Precise Asymptotic Analysis of the Tunstall Code

Michael Drmota*, Yuriy Reznik†, Serap A. Savari‡, and Wojciech Szpankowski§

* Institute of Discrete Mathematics and Geometry, TU Wien, Wiedner Hauptstr. 8–10, A-1040 Wien, Austria
Email: michael.drmota@tuwien.ac.at

† Qualcomm Inc., 5775 Morehouse Dr., San Diego, CA 92121,
Email: yreznik@qualcomm.com

‡ Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 48109,
Email: savari@eecs.umich.edu

§ Department of Computer Science, Purdue University, W. Lafayette, IN 47907,
Email: spa@cs.purdue.edu

***Abstract*— We study the Tunstall code using the machinery from the analysis of algorithms literature. In particular, we propose an algebraic characterization of the Tunstall code which, together with tools like the Mellin transform and the Tauberian theorems, leads to new results on the variance and a central limit theorem for dictionary phrase lengths. This analysis also provides a new argument for obtaining asymptotic results about the mean dictionary phrase length and average redundancy rates.**

I. Introduction

A variable-to-fixed length encoder partitions the source string over an m-ary alphabet $\mathcal{A}$ into a concatenation of variable-length phrases. Each phrase except the last one is constrained to belong to a given dictionary $\mathcal{D}$ of source strings; the last phrase is a non-null prefix of a dictionary entry. One common constraint on a dictionary is that it leads to a *unique* parsing of any string over $\mathcal{A}$ (see [13] for examples of dictionaries without this constraint). For the rest of the paper we will assume that all dictionaries are uniquely parsable. It is convenient to represent a uniquely parsable dictionary by a complete parsing tree $\mathcal{T}$, i.e., a tree in which every internal node has all m children nodes in the tree. The dictionary entries $d \in \mathcal{D}$ correspond to the leaves of parsing tree. The encoder represents each parsed string by the fixed length binary code word corresponding to its dictionary entry. If the dictionary $\mathcal{D}$ is has M entries, then the code word for each phrase has $\lceil \log_2 M \rceil$ bits. The best known variable-to-fixed length code is now generally attributed to Tunstall [22]; however, it was independently discovered by Khodak [7], Verhoeff [23], and possibly others. In this paper, we offer a new perspective and asymptotic analysis of the Tunstall code.

Tunstall's algorithm is simple to visualize through evolving parsing trees in which every edge corresponds to a letter from the source alphabet $\mathcal{A}$: starting from a tree with a root node and m leaves which together correspond to all of the symbols in $\mathcal{A}$). For J iterations we select the current leaf corresponding to a string of highest probability and grow m children out it, one for each symbol in $\mathcal{A}$. After these J steps, the parsing tree has J non-root internal nodes and $M = (m-1)J + m$ leaves, which each correspond to a distinct dictionary entry. The dictionary entries are prefix-free and can be easily enumerated.

Tunstall's algorithm has been studied extensively (see, e.g., the survey article [1]). Simple bounds for its redundancy have been independently obtained by Khodak [7] and by Jelinek and Schneider [6]. Tjalkens and Willems [19] were the first to look at extensions of this code to sources with memory. Savari and Gallager [11] proposed a generalization of Tunstall's algorithm for Markov sources and used renewal theory for an asymptotic analysis of average code word length and redundancy for memoryless and Markov sources. Savari [12] later published a non-asymptotic analysis of the Tunstall code for binary, memoryless sources with small entropies. Universal variable-to-fixed length codes were analyzed in [21], [10], [9], [8], [20], [24]; however, to the best of our knowledge the minimax redundancy for variable-to-fixed and variable-to-variable length codes has not been carefully studied.

Our goal is to establish the limiting distribution of the phrase length and provide a precise asymptotic analysis of the average redundancy of the Tunstall code. While the average redundancy of the Tunstall code for memoryless and Markov sources has been studied previously by Savari [11], [12], we provide here a new approach that allows us to precisely quantify oscillations involved in the redundancy for a certain class of sources. Our central limit theorem concerning the phrase length is new and has been derived in an analytic way that hopefully will serve as a template for the future analysis of variable-to-fixed length and variable-to-variable length codes.

