

Modelling of Convolutional Encoder with Viterbi Decoder for the Next Generation Broadband Wireless Access Systems

Nnochiri, I.U., M. Eng.,

Abstract— This research focuses on modelling of convolutional encoder with Viterbi decoder for next generation broadband wireless access systems. Path loss and path fading can introduce errors to the signal during propagation, thus channel coding is introduced to overcome these problems. Convolutional codes are non blocking codes that can be designed to either correct or detect errors arising from Path loss and path fading of signal during propagation. Viterbi decoder makes use of the Viterbi algorithm in decoding a bit stream which has been encoded with a convolutional code. It estimates actual bit sequence using trellis diagram. Matlab software was used in the simulation of the model which was carried out over an Additive White Gaussian Noise (AWGN) channel using the Binary Phase Shift Keying (BPSK) modulation technique. Using the BER value of 10^{-5} , a better coding gain of 5.4dB was actualised, and by adjusting the trace-back depth implemented in the Viterbi algorithm, a stronger bound for the modelled code rates for the Additive White Gaussian Noise channel was realized and set as a measuring yardstick.

Index Terms— Convolution, Encoding, Viterbi, Decoding, Evaluation

I. INTRODUCTION

Owing to the incessant demand for bandwidth by bandwidth-hungry application and digital communication equipment miniaturisation, the need to design a good encoders and decoders for the next generation wireless communications system became very important.

Broadband Wireless Access system is one technology that provides the users with an option to wired access such as Digital Subscriber Lines (DSL), fibre optic link and coaxial cable system with regards to coverage, speed, and capacity [1]. It is suggested to work efficiently in the 2 gigahertz – 11 gigahertz spectrum frequency aiming at 1000Mbit/s data rate for a fixed or slow dynamic user and 100Mbit/s for a high accelerating vehicle.

A high speed communication system tackles the resultant errors that occur when data is transmitted across an impaired channel. Such errors can occur in the form of fading, Inter-signal interference, ISI or noisy channel [2]. Therefore, for the next generation BWA system to obtain an efficient and reliable data communication, it must employ the use of a method which can efficiently and effectively locate and correct errors; so as to help forward the standard established

by IEEE for broadband wireless access systems. The operations involved in locating and correcting errors in a communication system is simply called Channel Coding (CC).

Convolutional encoders with Viterbi decoders are techniques used in correcting errors which are greatly deployed in communication systems to better the BER performance [3]. The source encoder converts the signals meant to be transmitted from analogue to digital format. Redundancy in the signal is removed here by source coding and the information is then further compressed or converted into a sequence of binary digits for onward storage or transmission. Encryption on its own is the process of adding redundancy for security purposes. The information sequence is transformed by the Channel Encoder into encoded sequence and redundant information incorporated into the generated binary data at encoder for the purpose of removing noise such that the sequential data can be accurately recovered at the receiving end.

These binary data are generated by the source encoder from the source. Therefore, the information sequence stored in the source encoder is changed by the channel encoder to a discrete encoded sequence known as a codeword. By modulating the channel encoder, data stream for transmission coming from the channel encoder are converted into waveforms of time duration T sec. making it suitable to be transmitted over the communication channel which can either be a wired or wireless system. On entering the channel which is a physical medium used for information transmission, the waveform is affected by noise existing in the channel which could be in form of thermal, crosstalk and switch impulse. This waveform is then reduced to a sequence of numbers by the demodulator in which the acquired channel errors are processed and an error free output sequence produced which matches the estimated transmitted data symbols and termed the received sequence. The channel encoder tries to reproduce the main signal sequence by using the knowledge of code employed by the channel encoder and the redundancy present in the received signal. The source decoder receives the output information from the channel encoder and translates this sequence of binary digits into an estimate of the output source which it then forwards to the information sink. Information sink stand as the final destination to the main signal that was transmitted.

Basically, what is actually implemented here or the main objective is that the Probability of bit error (P_b), or required Signal-to-Noise Ratio (SNR) is reduced at the expense of

Manuscript received November 28, 2014.

Nnochiri, I.U., M. Eng., Department of computer Science, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

trading more bandwidth than would be required if an un-coded signal were to be transmitted [4].

