Proof of the Median Paths

Thijs Veugen
TNO
Technical Sciences
The Hague, The Netherlands
thijs.veugen@tno.nl

Abstract

We consider the problem of coding for discrete memoryless channels with noiseless feedback. When studying Horstein's sequential coding scheme, Schalkwijk in 1971 found a regular behaviour of the so-called median paths for certain channel error probabilities, which led to the development of repetition strategies. We proof that Schalkwijk's block decoding scheme exactly follows Horstein's regular median paths.

1 Introduction

We consider the problem of coding for discrete memoryless channels with noiseless feedback [4], which is strongly related to Ulam's game [3, p. 281]. In this game, person A knows a secret integer m, which person B should guess by constructing sets one by one, and asking whether m is in it or not. Person A is allowed to lie a fixed number of times, and B wants to minimise the number of required questions.

During the game, person B is maintaining a probability distribution on the possible integers of A. We will show that the median of this distribution plays a very important role. In 1963, Horstein [1] developed a sequential coding scheme for the binary symmetric channel with noiseless feedback, by closely following the median of the receiver's distribution. Schalkwijk [2] in 1971 found out that, for certain channel error probabilities, these medians exhibit a regular pattern, which led to the development of a simple block coding scheme, achieving channel capacity for those particular channel error probabilities.

However, this regular behaviour of the median, visualised by median paths, has never been proven. In this paper we present a proof, which has been known since 1997, but has never been published.

2 Median paths

In 1963, Horstein [1] developed a sequential coding scheme for the binary symmetric channel with noiseless feedback. Let p, $0 \leq p < \frac{1}{2}$, be the error probability of the binary symmetric channel, and $q = 1 - p$. Suppose the sender has an (infinite) binary message, presented as a point m in the message interval $[0, 1]$, which he would like to send towards the receiver. Assuming each message being equally likely, the initial message distribution of the receiver will be a uniform distribution over the interval $[0, 1]$. The idea of Horstein was as follows. He introduced a parameter a, $0 < a < 1$. Before transmitting the n^{th} bit, the sender will compute the currently expected message distribution, from the viewpoint of the receiver, and a message point m_n, such that $\Pr(m \geq m_n) = a$. In case $m \geq m_n$, the sender will transmit a 1, and a 0, otherwise. Because of the assumed noiseless feedback link, the sender learns which output bits have been received by the receiver, which enables him to compute the expected message distribution.

In case the receiver just got a 1, the message points in the $[m_n, 1]$ message interval become more likely, and the message points in the $[0, m_n)$ interval become less likely. The channel output distribution $\underline{\pi}$ can be computed as $\pi_0 = aq + (1-a)p$ and $\pi_1 = (1-a)q + ap$, and thereby the transmission rate

$$R(a) = I(X;Y) = H(Y) - H(Y|X) = h(\pi_1) - h(p),$$

where I represents mutual information, H information entropy, and h binary entropy.

From this equation, we derive that the transmission rate is maximised by setting $a = \frac{1}{2}$, because it leads to a uniform output distribution. The maximal transmission rate equals channel capacity $1 - h(p)$. This value of a corresponds with the *median* of all message points. In this case, after receiving a 1, the probability density of all messages in the upper interval $[m_n, 1]$ increases with a factor $\frac{q}{\pi_1} = 2q$, while in the lower interval $[0, m_n)$, the probability density will decrease with a factor $\frac{p}{\pi_0} = 2p$.

Figure 1 depicts an example of a median tree, showing all possible medians during 7 transmissions, given a binary channel with error probability $p = 0.1$. All medians have different values, and their pattern seems to be irregular. In 1971, Schalkwijk [2] discovered that, for certain channel error probabilities, specific median values keep recurring, and the consecutive medians follow a more regular pattern. An example of a regular median tree, for error probability $p = (3 - \sqrt{5})/4$, is depicted in Figure 2.

As explained earlier, after receiving a 1, the median moves up, and goes down, otherwise. The medians from Figure 2 have an additional property, namely that after receiving 100 or 011, the median returns to exactly the same value as three transmissions earlier. Not very surprisingly, this occurs when $2p \cdot 2q \cdot 2q = 1$, which explains the value $p = (3 - \sqrt{5})/4$. Schalkwijk observed that a receipt of 1000 leads to the same median as after receiving 0, which could be interpreted as an erroneously transmitted 0, followed by three additional 0's trying to correct this transmission error. This led to the birth of *repetition* strategies [2, 4].

