



Africa Research Journal

Research Journal of the South African Institute of Electrical Engineers
Incorporating the SAIEE Transactions

SAIEE AFRICA RESEARCH JOURNAL Vol 103 No. 4 pp 149-178

www.saiee.org.za

December 2012
Volume 103 No. 4

President:

Mr Mike Cary

Deputy President:

Mr Paul van Niekerk

Secretary and Head Office

Ms Gerda Geyer

South African Institute for Electrical Engineers (SAIEE)

PO Box 751253, Gardenview, 2047, South Africa

Tel: (27-11) 487-3003

Fax: (27-11) 487-3002

E-mail: researchjournal@saiee.org.za

Senior Vice President:

Dr Pat Naidoo

Junior Vice President:

Mr Andre Hoffmann

Immediate Past President:

Mr Andries Tshabalala

Honorary Vice President:

Mr Roland Hill

EDITORS AND REVIEWERS

EDITOR-IN-CHIEF

Prof. B.M. Lacquet, Faculty of Engineering and the Built Environment,
University of the Witwatersrand, Johannesburg, SA, beatrys.lacquet@wits.ac.za

MANAGING EDITOR

Prof. S. Sinha, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA, researchjournal@saiee.org.za

SPECIALIST EDITORS

Communications and Signal Processing:

Prof. L.P. Linde, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA

Prof. S. Maharaj, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA

Dr O. Holland, Centre for Telecommunications Research, London, UK

Prof. F. Takawira, School of Electrical and Information Engineering,
University of the Witwatersrand, Johannesburg, SA

Prof. A.J. Han Vinck, University of Duisburg-Essen, Germany

Dr E. Golovins, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA

Computer, Information Systems and Software Engineering:

Dr M. Weststrate, Newco Holdings, Pretoria, SA

Prof. A. van der Merwe, Department of Informatics, University of Pretoria, SA

Prof. E. Barnard, Meraka Institute, CSIR, Pretoria, SA

Prof. B. Dvornatzky, Joburg Centre for Software Engineering,
University of the Witwatersrand, Johannesburg, SA

Control and Automation:

Dr B. Yuskel, Advanced Technology R&D Centre,
Mitsubishi Electric Corporation, Japan

Prof. T. van Niekerk, Dept. of Mechatronics,
Nelson Mandela Metropolitan University, Port Elizabeth, SA

Electromagnetics and Antennas:

Prof. J.H. Cloete, Dept. of Electrical and Electronic Engineering,
Stellenbosch University, SA

Prof. T.J.O. Afullo, School of Electrical, Electronic and Computer Engineering,
University of KwaZulu-Natal, Durban, SA

Dr R. Geschke, Dept. of Electrical and Electronic Engineering,
Stellenbosch University, SA

Dr B. Jokanović, Institute of Physics, Belgrade, Serbia

Electron Devices and Circuits:

Dr M. Božanić, Azoteq (Pty) Ltd, Pretoria, SA

Prof. M. du Plessis, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA

Dr D. Foty, Gilgamesh Associates, LLC, Vermont, USA

Energy and Power Systems:

Prof. M. Delimar, Faculty of Electrical Engineering and Computing,
University of Zagreb, Croatia

Prof. S.P. Chowdhury, Department of Electrical Engineering,
University of Cape Town, SA

Engineering and Technology Management:

Prof. J.H. Pretorius, Faculty of Engineering and the Built Environment,
University of Johannesburg, SA

Prof. L. Pretorius, Dept. of Engineering and Technology, University of Pretoria, SA

Engineering in Medicine and Biology

Prof. J.J. Hanekom, Dept. of Electrical, Electronic & Computer Engineering,
University of Pretoria, SA

Prof. F. Rattay, Vienna University of Technology, Austria

Prof. B. Bonham, University of California, San Francisco, USA

General Topics / Editors-at-large:

Dr P.J. Cilliers, Hermanus Magnetic Observatory, Hermanus, SA

Prof. M.A. van Wyk, School of Electrical and Information Engineering,
University of the Witwatersrand, Johannesburg, SA

Prof. J.A. Ferreira, Electrical Power Processing Unit, Delft University of Technology,
The Netherlands

O. Flower, University of Warwick, UK

Prof. H.L. Hartnagel, Dept. of Electrical Engineering and Information Technology,
Technical University of Darmstadt, Germany