II. EVOLUTION OF WIRELESS COMMUNICATION

Wireless in a nutshell could be said to be an encompassing word used to depict the transmission of signal across radio frequency channels via the air interface other than employing the use of a kind of physical media (e.g. cable) in bridging the communication service as seen in the last few decades.

By evolution of wireless communication we mean the classification of constant transitional change in the history of wireless communication based on the nature its services have experienced since the invention of the first generation (1G) network system. This also encompasses its irreversible compatibility transmitting technology and recent frequency bands. Prior to 1981, what we had was the 1G network (analogue) which was transited to digital, 2 generation network in 1992. This was followed by the 2.5G network (which has a technology that allows for the use of infrastructure and facilities only when a transaction is required due to its packet based nature) and then the 3G network in 2002 (which possesses a multi-media support, spread spectrum transmission and with nothing less than 200Kbps data rate) [5].

Therefore in summary, we can say that since 1980, data network have enjoyed a steady growth and success. For many companies, the corporate network is a crucial information resource that enables business successes. In the public domain, the internet provides vast possibilities for information retrieval and communication. Today, we face a great demand for high speed wireless access to these data networks which means that at least 10Mbps of data rate should be made available for mobile equipment.

The third generation wireless networks 3G such as Universal Mobile Telecommunication System (UMTS) enhanced by high speed packet access air interface have shown significant progress in mobile data rates, but still we are limited to few megabits per second (Mbps) user data rate. The most powerful available data network technology is based on IEEE 802.11a/b/g WLAN's family. It provides up to 54Mbps within a range of a few hundred meters. If a wider coverage is to be achieved, for example, a university campus or an entire city, many wireless LANs must be interconnected. In 2004, IEEE 802.16-2004 working group developed a wireless metropolitan area wireless network standard known as WiMAX (Worldwide interoperability for Microwave Access). WiMAX provides data rate of up to 100Mbps using one single base station that covers an area of several square kilometres. By the end of 2005, the amendment of the WiMAX standard was approved, as mobility and fast hand-out capability to the WiMAX standard. The IEEE 802.16-e is a solution for a Wireless Wide Area Network (WWAN) that equals the legacy 3GPP network but provides much data rate network for user applications. To complete our consideration, we can add two classes at the lower range level which are the Personal Area Network (PAN) which interconnects PDA, mobile phones, hands free kits etc. within a few meters. PAN standards includes Bluetooth (IEEE 802.15.1), universal wireless broadband (IEEE 802.15.3a) and ZigBee (IEEE 802.15.4). Finally, Wireless Body Area

Networks (WBAN) interconnects a handset to an intelligent health care device using Radio Frequency Identification (RFID) or Near-Field communication (NFC). With these entire wireless networks co-existing in the future, the need for a media independent handover is obvious. The IEEE 802.21 working group prepares for a serious handover between all wireless standards including the 3GPP networks [6]

Currently, we are in the era of 4G network which is an all-IP packet-switched mobile network, multi-carrier transmission and mobile ultra-broadband capable of supporting an array of multi-media services which includes moving pictures with towering speed and unrivalled quality, employing a focused wireless infrastructural network. Hence, if we can put together some services, such as voice service, then a variety of video, data and audio services will be obtainable. 4G is the fourth generation of mobile telecommunications technology, succeeding 3G. A 4G system, in addition to the usual voice and other services of 3G, provides mobile Ultra-broadband Internet access. It was originally envisioned by the defense Advanced Research Projects Agency (DARPA).

III. A SIMPLE CHANNEL MODEL

In practice, it is the responsibility of a communication system to transmit information from source to destination. The original information signal passes through different communication system components and then undergoes so many changes in its orientation and shape because of noise and attenuation.

As shown in the figure1 below, a communication system consists of source of information, transmitter, transmission medium (channel), receiver and destination (sink). To reach its destination, the signal generated by the source of information needs to pass through a certain medium of communication. Depending on the capacity of the channel (medium) used; the signal can be either detected or undetected at the destination. This is because of different factors which the signal encounters as it passes through the channel. Figure1 below shows the fundamental model of communication system.