To facilitate our analysis, we will focus upon another construction of the Tunstall code that was invented by Khodak [7] (see also [8]). Khodak independently discovered the Tunstall code using a rather different approach. Let p_i be the probability of the ith source symbol and let $p_{\min} = \min\{p_1, \ldots, p_m\}$. Khodak suggested choosing a real number $r \in (0, p_{min})$ and growing a complete parsing tree satisfying

$$p_{\min} r \le P(d) < r \,, d \in \mathcal{D}. \qquad (1)$$

It can also be shown (see, e.g., [6, Lemma 6] and [11, Lemma 2]) that the resulting parsing tree is exactly the same as a tree constructed by Tunstall's algorithm. The asymptotic relationship between r and the resulting number of dictionary entries M_r was studied in [11] and will be established here in a different way.

It follows from (1) that if y is a proper prefix of one or more entries of $\mathcal{D} = \mathcal{D}_r$, i.e., y corresponds to an internal node of $\mathcal{T} = \mathcal{T}_r$, then

$$P(y) \geq r.$$

As it is easier to directly characterize the internal nodes of the parsing tree $\mathcal{T}_r$ rather than its leaves, we shall approach the analysis of $\mathcal{D} = \mathcal{D}_r$ by representing the moment generating function of the phrase length in terms of the transform of the path lengths to internal nodes in $\mathcal{T}_r$. We will show that the moment generating function of the dictionary phrase length in the parsing tree satisfies certain recurrences that could surprisingly be analyzed through analytic algorithmic methods such as the Mellin transform and the Tauberian theorems. This analysis provides a precise asymptotic characterization of the behavior of the Tunstall code. In passing, we mention that this work directly extends recent analyses of fixed-to-variable codes (cf. [4], [5], [17], [18]) through tools of analytic algorithms and is hence in the domain of analytic information theory.

II. Main Results

Assume a memoryless source over an m-ary alphabet $\mathcal{A}$ generates an output sequence. Let $p_i > 0$ be the probability of the i^{th} letter of alphabet $\mathcal{A}$, $i \in \{1, \dots, m\}$, $p_{\min} = \min\{p_1, \dots, p_m\}$, and $p_{\max} = \max\{p_1, \dots, p_m\}$. Given a dictionary $\mathcal{D}$ and corresponding complete parsing tree $\mathcal{T}$, the encoder partitions the source output sequence into a sequence of variable-length phrases. Let $d \in \mathcal{D}$ denote a dictionary entry, $P(d)$ be its probability, and $|d|$ be its length. Our focus will be on the random variable $D = |d|$, the phrase length of a dictionary string. One of our goals is to investigate the moment generating function of the phrase length $D = D_r$ in Khodak's construction of the Tunstall dictionary with parameter r. That is, we consider

$$D(r, z) := \mathbf{E}[z^D] = \sum_{d \in \mathcal{D}_r} P(d) z^{|d|}.$$

Towards this end, we next introduce a second transform describing the probabilities of strings which correspond to internal nodes in the parsing tree $\mathcal{T}_r$. Let

$$S(r, z) = \sum_{y:\ P(y) \geq r} P(y) z^{|y|}. \tag{2}$$

Our first result considers *arbitrary* complete parsing trees, i.e., not necessarily Tunstall trees, and relates the transform for the probabilities of internal nodes to a function of the leaf probabilities.

Theorem 1: Let $\tilde{\mathcal{D}}$ be a uniquely parsable dictionary and $\tilde{\mathcal{Y}}$ be the collection of strings which are proper prefixes of one or more dictionary entries. Then for all $|z| \leq 1$,

$$\sum_{d \in \tilde{\mathcal{D}}} P(d) \frac{z^{|d|} - 1}{z - 1} = \sum_{y \in \tilde{\mathcal{Y}}} P(y) z^{|y|}. \tag{3}$$

Proof We use induction on the number of internal nodes in the corresponding dictionary tree. For the basis step, (3) is clearly true when $\tilde{\mathcal{D}} = \mathcal{A}$ since the only element of $\tilde{\mathcal{Y}}$ is the null string, which has probability one and length zero.