C.F. Landy, Engineering Systems Inc., USA

D.A. Marshall, ALSTOM T&D, France

Dr M.D. McCulloch, Dept. of Engineering Science, Oxford, UK

Prof. D.A. McNamara, University of Ottawa, Canada

M. Milner, Hugh MacMillan Rehabilitation Centre, Canada

Prof. A. Petroianu, Dept. of Electrical Engineering, University of Cape Town, SA

Prof. K.F. Poole, Holcombe Dept. of Electrical and Computer Engineering,
Clemson University, USA

Prof. J.P. Reyniers, Dept. of Electrical & Information Engineering,
University of the Witwatersrand, Johannesburg, SA

I.S. Shaw, University of Johannesburg, SA

H.W. van der Broeck, Phillips Forschungslabor Aachen, Germany

Prof. P.W. van der Walt, Stellenbosch University, SA

Prof. J.D. van Wyk, Dept. of Electrical and Computer Engineering, Virginia Tech, USA

R.T. Waters, UK

T.J. Williams, Purdue University, USA

Published by

SAIEE Publications (Pty) Ltd, PO Box 751253, Gardenview, 2047,

Tel. (27-11) 487-3003, Fax. (27-11) 487-3002,

E-mail: SAIEEPublications@saiee.org.za

Additional reviewers are approached as necessary

**ARTICLES SUBMITTED TO THE SAIEE AFRICA RESEARCH JOURNAL ARE FULLY PEER REVIEWED
PRIOR TO ACCEPTANCE FOR PUBLICATION**

The following organisations have listed SAIEE Africa Research Journal for abstraction purposes:

INSPEC (The Institution of Electrical Engineers, London); 'The Engineering Index' (Engineering Information Inc.)

Unless otherwise stated on the first page of a published paper, copyright in all materials appearing in this publication vests in the SAIEE. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, magnetic tape, mechanical photo copying, recording or otherwise without permission in writing from the SAIEE. Notwithstanding the foregoing, permission is not required to make abstracts on condition that a full reference to the source is shown. Single copies of any material in which the Institute holds copyright may be made for research or private use purposes without reference to the SAIEE.

This journal publishes research, survey and expository contributions in the field of electrical, electronics, computer, information and communications engineering. Articles may be of a theoretical or applied nature, must be novel and must not have been published elsewhere.

Nature of Articles

Two types of articles may be submitted:

- Papers: Presentation of significant research and development and/or novel applications in electrical, electronic, computer, information or communications engineering.
- Research and Development Notes: Brief technical contributions, technical comments on published papers or on electrical engineering topics.

All contributions are reviewed with the aid of appropriate reviewers. A slightly simplified review procedure is used in the case of Research and Development Notes, to minimize publication delays. No maximum length for a paper is prescribed. However, authors should keep in mind that a significant factor in the review of the manuscript will be its length relative to its content and clarity of writing. Membership of the SAIEE is not required.

Process for initial submission of manuscript

Preferred submission is by e-mail in electronic MS Word and PDF formats. PDF format files should be 'press optimised' and include all embedded fonts, diagrams etc. All diagrams to be in black and white (not colour). For printed submissions contact the Managing Editor. Submissions should be made to:

The Managing Editor, SAIEE Africa Research Journal,

PO Box 751253, Gardenview 2047, South Africa.

E-mail: researchjournal@saiee.org.za

These submissions will be used in the review process. Receipt will be acknowledged by the Editor-in-Chief and subsequently by the assigned Specialist Editor, who will further handle the paper and all correspondence pertaining to it. Once accepted for publication, you will be notified of acceptance and of any alterations necessary. You will then be requested to prepare and submit the final script. The initial paper should be structured as follows:

- TITLE in capitals, not underlined.
- Author name(s): First name(s) or initials, surname (without academic title or preposition 'by')
- Abstract, in single spacing, not exceeding 20 lines.
- List of references (references to published literature should be cited in the text using Arabic numerals in square brackets and arranged in numerical order in the List of References).
- Author(s) affiliation and postal address(es), and email address(es).
- Footnotes, if unavoidable, should be typed in single spacing.
- Authors must refer to the website: <http://www.saiee.org.za/arj> where detailed guidelines, including templates, are provided.

Format of the final manuscript

The final manuscript will be produced in a 'direct to plate' process. The assigned Specialist Editor will provide you with instructions for preparation of the final manuscript and required format, to be submitted directly to:

The Managing Editor, SAIEE Africa Research Journal, PO Box 751253, Gardenview 2047, South Africa.

E-mail: researchjournal@saiee.org.za

Page charges

A page charge of R200 per page will be charged to offset some of the expenses incurred in publishing the work. Detailed instructions will be sent to you once your manuscript has been accepted for publication.

Additional copies

An additional copy of the issue in which articles appear, will be provided free of charge to authors.