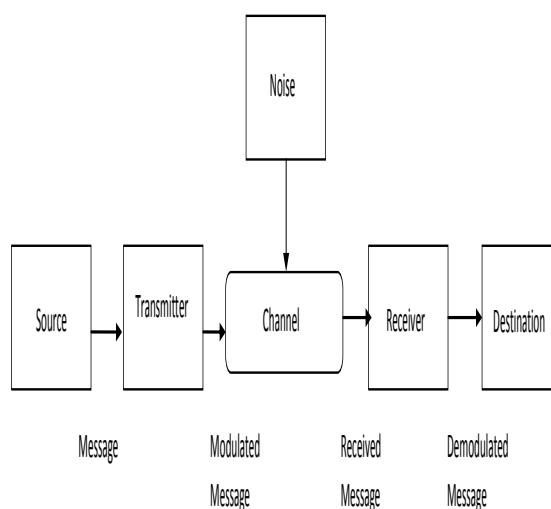


Figure 1: The Fundamental Model of Communication System

A. The Channel

Each type and nature of a communication channel determines to a great extent which kind of transmitted information that will be affected by noise or deteriorated in any communication system. Many non-wired systems display large characteristic gains when the receiver knows the channel properties. The degree of obtaining these gains relies on the receiver's ability to accurately evaluate the channel parameters [7]. The following channel types discussed below exists in the non-wired communication systems.

a) Raleigh fading

This models an environment where many objects that scatter the radio signal before it gets to the receiver exist and it is said to take care of the fading in a non-wired communication channel. The two kinds of fading include;

Long-term fading – This is order wise known as slow or log-normal fading exhibiting an envelope of fading information relating to the distance and received power.

Short-term fading – This can be called fast or Raleigh fading and it occurs basically as a result of reflection of transmitted signal. It can also be referred to the rapid fluctuations of a received signal that is set over an average value which changes gradually with the receiver movement.

The total information delivered to the receiver is composed of the product of Long-term and Short-term fading behaviours. The probability density function, Pr of the channel response R is given as

$$p_R(r) = \frac{2r}{\Omega} e^{-r^2/\Omega}, r \geq 0$$

Where $\Omega = E(R^2)$.

b) Gaussian Channel

This is basically employed in modelling the impairment produced at the receiving end for a perfect transmitting path. It is also called the AWGN channel. It describes very well and accurately many physical time-varying channels and also very useful in the provision of the system performance upper bound. It is assumed that the noise here possess a consistent Power Spectra Density, PSD across the channel bandwidth and a Gaussian Probability Density Function, PDF (though this rather is not responsible for fading).

In a situation where the user is stationed at a place and fading still occurs, then the channel can also be seen as Gaussian and the fading effects expressed as a path loss. The Gaussian channel due to its relevance for both cases of single receiver and transmitter have been given a considerable recognition in multiple access literature though its model did not adequately explain the numerous access channels in which many users transmit frequently [8]. The model which is of great desire here is one in which the active number of the network patronisers is a random variable and the Gaussian signal characterised on the basics that some users are transmitting. The additive noise channel of Figure 2 as shown below is said to be the simplest mathematical model for a communication channel.

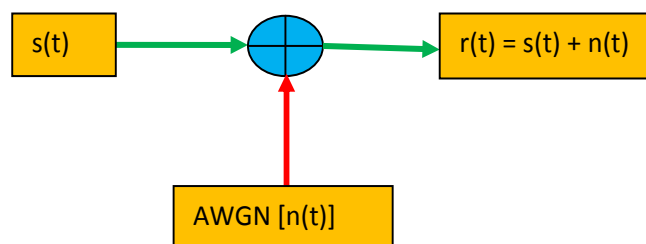


Figure 2: Additive noise channel

$s(t)$ which is the transmitted signal is corrupted by the additive random noise $n(t)$. The source of additive noise could be linked to the thermal noise innate in electronic components and the amplifiers at the receiver end or even arising from interference in the channel. The noise is characterised statistically as a Gaussian noise process and its mathematical model is therefore known as the additive Gaussian noise channel which can be represented mathematically as

$$r(t) = s(t) + n(t)$$

During the transmission process through the channel, the signal suffers attenuation given the received signal as

$$r(t) = \alpha s(t) + n(t)$$

Where α is the attenuation factor, $s(t)$ the transmitted signal and $n(t)$ additive random noise.

c) Ricean Channel

This is a theoretical model of a non-wired propagation phenomenon in which the prevailing pathway (usually the direct line-of-sight, LOS pathway) exists together with numerous other irregular pathways between the transmitter and receiver. The predominant pathway reduces the delay spread which in turn lowers the fading depth significantly thereby suggesting enough lower fading margins in designing the system.