For the inductive step, suppose that (3) is true for all dictionaries with parsing trees having k internal nodes. Let $\tilde{\mathcal{D}}$ be a dictionary with a corresponding proper prefix set $\tilde{\mathcal{Y}}$ having $k+1$ elements. Choose $y_0 \in \tilde{\mathcal{Y}}$ to have maximum length so that its single letter extensions correspond to the dictionary entries $d_1, d_2, \dots, d_m \in \tilde{\mathcal{D}}$. Observe that $P(y_0) = P(d_1) + P(d_2) + \ldots + P(d_m)$. We next define an auxiliary dictionary $\tilde{\mathcal{D}}'$ with $\tilde{\mathcal{D}}' = \tilde{\mathcal{D}} \cup \{y_0\} \setminus \{d_1, \dots, d_m\}$. Then $\tilde{\mathcal{D}}'$ has a corresponding proper prefix set $\tilde{\mathcal{Y}}' = \tilde{\mathcal{Y}} \setminus \{y_0\}$ with k elements.

Using the inductive hypothesis, we have

$$\begin{aligned}
\sum_{y \in \tilde{\mathcal{Y}}} P(y) z^{|y|} &= \sum_{y \in \tilde{\mathcal{Y}}'} P(y) z^{|y|} + P(y_0) z^{|y_0|} \\
&= \sum_{d \in \tilde{\mathcal{D}}'} P(d) \frac{z^{|d|} - 1}{z - 1} + P(y_0) z^{|y_0|} \\
&= \sum_{d \in \tilde{\mathcal{D}}' \setminus \{y_0\}} P(d) \frac{z^{|d|} - 1}{z - 1} \\
&\quad + P(y_0) \left(z^{|y_0|} + \frac{z^{|y_0|} - 1}{z - 1} \right) \\
&= \sum_{d \in \tilde{\mathcal{D}}' \setminus \{y_0\}} P(d) \frac{z^{|d|} - 1}{z - 1} \\
&\quad + (P(d_1) + \ldots + P(d_m)) \left(\frac{z^{|y_0|+1} - 1}{z - 1} \right) \\
&= \sum_{d \in \tilde{\mathcal{D}}} P(d) \frac{z^{|d|} - 1}{z - 1}.
\end{aligned}$$

This completes the proof of the lemma. ■

Since $\mathbf{E}[D] = \sum_{d \in \tilde{\mathcal{D}}} P(d)\,|d|$, Theorem 1 offers a new proof of the well-known result that

$$\mathbf{E}[D] = \sum_{y \in \tilde{\mathcal{Y}}} P(y),$$

and it provides a new result for uniquely parsable dictionaries that

$$\mathbf{E}[D(D-1)] = 2 \sum_{y \in \tilde{\mathcal{Y}}} P(y) |y|.$$

Furthermore, Theorem 1 and equation (2) imply that for the Tunstall code

$$D(r, z) = 1 + (z - 1) S(r, z).$$

Thus we can express the moment generating function for the phrase length of a Tunstall dictionary entry in terms of the transform describing the probabilities of proper prefixes of the dictionary entries. As we will discuss below, this relationship enables us to exploit a recurrence description for our analysis of the Tunstall code.

Let $v = 1/r$, z be a complex number and define $\tilde{S}(v, z) = S(v^{-1}, z)$. We restrict our attention here to a binary alphabet $\mathcal{A}$ with $0 < p_1 < p_2 < 1$.

Let $A(v)$ devote the number of source strings with probability at least v^{-1}; i.e.,

$$A(v) = \sum_{y:P(y)\geq 1/v} 1. \tag{4}$$

The functions $A(v)$ and $\tilde{S}(v,z)$ satisfy the following recurrences.

$$A(v) = \begin{cases} 0 & v < 1, \\ 1 + A(vp_1) + A(vp_2) & v \geq 1 \end{cases}$$

and

$$\tilde{S}(v,z) = \begin{cases} 0 & v < 1, \\ 1 + zp_1\tilde{S}(vp_1,z) + zp_2\tilde{S}(vp_2,z) & v \geq 1, \end{cases}$$

since every binary string either is the empty string, a string starting with the first source letter $a_1 \in \mathcal{A}$, or a string starting with the letter $a_2 \in \mathcal{A}$. This partition directly leads to the recurrences above. Observe that $A(v)$ represents the number of internal nodes in Khodak's construction with parameter v^{-1} of a Tunstall tree, that is $M_r = A(v)+1 = |\mathcal{D}_r|$ is the dictionary size. Further, $\mathbf{E}[D_r] = \tilde{S}(v,1)$ is the corresponding expected value of the phrase length ($r = 1/v$).