3 Proof

For each integer value of k, $k \geq 3$, regular median trees can be found, where the median returns to its original value after receiving 10^{k-1} or 01^{k-1}. This occurs exactly for the solution p of

$$(2p) \cdot (2q)^{k-1} = 1.$$

In the corresponding repetition block-coding strategy, the sender is transmitting a fixed-length message, which does not contain the subsequences 10^k and 01^k. Each time a transmission error occurs, which can be detected because of the noiseless feedback link, the sender repeats the erroneously received symbol k times. This is applied in a recursive way, so a transmission error during a repetition sequence leads to k additional symbols to be sent.

The question is: why does an 'error correction', implemented as substituting, from right to left, any subsequence 10^k or 01^k in the received sequence by 0 or 1, respectively, not affect the values of the corresponding medians?

Because of the way (regular) median trees are constructed, we have the following **Property**:
Suppose we have an arbitrary receiver's distribution [1] to start with. Let m_i, $0 \leq i \leq k+1$, be the medians after receiving the first i symbols of the sequence 10^k. By definition of Horstein's scheme (assume the median goes up when receiving 1), it follows that $m_0 < m_1$, $m_1 > m_2 > \ldots > m_{k+1}$, and $m_k = m_0$. Furthermore, the receiver's distribution after transmission k is equal to the original distribution, when restricted to the interval $[0, m_0]$.