If the page charge is honoured the authors will also receive 10 free reprints without covers.

Copyright

Unless otherwise stated on the first page of a published paper, copyright in all contributions accepted for publication is vested in the SAIEE, from whom permission should be obtained for the publication of whole or part of such material.

VOL 103 No 4
December 2012

SAIEE Africa Research Journal



SAIEE AFRICA RESEARCH JOURNAL EDITORIAL STAFF IFC

Reverberation Chamber Low Frequency Field Uniformity Improvement with
Conducting Pyramidal Structures

by *M. Roman, R. R. van Zyl, H. C. Reader and J. Andriambeloson* 150

A Comparative Study on the Distance-Optimality for Distance-Preserving
Mappings Constructions with Cyclic-Shift Prefix Technique

by *K. Ouahada, L. Cheng and H. C. Ferreira* 155

Modified Earliest Decoding in Networks that Implement Random Linear Network
Coding

by *S. von Solms and A.S.J. Helberg* 165

Investigating the Effects of Subsidy on Network Resource Utilization in Multi-tier
Communities

by *M. Sumbwanyambe and A. L. Nel* 172

GBi electric
low voltage

 Eskom

REVERBERATION CHAMBER LOW FREQUENCY FIELD UNIFORMITY IMPROVEMENT WITH CONDUCTING PYRAMIDAL STRUCTURES

M. Roman*, R. R. van Zyl*, H. C. Reader** and J. Andriambeloson**

* *Department Electrical Engineering, Cape Peninsula University of Technology, 7535 Bellville, South Africa E – mail: vanzylr@cput.ac.za*

** *Department of Electrical and Electronic Engineering, University of Stellenbosch, 7600 Stellenbosch, South Africa E-mail: hcreader@sun.ac.za*

Abstract: Reverberation chambers are commonly used in industry for measuring electromagnetic interference because they are cost-effective. We attempt to stir modes more evenly inside reverberation chambers by introducing passive field uniformity enhancement (PFUE) structures. These structures augment the existing rotating stirrers. Measured results for a 2.445 x 2.475 x 3.72 m reverberation chamber are presented. The results show that the inclusion of passive structures improves the field uniformity in the low 300 - 400 MHz band. The standard deviation of the electric field distribution in this band is improved by up to 0.93 dB. The PFUE structures also reduced the dynamic range, in the worst case, by 2 dB at high frequencies.

Keywords: Field uniformity, lowest usable frequency, reverberation chamber, standard deviation.

1. INTRODUCTION

The proliferation of electronic systems leads to an increasingly contaminated electromagnetic (EM) environment. Electromagnetic compatibility (EMC) standards regulate radiation by these systems. Electronic equipment which radiates unwanted EM waves is reciprocally susceptible to electromagnetic interference (EMI) from other sources [1].

A widely-used facility to test EMI is a reverberation chamber (RC). RCs have gained much popularity in industry as a complement to well-established radiated interference sites and systems such as anechoic chambers (ACs), Gigahertz transverse electromagnetic cells (GTEM) and open area test sites (OATS). RCs are cost-effective and they function independently of weather conditions. EMI and EMS (susceptibility) measurements can therefore be conducted at any time; a high reproducibility rate of measurements is also obtainable [2].

An RC, also called a mode-stirred chamber, is an electrically large shielded cavity [3], [4]. It is usually equipped with one or several mechanical rotating stirrers which provide a statistically-isotropic field with a Rayleigh amplitude distribution when excited by an internal EM source [5], [6]. Dawson *et al.* [7] have described the optimisation of the stirrer shapes. In this paper, we investigate passive field uniformity enhancement (PFUE) structures in conjunction with the stirrers. The objective of using combined active stirring and PFUE structures is to obtain a better field uniformity at lower frequencies, which reduces the lowest usable frequency (LUF) for a particular RC.

An RC test setup consists of a receive antenna, stirrers, computer, spectrum analyser or EMI receiver, and the device-under-test (DUT) as shown in Figure 1. The transmit antenna is only used during chamber calibration. In an ideal RC, the fields are uniformly distributed, which implies that a high number of modes are excited within the cavity. The chamber's volume and stirrers play an important role in how many modes can be supported. The field strength distribution is considered adequately uniform if the distribution meets the 0 - 3 dB standard deviation limit stipulated by the IEC 61000-4-21 standard [8] for frequencies above 400 MHz (below 400 MHz the standard deviation limit increases). The size of an RC is usually determined by the working volume and the LUF that the end-user requires.