These recurrences can be studied through the Mellin transform, see [2], [18]; in particular we find that as $v \to \infty$,

$$M_r = A(v) + 1 = \frac{v}{H} + o(v), \tag{5}$$

if $\ln p_2/\lg p_1$ is irrational and

$$M_r = A(v) + 1 = \frac{Q_1(\log v)}{H} v + O(v^{1-\eta}) \tag{6}$$

for some $\eta > 0$ if $\ln p_2/\ln p_1$ is rational, where

$$Q_1(x) = \frac{L}{1-e^{-L}} e^{-L\langle \frac{x}{L} \rangle} \tag{7}$$

and $L > 0$ is the largest real number for which $\ln(1/p_1)$ and $\ln(1/p_2)$ are integer multiples of L; $H = p_1 \ln(1/p_1) + p_2 \ln(1/p_2)$ is the entropy rate in *natural* units and $\langle y \rangle = y - \lfloor y \rfloor$ is the fractional part of the real number y. Observe that when $\log p_2/\log p_1 = b/d$ for some integers b, d such that $\gcd(b,d) = 1$ we can also write $L = \log(1/p_2)/b = \log(1/p_1)/d$.

Similarly,

$$\mathbf{E}[D_r] = \tilde{S}(v,1) = \frac{\log v}{H} + \frac{H_2}{2H^2} + o(1) \tag{8}$$

in the irrational case and

$$\mathbf{E}[D_r] = \tilde{S}(v,1) = \frac{\log v}{H} + \frac{H_2}{2H^2} + \frac{Q_2(\log v)}{H} + O(v^{-\eta}) \tag{9}$$

for some $\eta > 0$ in the rational case, where

$$Q_2(x) = L \cdot \left(\frac{1}{2} - \left\langle \frac{x}{L} \right\rangle \right) \tag{10}$$

and $H_2 = p_1 \ln(1/p_1)^2 + p_2 \ln(1/p_2)^2$.

Whereas the proof in the rational case is elementary (i.e., complex analysis is not used), the irrational case requires the use of Wiener's Tauberian theorem (cf. [2]).

In order to obtain distributional results on D we have to analyze $\tilde{D}(v,z) = D(1/v,z) = 1 + (z-1)\tilde{S}(v,z)$ uniformly for z in a neighborhood of $z = 1$. Although the recurrence for $\tilde{S}(v,z)$ looks similar to that of $A(v)$ its analysis is more complex. The main technical problem lies in the slow convergence rates of certain series. In this conference paper, we merely sketch our approach.

The Mellin transform $F^*(s)$ of a function $F(v)$ is defined as (cf. [18])

$$F^*(s) = \int_0^\infty F(v) v^{s-1} dv.$$

A simple calculation shows

$$\tilde{D}^*(s,z) = \frac{1-z}{s(1 - zp_1^{1-s} - zp_2^{1-s})} - \frac{1}{s}, \qquad \Re(s) < s_0(z),$$

where $s_0(z)$ denotes the real solution of $zp^{1-s} + zq^{1-s} = 1$.

In order to find the asymptotics of $\tilde{D}(v,z)$ as $v \to \infty$ we compute the inverse transform of $\tilde{D}^*(s,z))$:

$$\tilde{D}(v,z) = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\sigma - iT}^{\sigma + iT} \tilde{D}^*(s,z) v^{-s}\, ds,$$

where $\sigma < s_0(z)$. For this purpose it is usually necessary to determine the polar singularities of the meromorphic continuation of $\tilde{D}^*(s,z)$ to the range $\Re(s) \geq s_0(z)$, that is, we have to analyze the set

$$Z(z) = \{s \in \mathbf{C} : zp^{1-s} + zq^{1-s} = 1\} \tag{11}$$

of all complex roots of $zp^{1-s} + zq^{1-s} = 1$ It can be proved [18] that $Z(z) = \{s_k(z) : k \in \mathbf{Z}\}$ with $\Re(s_k(z)) \geq s_0(z)$ and $(2k-1)\pi/\ln p_1 \leq \Im(s_k(z)) \leq (2k+1)\pi/\ln p_1$.