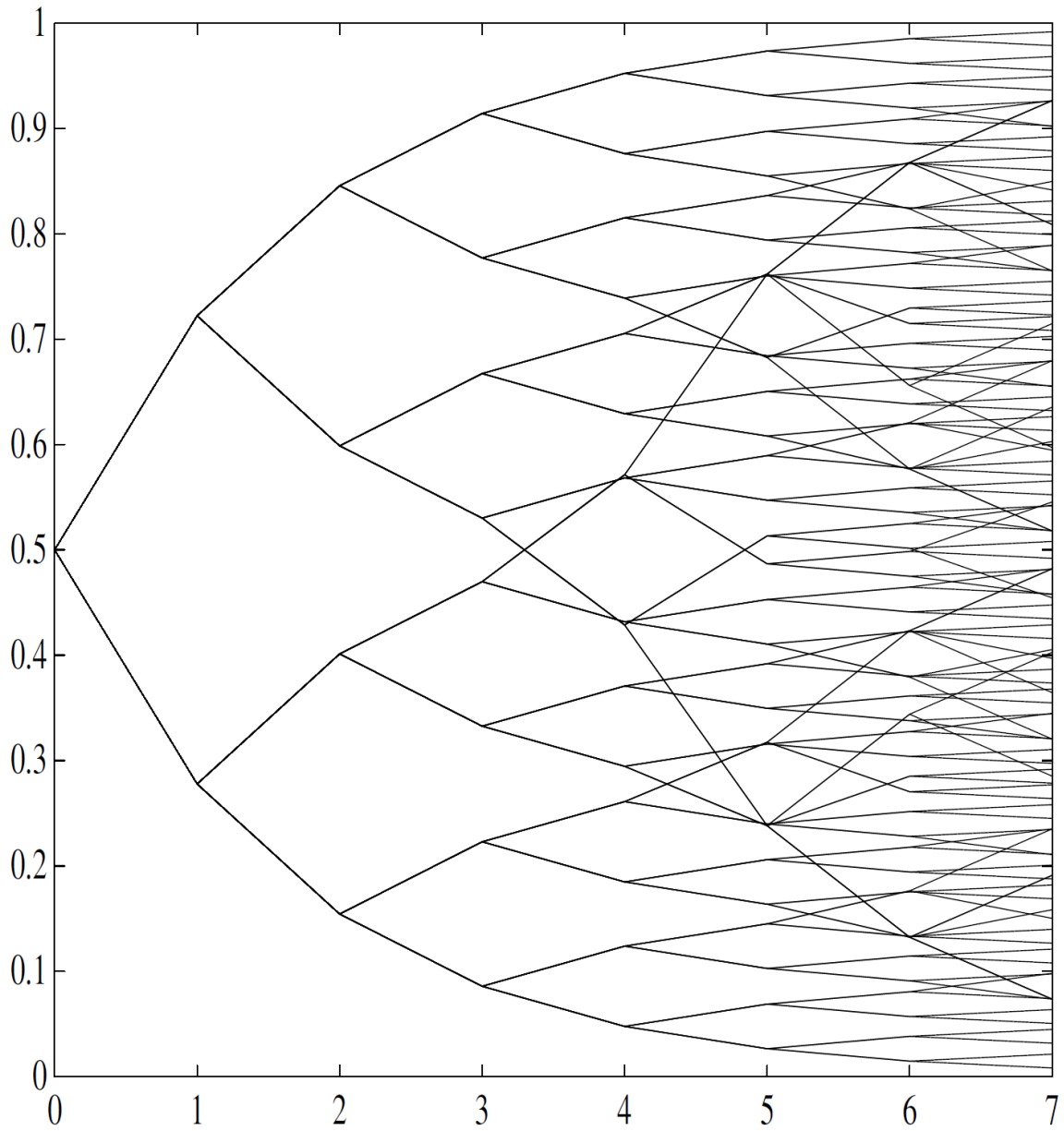

Figure 1: Median tree for $p = 0.1$

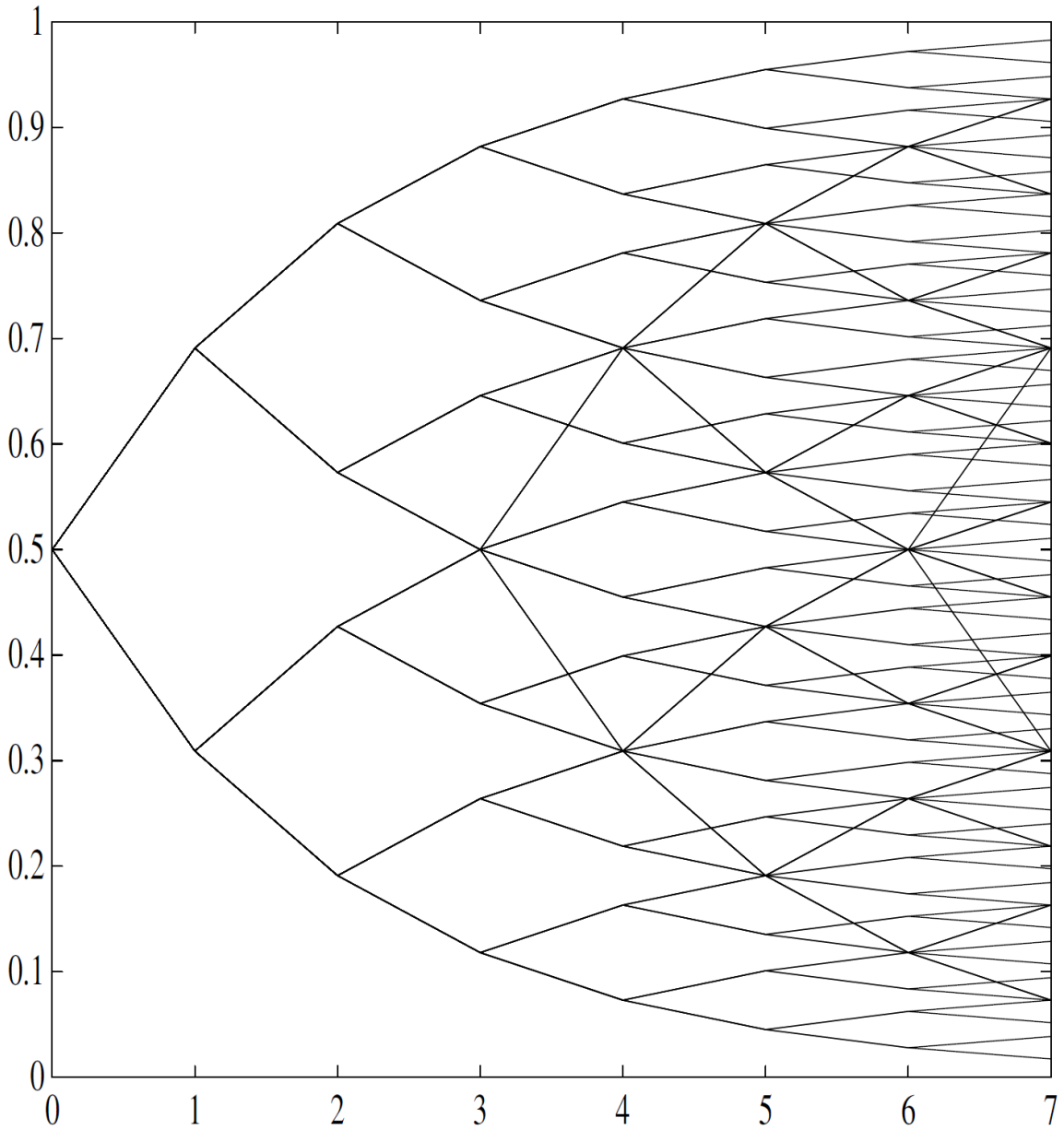

Figure 2: Median tree for $p = (3 - \sqrt{5})/4$

Using this property, we can proof the following lemma.