Two chambers are considered in this paper: a screened room at the Cape Peninsula University of Technology (CPUT), and a reverberation chamber at the University of Stellenbosch (US). The CPUT chamber is to be converted to an RC with the addition of rotating stirrers and PFUE structures. The US RC is slightly bigger than the CPUT chamber, and is being used for EMC measurements. PFUE structures were added to the chamber to determine the effect of these structures on the chamber's performance, especially at lower frequencies.

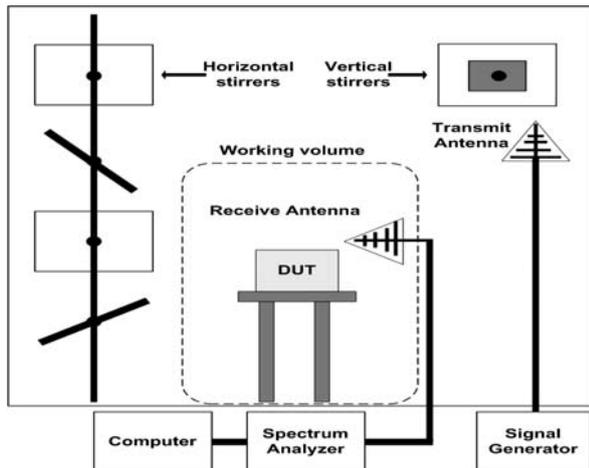


Figure 1: A test set-up in an RC having stirrers, DUT, antennas and a defined working volume (adapted from [3]).

2. FIELD UNIFORMITY ENHANCEMENT TECHNIQUES

2.1 Mode Stirrers

The use of rotating, conductive plates to ‘stir’ the electromagnetic modes inside an RC distributes the internal fields more evenly. There are also less commonly used alternative methods of mode stirring inside RCs, for example, frequency stirring, source stirring, or the use of a vibrating intrinsic RC [4], [9]. To obtain optimum field uniformity, the stirrers should have a complex shape and be electrically large, but a compromise between stirrer size and the usable working volume is necessary. Having more than one stirrer inside an RC is also beneficial. In a conventional RC, two techniques of stirring exist, namely that of mode-stirring and mode-tuned. In the mode-stirred technique, the stirrer rotates continuously and measurements are taken over a full revolution at a single frequency. In the mode-tuned technique, the stirrer is rotated in fixed steps and measurements are taken after every incremental rotation at a single frequency.

2.2 Field Uniformity

The field uniformity of an RC can be quantified by the standard deviation of the field strength distribution as sampled throughout the working volume inside the RC. The standard deviation σ of the field strength distribution along one axis, or combining the three axes measurements, is calculated, respectively, using the following formulae [10]:

$$\sigma_{\xi} = \sqrt{\frac{\sum_{i=1}^8 (\tilde{E}_{\xi,i} - \langle \tilde{E}_{\xi} \rangle)^2}{8 - 1}} \quad (1)$$

$$\sigma_{xyz} = \sqrt{\frac{\sum_{\xi=\{x,y,z\}} \sum_{i=1}^8 (\tilde{E}_{\xi,i} - \langle \tilde{E}_{xyz} \rangle)^2}{24 - 1}} \quad (2)$$

where $\langle \tilde{E}_{\xi} \rangle$ and $\langle \tilde{E}_{xyz} \rangle$ are the arithmetic means of the normalised maximum E-field measurements. The number 8 originates from field uniformity measurement methods and is the number of arbitrary locations used inside the RC at which the fields are measured. The standard deviation calculated in (2) is expressed in dB as [10]:

$$\sigma (dB) = 20 \log_{10} \left(\frac{\sigma_{xyz} + \langle \tilde{E}_{xyz} \rangle}{\langle \tilde{E}_{xyz} \rangle} \right) \quad (3)$$

The IEC 61000-4-21 standard requires that standard deviation falls within the range of 0 - 3 dB above 400 MHz.

2.3 Lowest Usable Frequency (LUF)

The LUF is directly related to the RC size, hence, a larger chamber leads to a lower LUF. The LUF corresponds to the lowest frequency where the chamber conforms to the IEC 61000-4-21 standard. One definition states that the LUF is 3 to 6 times the lowest resonant frequency of the chamber [11].

3. PASSIVE MODE-STIRRER STRUCTURES

The objective of incorporating passive mode stirrer structures in a conventional RC is to improve field uniformity for a lower LUF. This is especially relevant to small chambers, such as the chamber at CPUT which has dimensions $1.863 \times 2.443 \times 2.473$ m (x y z respectively). PFUE structures that were investigated through simulation and practical measurement are line, groove and cone structures, as depicted in Figures 2, 3 and 4, respectively. These structures were placed at different positions throughout the chamber and also in different combinations, to determine the optimum placement, whilst considering the practicalities of the actual chamber. The maximum protrusion of the PFUE structures is related to the LUF that is required. In the current investigation, these protrusions were one wavelength at the LUF, which is required to be 1 GHz for the planned CPUT RC.