Hence, by the residue theorem we obtain $\tilde{D}(v,z) = \lim_{T\to\infty} F_T(v,z)$ for every $\sigma > s_0(z)$ with $\sigma \notin \{\Re(s) : s \in Z(z)\}$ where

$$\begin{aligned} & F_T(v,z) \\ &= -\sum_{s'\in Z(z),\ \Re(s')<\sigma, |\Im(s')|>T} \mathrm{Res}(\tilde{D}^*(s,z)\, v^{-s},\ s = s') \\ &+ \frac{1}{2\pi i} \int_{\sigma - iT}^{\sigma + iT} \left(\frac{1-z}{s(1 - zp_1^{1-s} - zp_2^{1-s})} - \frac{1}{s} \right) v^{-s}\, ds \\ &= -\sum_{s'\in Z(z),\ \Re(s')<\sigma, |\Im(s')|>T} \frac{(1-z)v^{-s'}}{zs'p_1^{1-s'}\ln p_1 + zs'p_2^{1-s'}\ln p_2} \\ &+ \frac{1}{2\pi i} \int_{\sigma - iT}^{\sigma + iT} \left(\frac{1-z}{s(1 - zp_1^{1-s} - zp_2^{1-s})} - \frac{1}{s} \right) v^{-s}\, ds \end{aligned}$$

provided that the series of residues converges and the limit as $T \to \infty$ of the last integral exists. The problem is that neither the series nor the integral above are absolutely convergent since the integrand is only of order $1/s$.

After some careful analysis, it is possible to derive the following results.

Theorem 2: Let D_r denote the phrase length in Khodak's construction of the Tunstall code with a dictionary of size M_r

over a biased memoryless source. Then as $M_r \to \infty$

$$\frac{D_r - \frac{1}{H}\ln M_r}{\sqrt{\left(\frac{H_2}{H^3} - \frac{1}{H}\right)\ln M_r}} \to N(0,1)$$

where $N(0,1)$ denotes the standard normal distribution. Furthermore, we have $E[D] = \frac{\ln M_r}{H} + O(1)$ and

$$\mathbf{Var}[D_r] = \left(\frac{H_2}{H^3} - \frac{1}{H}\right)\ln M_r + O(1)$$

for large M_r.

By combining (5) and (8) resp. (6) and (9) we can be even more precise. In the irrational case (i.e., for $\log p_2/\log p_1 = b/d$) we have

$$\mathbf{E}[D_r] = \frac{\ln M_r}{H} + \frac{\ln H}{H} + \frac{H_2}{2H^2} + o(1)$$

and in the rational case

$$\begin{aligned}\mathbf{E}[D_r] &= \frac{\ln M_r}{H} + \frac{\ln H}{H} + \frac{H_2}{2H^2} + \frac{Q_2(\ln v) - \ln Q_1(\ln v)}{H} \\ &+ O(M_r^{-\eta}).\end{aligned}$$

Note that (7) and (10) yield

$$Q_2(\ln v) - \log Q_1(\ln v) = -\ln L + \ln(1-e^{-L}) + \frac{L}{2}$$

so that there is actually no oscillation.

As a direct consequence, we can derive a precise asymptotic formula for the average redundancy of the Tunstall code that was defined in [11] by

$$\mathcal{R}_M = \frac{\ln M}{\mathbf{E}[D]} - H\,.$$

The following result is a consequence of the above derivations.

Corollary 1: Let $\mathcal{D}_r$ denote the dictionary in Khodak's construction of the Tunstall code of size M_r. If $\ln p_1/\ln p_2$ is irrational then

$$\mathcal{R}_{M_r} = \frac{H}{\ln M_r}\left(-\frac{H_2}{2H} - \ln H\right) + o\left(\frac{1}{\ln M_r}\right).$$

In the rational case we have

$$\begin{aligned}\mathcal{R}_{M_r} &= \frac{H}{\ln M_r}\Big(-\frac{H_2}{2H} - \ln H \\ &\quad + \ln L - \ln(e^L - 1) + \frac{L}{2}\Big) + O\left(M_r^{-\eta}\right),\end{aligned}$$

for some $\eta > 0$, where $L > 0$ is the largest real number for which $\ln(1/p_1)$ and $\ln(1/p_2)$ are integer multiples of L.

In passing we observe that the Corollary 1 is a special case of Theorems 5 and 12 of [11].

