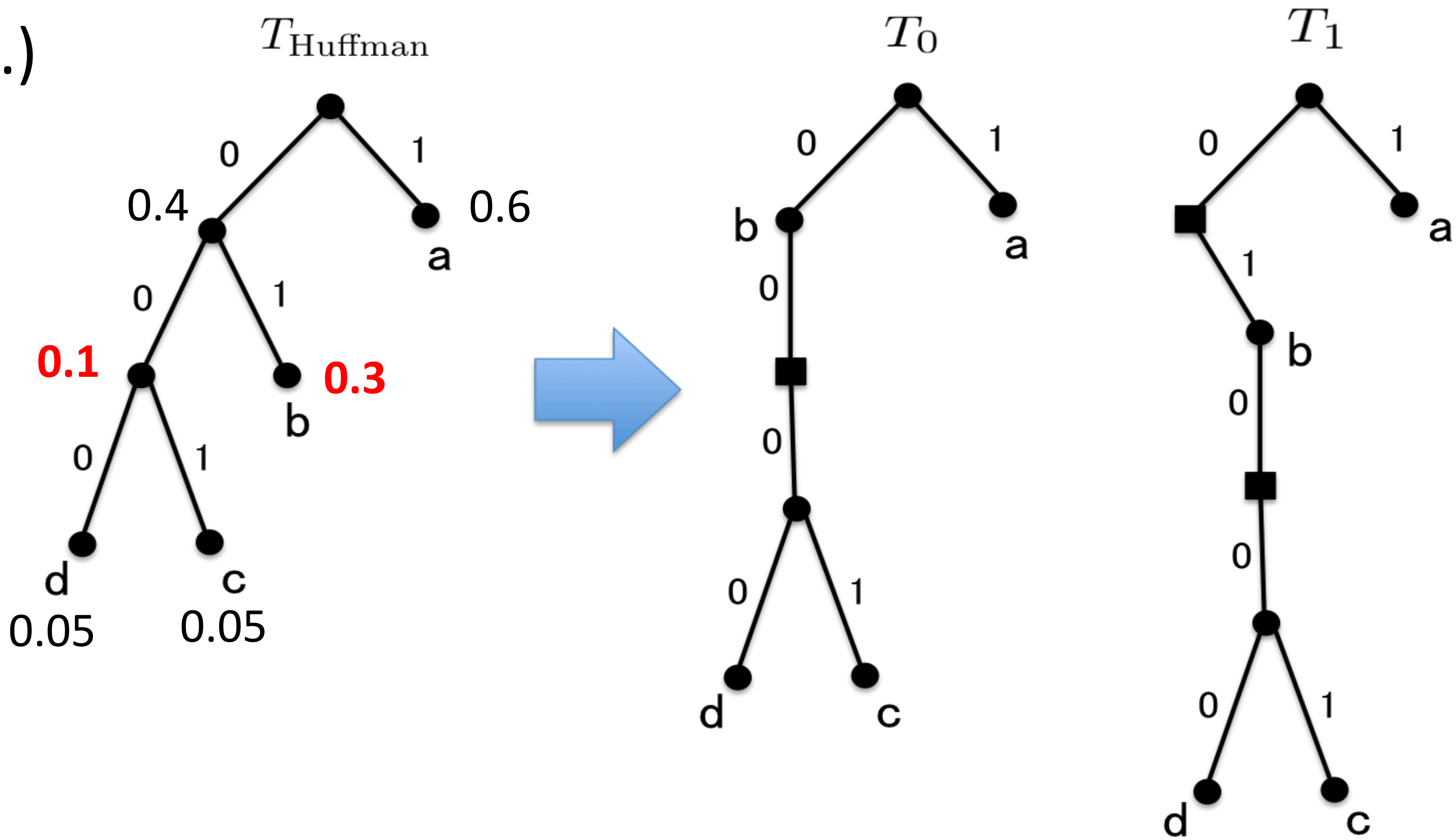
Only if $\frac{-1+\sqrt{5}}{2} \leq q_1$



Proof outline (4/6)