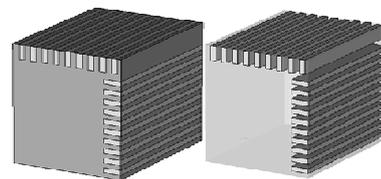


Figure 2: Cavity fitted with line structures on the ceiling and on the side walls.

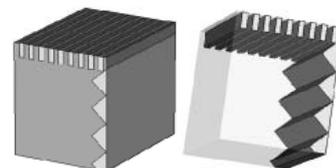


Figure 3: Cavity fitted with line structures on the ceiling and groove structures on the side walls.

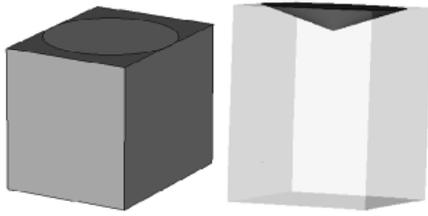


Figure 4: Cavity fitted with a cone structure on the ceiling.

4. SIMULATION SET-UP

Electromagnetic characterisation of the CPUT chamber was done through CST simulations [12]. Matched Yagi-Uda antennas were used as excitation sources. They were designed for each of the specified frequencies to ensure that the radiated power was the same over the frequency band. Simulations with the Yagi-Uda antenna took, on average, one to three days to complete. The antennas were orientated in both the horizontal and vertical to mimic real-life measurement procedures. The simulated frequencies were chosen to be harmonically independent. The simulated frequency range was 1 - 2 GHz for the CPUT RC.

During the simulation the Perfect Electric Conductor (PEC) was chosen as the background material in all directions. The empty cavity was constructed of vacuum material and the antenna, PFUE structures and stirrers were made of aluminium. The antenna excitation was an S-parameter discrete edge port with an impedance of 50Ω . The Fast Perfect Boundary Approximation (FPBA), which is a hexahedral meshing option, was chosen as it reduces the number of mesh cells normally used by a factor of three when compared to the PBA option. A high performance PC with the following specification was used: Intel (R) Xeon CPU E5620@2.40GHz (2 processors) with a 1.5 TB hard drive and 48 GB DDR3 memory.

5. SIMULATION RESULTS

Simulations were based on the CPUT chamber. A working volume of 1 m^3 was chosen, defined within 0.5 to 1.5 m in the three axes, relative to the one corner. The standard deviation of the electric field strength distribution inside the CPUT chamber was simulated for each of the three PFUE structures. The vertical and horizontal stirrers were rotated simultaneously in steps of 30° angles. This leads to 12 simulations for the cavity fitted with stirrers only as well as 12 simulations for the cavity fitted with stirrers and passive structures.

The simulation results are presented in Figure 5. It is clear that the groove structures, placed on the ceiling of the chamber, exhibit the lowest standard deviation of the four scenarios over the frequency band. The improvement in standard deviation is marginal, showing a reduction of about 0.4 dB across the band when compared to the empty cavity where no mode stirring takes place. Hence,

these passive structures alone do not ensure compliance with the IEC 61000-4-21 standard. However, this does raise the plausible question that, since they improve field uniformity, will the use of passive structures in addition to active stirring with rotating stirrer plates further improve the field uniformity? If this is true, a lower LUF for the same RC may be realised. This hypothesis was tested through measurement of the US RC.

6. MEASUREMENT RESULTS

The structure with the best performance in the CPUT RC, namely the groove structure, was adapted to the US RC since, being bigger, is operating at a lower frequency range than the envisaged CPUT RC. The groove structures protrusions were redesigned to be 0.47 m, which is one wavelength at 646 MHz as this is the lowest frequency at which this chamber had been experimentally characterised by Wiid [3]. Before the actual grooves were designed, it was decided to replace the groove structures with pyramid structures. The notion was that the pyramid structures could scatter the fields even more inside the RC since they have groove-like protrusions for both horizontal and vertical positions.

The 8 pyramid structures were manufactured from galvanised steel since aluminium would be prohibitively expensive. Simulations verified that the use of galvanised steel does not affect the field uniformity as compared with using aluminium, but the loss factor of the chamber may increase.