Acknowledgment

The work of M. Drmota was supported in part by the Austrian Science Foundation FWF Grant No. S9604. The work of S. Savari was supported in part by NSF Grant No. CCF-0430201. The work of W. Szpankowski was supported in part by the NSF Grants CCR-0208709, CCF-0513636, and DMS-0503742, NIH Grant R01 GM068959-01, and the AFOSR Grant FA8655-04-1-3074.

References

[1] J. Abrahams, "Code and parse trees for lossless source encoding," *Communications in Information and Systems* 1(2):113-146, April 2001.
[2] V. Choi and M. J. Golin, Lopsided trees. I. Analyses. *Algorithmica* 31 (2001), no. 3, 240–290.
[3] T. M. Cover and J. M. Thomas, *Elements of Information Theory*, (John Wiley & Sons, New York, 1991).
[4] M. Drmota and W. Szpankowski, Precise Minimax Redundancy and Regrets, *IEEE Trans. Information Theory*, 50, 2686-2707, 2004.
[5] P. Jacquet and W. Szpankowski, Markov Types and Minimax Redundancy for Markov Sources *IEEE Trans. Information Theory*, 50, 1393-1402, 2004.
[6] F. Jelinek, and K. S. Schneider, On Variable-Length-to-Block Coding, *IEEE Trans. Inform. Theory*, **18** (6) (1972) 765–774.
[7] G. L. Khodak, Connection Between Redundancy and Average Delay of Fixed-Length Coding, *All-Union Conference on Problems of Theoretical Cybernetics* (Novosibirsk, USSR, 1969) 12 (in Russian)
[8] R. E. Krichevsky, Universal Data Compression and Retrieval. (Kluwer, Norwell, MA, 1993).
[9] R. E. Krichevsky and V. K. Trofimov, The Performance of Universal Encoding, *IEEE Trans. Information Theory*, **27** (1981) 199–207.
[10] J. C. Lawrence, A New Universal Coding Scheme for the Binary Memoryless Source, *IEEE Trans. Inform. Theory*, 23 (4) (1977) 466–472.
[11] S. A. Savari, Robert G. Gallager; Generalized Tunstall codes for sources with memory, IEEE Trans. Info. Theory, vol. IT-43, pp. 658 - 668, March 1997.
[12] S. A. Savari, Variable-to-Fixed Length Codes for Predictable Sources, Proc IEEE Data Compression Conference (DCC'98), Snowbird, UT, March 30 - April 1, 1998, pp. 481–490.
[13] S. A. Savari, Renewal Theory and Source Coding, *Proceedings of the IEEE*, vol. 88, no. 11, pp. 1692-1702, November 2000.
[14] W. Schachinger, Limiting distributions for the costs of partial match retrievals in multidimensional tries. *Random Structures and Algorithms* 17 (2000), no. 3-4, 428–459.
[15] J. P. M. Schalkwijk, An Algorithm For Source Coding, *IEEE Trans. Inform. Theory*, 18 (3) (1972) 395–399.
[16] V. M. Sidelnikov, On Statistical Properties of Transformations Carried out by Finite Automata, *Cybernetics*, 6 (1965) 1–14 (in Russian)
[17] W. Szpankowski, Asymptotic Average Redundancy of Huffman (and other) Block Codes, *IEEE Trans. Information Theory*, 46, 2434-2443, 2000.
[18] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, Wiley, New York, 2001.
[19] T. J. Tjalkens and F. M. J. Willems, "Variable to fixed-length codes for Markov sources," *I.E.E.E. Trans. Inform. Theory* IT-33, 246-257, 1987.
[20] T. Tjalkens and F. Willems, A Universal Variable-to-Fixed Length Source Code Based on Lawrence's Algorithm, *IEEE Trans. Information Theory*, 38, 247-253, 1992.
[21] V. K. Trofimov, Universal Variable-Length to Block Codes for Bernoulli Sources, *Methods of Discrete Analysis*, 29 (Novosibirsk, 1976) 87–100 (in Russian)
[22] B. P. Tunstall, Synthesis of Noiseless Compression Codes, Ph.D. dissertation, (Georgia Inst. Tech., Atlanta, GA, 1967)
[23] J. Verhoeff, A new data compression technique, *Annals of Systems Research*, **6** (1977) pp. 139-148.
[24] K. Visweswariah, S. R. Kulkarni, and S. Verdu, "Universal variable-to-fixed length source codes," IEEE Trans. Info. Theory, vol. IT-47, pp. 1461 - 1472, May 2001.