Lemma 1. *Let $\underline{x} = x_1 \dots x_N$ be an arbitrary received sequence without forbidden subsequences. Let m_i, $0 \leq i \leq N$, be the medians after receiving $x_1 \dots x_i$. If $x_1 = 0$, then $m_i \leq m_0$ for $0 \leq i \leq N$.*

Proof: Assume $x_1 = 0$. We proof by natural induction $m_i \leq m_0$ for $0 \leq i \leq N$. The basis is $n = 0$. The induction step is as follows. Suppose that for some n, $n \geq 0$, we have $m_i \leq m_0$ for all i, $0 \leq i < n$. Assume $m_n > m_0$. Since $m_{n-1} \leq m_0 < m_n$, it follows that $x_n = 1$. If $x_{n-1} = 0$, then $m_{n-2} \geq m_n$, according to the property above, which would contradict $m_{n-2} \leq m_0$, so $x_{n-1} = 1$. Similarly, it is shown that $x_{n-2} = 1$, and $x_{n-3} = 1$, etc., until $x_{n-k+1} = x_{n-k+2} = \dots = x_n = 1$. Now, because $\underline{x}$ contains no forbidden subsequences, we conclude $x_{n-k} = x_{n-k-1} = \dots x_1 = 1$, which contradicts our first assumption $x_1 = 0$. So it must be that $m_n \leq m_0$. □

This lemma is used in the proof of our main theorem, which answers the research question.

Theorem 1. *Let $\underline{x} = x_1 \dots x_N$ be an arbitrary received sequence. Let m_i, $0 \leq i \leq N$, be the medians after receiving $x_1 \dots x_i$. Let $x_n \dots x_{n+k}$ be the rightmost forbidden subsequence. Let $\underline{x'} = x'_1 \dots x'_{N-k}$ be the sequence after substitution of $x_n \dots x_{n+k}$ by x_{n+k} in $\underline{x}$. Let m'_i, $0 \leq i \leq N-k$, be the medians after receiving $x'_1 \dots x'_i$. Then*

$$m'_i = m_{i+k}, n \leq i \leq N-k.$$

Proof: Since $x_1 \dots x_{n-1} = x'_1 \dots x'_{n-1}$, it follows that $m_{n-1} = m'_{n-1}$. Furthermore, by the previous property, $m_{n+k-1} = m_{n-1}$. Assume w.l.o.g. $x_n = 1$. Since $x_n \dots x_{n+k}$ is a forbidden subsequence, $x_{n+k} = 0$. Because $x_{n+k} \dots x_N$ is a sequence without forbidden subsequences, it follows from our Lemma that $m_i \leq m_{n+k-1}$ for $n+k-1 \leq i \leq N$. Due to our property, the receiver's distribution after receipt of x_{n+k-1} is equal to the receiver's distribution after receiving x_{n-1}, when restricted to the interval $[0, m_{n+k-1}]$. Since each median after transmission $n+k-1$ lies in this interval, it is easily shown that $m'_i = m_{i+k}$ for $i = n-1, n, \dots, N-k$, respectively. □

The theorem explains that substituting the rightmost forbidden subsequence does not affect the corresponding median values. Therefore, the entire decoding process as a whole will not affect the median values of the original message (assuming all transmission errors have been corrected). This affirms that Schalkwijks repetition block coding strategies naturally fit into Horstein's sequential coding scheme.

4 Conclusion

We introduced Horstein's sequential coding scheme for the binary symmetric channel with noiseless feedback. We showed the existence of regular median trees, which led Schalkwijk to the discovery of repetition block coding strategies. Although at first sight, these repetition strategies seem to form a natural extension of Horstein's scheme, it is not straightforward to see that the behaviour of the medians is in line with the entire decoding process. We were able to prove this, affirming their conceptual coherence.

References

[1] Michael Horstein. Sequential transmission using noiseless feedback. IEEE Transactions on Information Theory, 9: 136-143, July 1963.

[2] J. Pieter M. Schalkwijk. A class of simple and optimal strategies for block coding on the binary symmetric channel with noiseless feedback. IEEE Transactions on Information Theory, 22(9): 1369-1373, May 1971.

[3] S.M. Ulam. Adventures of a mathematician. Charles Scribner's sons, 1976.

[4] Thijs Veugen. Multiple-repetition coding for channels with feedback. PhD thesis. Eindhoven, University of Technology. June 1997.