B. Channel Coding

Channel coding for error detection and correction helps the communication system designers to reduce the effects of a noisy transmission channel.

The main aim of CC is to make available an efficient and reliable error correcting capability over an impairment channel, in the presence of a physically implementable decoder, to complement the success of J. A. Viterbi as proposed in 1967 [9], when he developed a decoding algorithm for convolutional codes. Though the Viterbi algorithm was relatively simple, it still met the criteria of exhibiting behaviour almost like that of a Maximum Likelihood Decoding (MLD) in practical decoders. Three types of CC exists which are Automatic Repeat request (ARQ), Forward Error Correction (FEC), and Hybrid Automatic Repeat Request (HARQ).

a) Automatic Repeat reQuest (ARQ)

This error correction technique combines error detection with the respect to retransmit an erroneous data. Here the received data block is checked for an existence of error in which if an error is indicated, the system will automatically request a

retransmission of the sent data. This process is continued unless the transmitted data is certified to be error free. Due to this repeated request, we can say that ARQ technique needs the presence of a feedback channel and used for transmitting non real-time data.

b) Forward Error Correction (FEC)

This channel coding technique does not require a feedback and the data to be transmitted is first coded using an error-correcting code prior to its transmission. This extra signal combined with the code is in-turn utilised at the receiving end to recover the transmitted original data. FEC have enjoyed a wide range of application with respect to error-control. The two main types of FEC code are Block codes and Convolutional codes.

i. Block codes:

They are suitable for burst errors and deal with large blocks of about a couple of hundred bytes. They are used in detecting errors only and tend to waste a high rate of bandwidth due to the addition of extra bits to the original transmitting data. Block codes could be classified as Hamming codes, Reed-Solomon codes, Cyclic codes etc.

If a block of 'k' signal bits are encoded to yield a code word of 'n' bits (such that 'n' is greater than 'k') therefore, for each 'k' information bit arrays, a unique code word of 'n' bits exists.

ii. Convolutional Codes:

Convolutional codes are linear codes over the field of one sided infinite sequences. Its usage is regularly seen in the correction of errors existing in a badly impaired channel due to their high affinity to error correction. These codes are recently majorly used in place of block codes when FEC is needed and have been registered to perform exceptionally well when run with Viterbi decoder which can be in the form of soft decision decoding or probabilistic decoding algorithm. The major difference between convolutional codes and block codes is that in block codes, the data sequence is first mapped out into individual blocks before it is encoded, whereas in convolutional codes, there is a direct mapping of that continuous information bit sequence to an encoder output bit. Figure1 below represents a generalized block diagram of a convolutional encoder.

c) Hybrid Automatic Repeat ReQuest (HARQ)

These error correcting codes according to Lin et al [10] employs the combination of both FEC and ARQ systems in order to attain an excellent reliability as exhibited by ARQ system and high throughput as also exhibited by FEC systems. With the HARQ bringing together these two properties of FEC and ARQ, the shortcomings experienced when either of the codes is used independently are overcome. Unlike ARQ system which discards signals that were previously received due to the presence of error, HARQ is rather suggested to improve the performance of the signal by adding all the received signals together so as to decode the message transmitted. Two types of improved HARQ exists which are HARQ with chase combining (HARQ-CC) and HARQ with incremental redundancy (HARQ-IR).

The major benefits of HARQ are that its implementation is quite easy and it also creates a good environment if trade-off is required between throughput and reliability. It is generally

applied in diverse systems including HSDPA, LTE, WiMAX and WLAN together with coding and adaptive modulation. According to Rappaport and Theodre [11], with reference to Figure 1, convolutional codes are generated by simply transmitting the data sequence across a finite shift register with 'N' k-bits stages and a linear algebraic function generator 'm'. A shift of data 'k' bits at any particular time in the shift register is recorded for every input of the data into the shift register and an output bit of 'n' bits is got for each 'k' bit user input.