The structures were intended to be placed on the ceiling but this was not feasible in the US chamber. It was therefore decided to place these passive pyramid structures underneath the horizontal stirrer as this space was not being utilised, and did not interfere with the original working volume of the RC. The placement of the pyramidal PFUE structures in relation to the existing mode stirrers is shown in Figures 6 and 7.

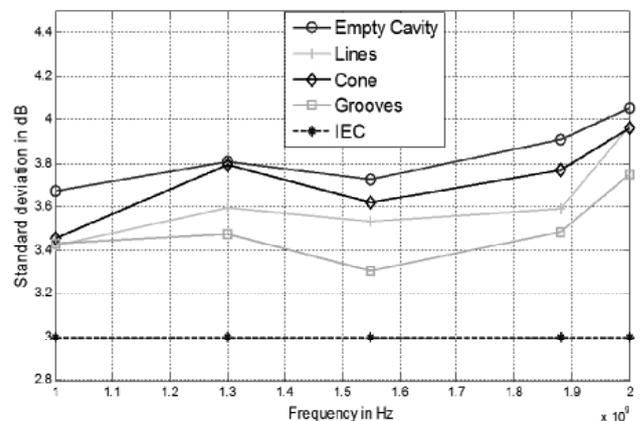


Figure 5: Simulated standard deviation of the electric field distribution within the CPUT chamber for the empty cavity, and the inclusion of the PFUE structures (cone, line and groove) for a vertically positioned dipole antenna.

The measured results indicating field uniformity are presented in Figures 8 and 9, showing the standard deviation of the two scenarios, compared to the IEC limit. Figure 10 shows the chamber calibration factor (CCF) which compares the induced losses for the two scenarios.

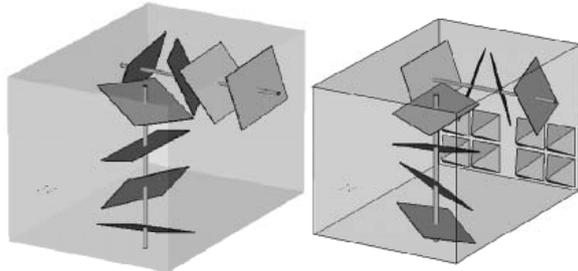


Figure 6: US RC fitted with stirrers only and with eight pyramid structures with dimensions: 0.47 x 0.47 x 0.47 m.

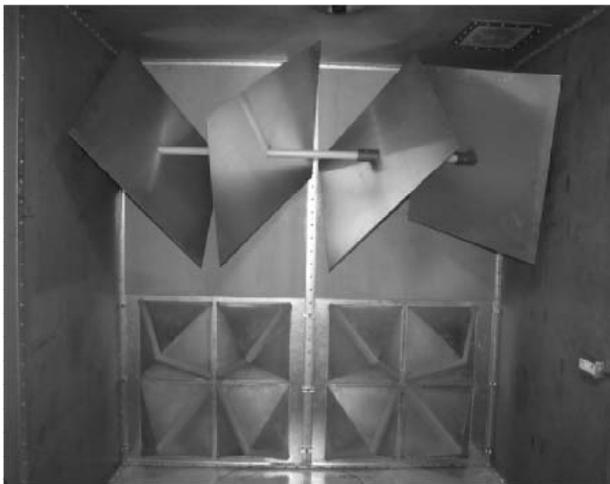


Figure 7: Photograph of the US RC fitted with eight pyramidal PFUE structures below the horizontal stirrer.

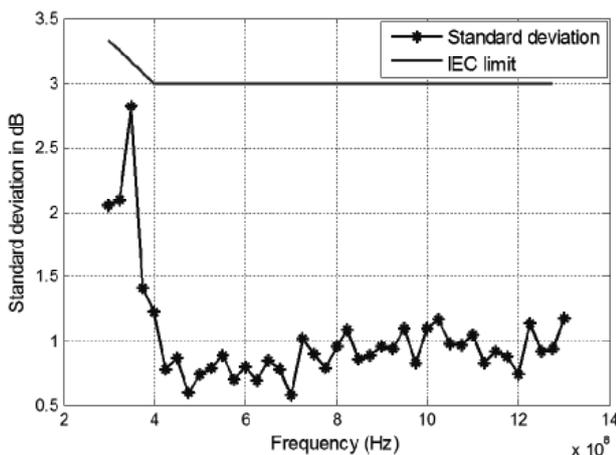


Figure 8: Measured standard deviation of the electric field for the US RC fitted with stirrers only for 3 independent data samples.