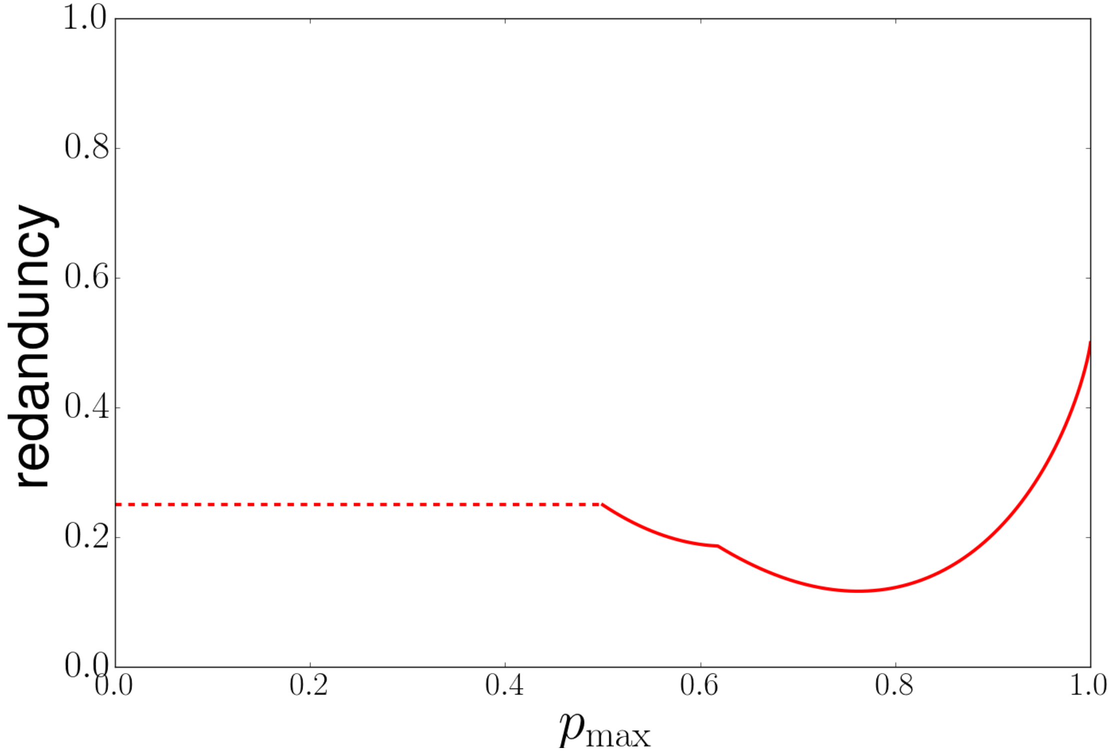
・Simple two-stage construction of sub-optimal AIFV code trees from Huffman tree

Ex.)



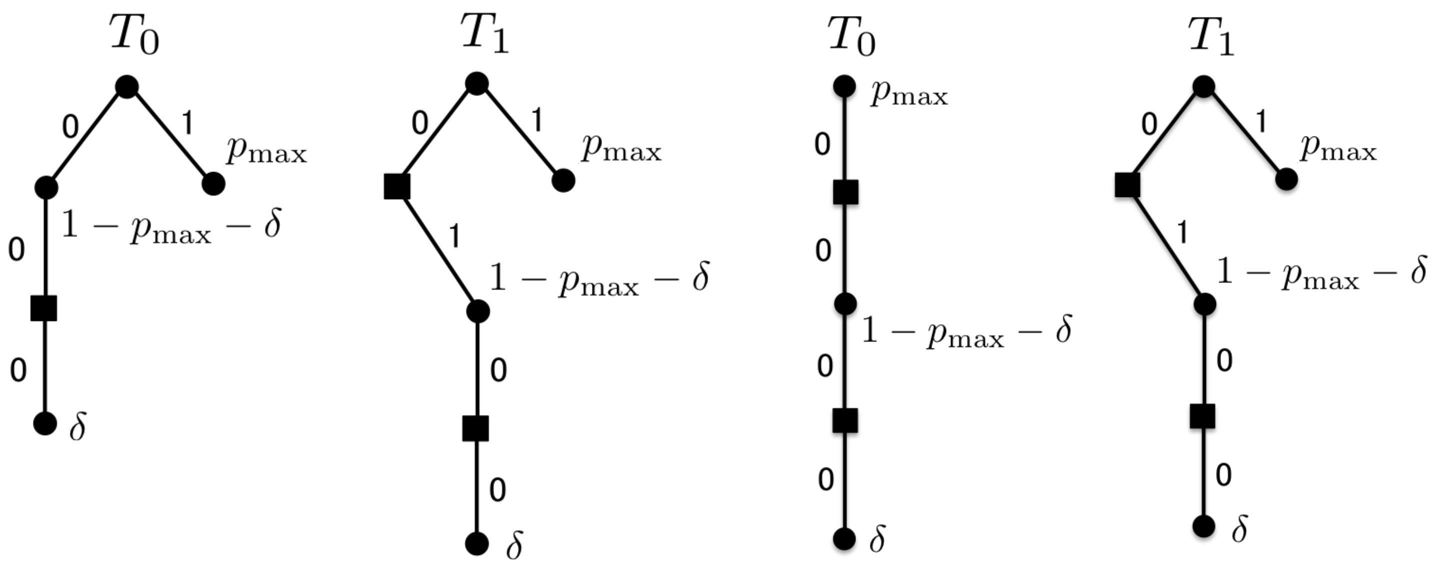
Proof outline (5/6)

- Upper bounds for sub-optimal AIFV code can be evaluated → tight for $p_{\max} \geq \frac{1}{2}$. **Why?**



Proof outline (6/6)

Optimal AIFV trees for $(p_{\max}, 1-p_{\max}-\delta, \delta)$ coincides with worst-case trees of sub-optimal AIFV codes.



(a) $\frac{1}{2} \leq p_{\max} \leq \frac{\sqrt{5}-1}{2}$. (b) $\frac{\sqrt{5}-1}{2} \leq p_{\max} \leq 1$.

→ The bound of sub-optimal trees is tight for $p_{\max} \geq \frac{1}{2}$.

Conclusion

1. Worst-case redundancy of binary AIFV codes is $\mathbf{1/2}$.
2. Worst-case redundancy in terms of $p_{\max} = p \geq 1/2$.

Theoretical justification for **superior performance of AIFV codes** over Huffman codes.

Further extension

If the codes are allowed to use **3 and 4 code trees**, worst-case redundancy is **1/3 and 1/4**, respectively.