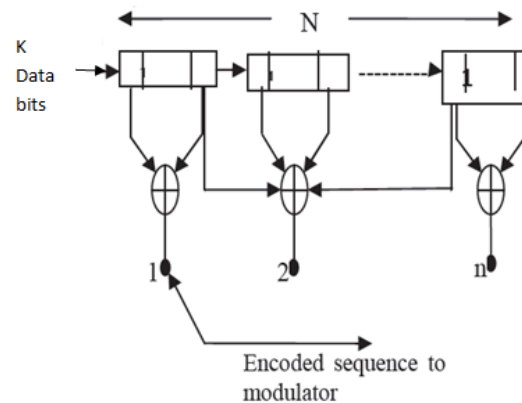


Figure 3: A generalised block diagram of Convolution encoder.

C. Applications of Convolutional Codes

a) WiMAX

This is a new wireless technology which is based on the air interface standard IEEE 802.16 WMAN designed just like the normal cellular network which uses a point to multipoint base station of configuration to provide a service of high throughput broadband connection covering a radius of over many kilometres. Though it exhibits a slightly higher BER at low SNR, its range and non-line of sight makes the system attractive as it has been configured to complement both mobile and fixed broadband applications. Its applications can be seen in many areas like high-speed enterprise connectivity, cellular backhaul to mention but a few based on its possession of high spectrum efficiency and reliability in multipath propagation. Many researches have been carried out for different coding stages in WiMAX and the system has been proved to depend on OFDMA PHY layer as specified by the IEEE 802.16 STD [12]. One major physical application of WiMAX in convolutional encoder is seen in SAL50300E product. This is a WiMAX compactible high-speed convolutional encoder which has a high speed convolutional encoder as specified in IEEE 802.16 – 2004. It has a basic coding constraint length 7, rate 1/2 transparent code suitable with channels of predominantly Gaussian noise.

b) Long Term Evolution (LTE)

This is the next generation network beyond 3G for mobile broadband standardization by 3GPP which is aimed at providing a capacity that will support demand for connecting from a new generation of consumers devices fashioned to a new mobile application.

c) 4G Networks

Advances in the network environment have added significantly to the creation of the present day society and with the internet as an integral part of the modern day society, it is hard to imagine industrial, social or daily life without it, yet the existing internet architecture is being stretched to its limits. According to recent statistics, internet traffic is growing at 40% annually and going by the basics of the current trend, it is expected to be a thousand times larger than what it is today in the late 2020. This implies that the power consumption of the Information and Communication Technology (ICT) will sky rocket.

D. Viterbi Decoding

This was first revealed in an IEEE transaction in 1967 [9] having been developed by Andrew J. Viterbi. It makes use of the Viterbi algorithm in decoding a bit stream which has been encoded with a convolutional code. It estimates actual bit sequence using trellis diagram. The decoder examines an entire received of a given length and computes a metric for each path and makes a decision based on this metric. All paths are followed until two paths converge on one node. Then the path with a higher metrics is kept, and the one with a lower metric is discarded. The paths selected are called the survivor. As it can be seen in Figure 3 below, our proposed block diagram of Viterbi decoder is made up of two major working blocks which are the Add and Compare Select (ACS) module and the Path Memory (PM) module with the former handling the Branch Metric calculations, Path Metric calculations and Add- Compare-Select, the latter keeps record and outputs the decoded information bits of the surviving path.

1. Branch Metric Calculation

Under this part of Viterbi algorithm, an assumption in its design has it that only two paths can lead to any other state of convolutional encoder. This is seen as a calculation of the difference in the input pair of bits and the other four obtainable ideal pairs which are; '00', '01', '10', '11'.

2. Path Metric Calculation

Under this part, a calculation of a metric is made for the pathway, with the least metric (survivor path) terminating in this state for every encoder. It computes the metrics for '2^{L-1}' path, choosing one of the paths as the optimal and stores the decisions result in the back tracing unit.