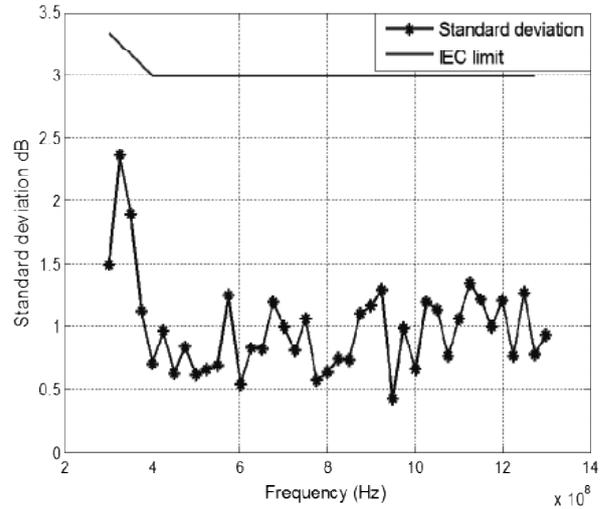


Figure 9: Measured standard deviation of the electric field for the US RC fitted with stirrers and pyramid structures.

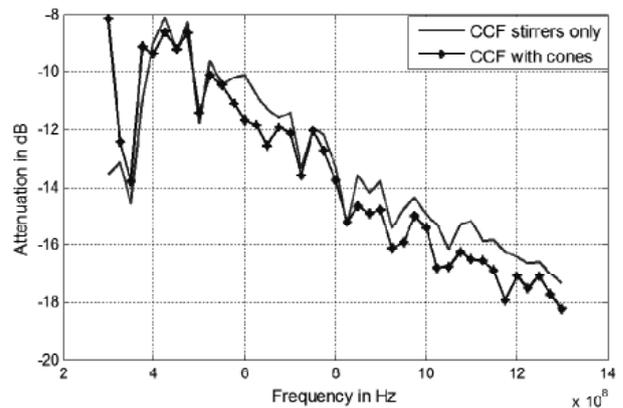


Figure 10: Comparing the measured chamber calibration factor of the RC with stirrers only against the RC with stirrers and pyramid structures.

From the measured results in Figures 8 and 9 it is clear that the combination of passive and active stirring mechanisms improves the field uniformity by up to 0.93 dB over the frequency range of 300 - 400 MHz. Figure 10 shows that for the worst case, there is up to 2 dB additional loss at high frequencies when compared to the stirrers only case. Loss is expected due to the loading effect of RCs. This loss will, however, not adversely affect the RC's performance as it can be incorporated in the calibration of the chamber.

7. CONCLUSION

This paper principally investigated the addition of PFUE techniques in active RCs through simulation and measurement. The PFUE structures are shown to yield better field uniformity at the lower end of the chamber frequency range. Dawson et al. [7] achieved similar field uniformity with optimised stirrer geometry. It would be of interest to combine both stirrer geometry optimisation and PFUE inclusion to push the lower usable band limit of reverberation chambers.

Based on our measured results, with a limited PFUE inclusion, a valuable improvement of up to 0.93 dB in standard deviation was achieved in the low frequency range of 300 - 400 MHz. However, no improvement in field uniformity was noticed at higher frequency ranges. These structures are also cost-effective. We expect further gains by incorporating these PFUE structures on the ceiling and other available wall space.

8. REFERENCES

- [1] R. De Leo, & V. Primiani. "Radiated immunity tests: reverberation chamber versus anechoic chamber results". *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 4, pp. 1169–1174, August 2006.
- [2] H. Streitwolf, R. Heinrich, H. Behnke, L. Dallwitz & U Karsten. "Comparison of radiated immunity tests in different EMC test facilities". *18th International. Zurich Symposium on EMC*, pp. 229-232, September 2007.
- [3] P.G. Wiid. "Investigating Cost-Effective EMC Methods". University of Stellenbosch. 2005.
- [4] E. Voges & T. Eisenburger. "Electrical Mode Stirring in Reverberation Chambers by Reactively Loaded Antennas". *IEEE Transactions on Electromagnetic Compatibility*, vol. 49, no. 4, pp. 756-761, November 2007,
- [5] D.A. Hill. "Boundary Fields in Reverberation Chambers" *IEEE Transactions on Electromagnetic Compatibility*, vol. 47, no. 2, pp. 281-290, May 2005.
- [6] J. Sanchez-Heredia, M. Garcia-Fernandez, M. Gruden, P. Hallbjorn, A. Rydberg, & D. Sanchez-Hernandez. "Arbitrary fading emulation using mode-stirred reverberation chambers with stochastic sample handling". *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, pp. 152–154, April 2011.
- [7] J. F. Dawson, M. O. Hatfield, L. Arnaut, N. Eulig. "Reverberation (Mode-stirred) chambers for electromagnetic compatibility". www.compliance-club.com/archive/old_archive/030530.htm.
- [8] *IEC 61000-4-21*, Electromagnetic Compatibility (EMC)—Part 4-21: Testing and measurement techniques—Reverberation chamber test methods, 2003.
- [9] J. Skrzypczynski. "On Shielding Effectiveness Measurements Using Dual Vibrating Intrinsic Reverberation Chamber". *2010 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 112 – 116, September 2010.
- [10] C. Bruns. "Three dimensional simulation and experimental verification of a reverberation chamber". Zurich: Universitat Frediriciana Karlsruhe. 2005.
- [11] L. Xiaoqiang, W. Guanghui, Z. Yongqiang & Z. Chenghui. "Effects of Stirrer on the field uniformity at Low Frequency in a Reverberation Chamber and its Simulation". *International Symposium on Computer Science and Computational Technology (ISCST '08)*, vol. 2, pp. 517 - 519. December 2008.
- [12] CST Computer Simulation Technology. CST MICROWAVE STUDIO – Transient Solver. <http://www.cst.com/Content/Products/MWS/TransientSolver.aspx>.