3. Back Tracing

This takes care of hardware implementation with minimal storage of information with respect to the pathway exhibiting a minimum metric path. It stores just a bit decision each time a minimum metric path is selected between the two. The trace back unit uses a covers direction to reinstate the maximum probability path from the result handed in by the path metric unit, making the Viterbi decoder to employ a trace back pattern of First-In-Last-Out (FILO) to recover the data. Figure4 below exhibits these above three discussed different parts of Viterbi algorithm.

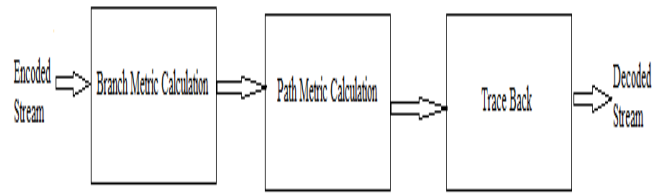


Figure 4: Viterbi Decoder Data Flow

IV. METHODOLOGY

Using the MATLAB software as required and employing the knowledge of analytical theory of the coding fundamental principles, the convolutional encoder and Viterbi decoder was modeled as shown in Figure5.

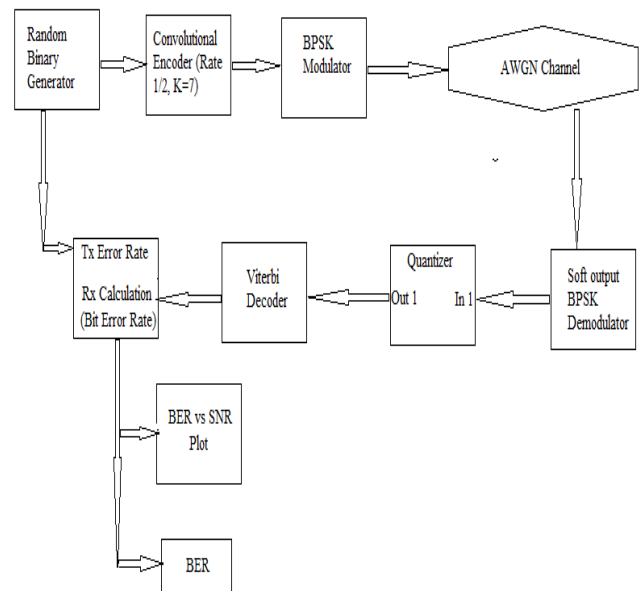


Fig-5: A communication system model block diagram exhibiting the Convolutional Encoder and Viterbi Decoder

A. Specifications for Modelling

- i. A binary convolutional encoder of rate 1/2 code, 6 memory storage units, constraint length K of 7
- ii. A soft input Viterbi decoder to take care of the convolutional encoder in 'T'.
- iii. BPSK modulation technique.
- iv. A binary random data generator as information production unit which should be able to hand in at least 5 million bits of information so as to account for a useful BER data.
- v. An SNR bit E_b/N_o of 0dB to 10dB.

V. RESULTS ANALYSIS AND DISCUSSIONS

In Figure 6 below, convolutional encoder data simulation was carried out on an input sequence of 1 million bits ranging from 0 to 10dB SNR values and 0.5 line spacing in other to obtain a good performance curve. As can be seen from the graph label, the curve of the convolutional encoder of rate 1/2 and K=7,

with Viterbi decoder using hard decision decoding of two-level quantization signals which is converted to only 'ones' and 'zeros' over an AWGN channel is marked with blue in the graph below, and subsequently, curves of 2-bits, 3-bits and 4-bits soft decision decoding are presented in the same Figure for comparison. The reference curve being the theoretical BER 'uncoded' is also present for use in the verification, comparison and analysis of the differences in the coding gain of the individual curves. From the hard decision decoding curve, the coding gain in SNR at a BER of 10^{-5} rated about 3.2dB presenting a decrease in the amount of transmit power up to a factor of 2 in comparison with the theoretical signal. We also observed that when soft decision decoding was implemented, which involved the quantization of signals into levels order than just 'zeros' and 'ones', the gain received increased which means that there was an improvement in the reduction of transmit power required.

When 2-bit soft decision decoding which involves 4-levels (00, 01, 10, and 11) of quantization was implemented, an additional gain of 1dB SNR at a BER of 10^{-5} was got in comparison to when hard decision decoding was used. This also improved to 2.1dB when 3-bit soft decision decoding was implemented and compared with that of hard decision decoding. On increasing to a 4-bit soft decision decoding, little or no significant change was recorded when compared with its 3-bit counterpart.

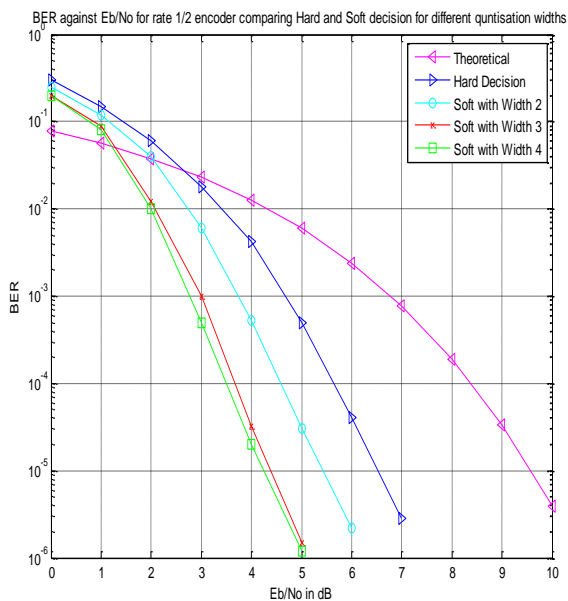


Figure 6: BER vs SNR curve of different quantization widths for rate 1/2 Binary

VI. CONCLUSION

This paper have carefully covered modeling of configurable rate convolutional encoder with Viterbi decoder from a mother code rate 1/2 and a constraint length 7 convolutional code. The whole system performance results were proved using some already established error performance bounds standard in which the achieved results exhibited a tighter upper bound for the model. The Convolutional encoders with Viterbi Decoder have proved to be a veritable tool for reducing the effects of noisy transmission channels.

REFERENCES

- [1]. Mike Rude, MMDS, Wireless Broadband Access and the Last Mile, A white paper, May 1, 2003.
- [2]. Keattisak Sripimanwat, Error Control Coding and its Applications, Lecture notes, Electrical Engineering, King's Mongkut Institute of Technology. 2005.
- [3]. W. Chen. (2006), 'RTL Implementation of Viterbi Decoder', MSc. Thesis, Linköping University
- [4]. B Sklar, Digital Communications, Prentice-Hall International Editions, 1988.
- [5]. Wikipedia, (2010) 4G [Available at: <http://en.wikipedia.org/wiki/4G>] [Viewed on November 20, 2014].
- [6]. The Future of Mobile Media Communication, (2010) 4G [Available at: <http://www.business-intracive.com>] (viewed on 05/11/2014 by 12:34pm).
- [7]. Benny Bing, "Broadband Wireless Access", Kluwer Academic Publishers, 2002.
- [8]. Abramson, N., Multiple Access Communications, IEEE Press, 1993.
- [9]. G. D. Forney, Jr., "The Viterbi algorithm," Proc. IEEE, vol. 61, no. 3, pp. 268-278, March 1973.
- [10]. Shu Lin, Daniel J. Costello, Jr., Michael J. Miller, "Automatic repeat request-error control schemes", IEEE communication magazine, December 1984, vol.22, No. 12, pp. 5-17
- [11]. Theodddore S. Rappaport "Wireless communications principle and practice", Second edition, Prentice Hall, 2002.
- [12]. A. Ebian, M. Shokair and K. H. Awadalla, 'Progress in Electronics Research C', Vol. 5, pp. 125-133, 2008.

ABOUT THE AUTHORS



Engr. (Mrs) Ifeoma U. Nnochiri is a lecturer in the department of computer Science, Michael Okpara University of Agriculture Umudike, Abia State, Nigeria. She holds a Bachelor degree (B.Eng) and Master's degree (M.Eng) in Computer Science & Engineering. More so, she is at the point of completing her Doctorate degree (PhD) in the same field. Her areas of interest include artificial intelligence, fuzzy logic and neural network, Embedded Systems etc