A COMPARATIVE STUDY ON THE DISTANCE-OPTIMALITY FOR DISTANCE-PRESERVING MAPPINGS CONSTRUCTIONS WITH CYCLIC-SHIFT PREFIX TECHNIQUE

K. Ouahada, ‡ L. Cheng and H. C. Ferreira *

* *Department of Electrical and Electronic Engineering Science, University of Johannesburg, South Africa. Email: {kouahada,hcferreira}@uj.ac.za.*

‡ *School of Electrical and Information Engineering Science, University of the Witwatersrand, Johannesburg, South Africa. Email: ling.cheng@wits.ac.za.*

Abstract: This paper presents a comparative study on the distance optimality for Distance-Preserving Mappings codes generated from different constructions. Few constructions are considered to investigate the reasons behind making codes optimum when they reach the upper bound on the sum of the Hamming distances for certain lengths of the permutation sequences. The technique of cyclic-shift prefix or suffix of permutation symbols is used to investigate the conditions making distance-preserving mappings codes optimum.

Key words: Distance-preserving mappings, Distance optimality, Cyclic codes, Suffix/Prefix.

1. INTRODUCTION

A block code, where a cyclic shift of every codeword yields another codeword belonging to the same block code is called a cyclic code [1]. Gilbert [2] defined a cyclic-permutable code as a binary block code of block length n such that each codeword has a cyclic order n and the codewords are cyclically distinct. Later Maracle and Wolverson [3] proposed an efficient algorithm for generating these codes [4], where all these codes are generated from binary block codes.

Using the technique of a cyclic shift of a symbol in a permutation sequence, we aim in this paper to study the reasons and conditions behind any distance-preserving mapping constructions [5–8] that can reach the upper bound on the sum of the Hamming distances.

Since we are interested in the study of distance-preserving mapping codes [9] using the technique of cyclic shift of codewords symbols, we will focus in the paper only on the resultant codes from the mapping called as non-binary cyclic shift block codes.

The property of the cyclic shift symbols of a permutation sequence is based on the shift of the symbols in different positions in such a way that we create codewords from one another with different sequences. Thus the Hamming distance between the two codewords reaches the maximum.

Using this property we try to understand and analyze the reasons and conditions that can lead to the upper bound on the sum of the Hamming distances for a permutation block code. Another objective is the investigation of the conditions under which any cyclic shift of symbols in a codeword does not result in repetitive codewords in the same codebook. Hence the use of the prefix/suffix technique will help solving such a problem as it will be detailed in the following sections.

In Section 2 a few definitions [5, 11] used in this paper are presented. In Section 3 we present briefly distance-preserving mappings. Section 4 presents a few examples of cyclic shift codes and the distance effect. Section 5 presents the distance optimality for DPMs. Section 6 presents the technique of cyclic shift of a prefix or a suffix [10]. We conclude with some final remarks in Section 7.

2. PRELIMINARY

Definition 1 A linear code C is an h -dimensional vector sub-space of \mathbb{Z}_q^n , $0 \leq h \leq n$. The elements of a code are called codewords, n is called the length and h is called the dimension, where,

- $q \in \mathbb{N}$ such that $q = p^h$, p prime. We denote by \mathbb{F}_q the field with q elements.
- \mathbb{F}_q^n forms an n -dimensional vector space over \mathbb{F}_q .

Definition 2 An $\mathbb{F}_q[n, h, d]$ code (linear) C is called cyclic if any cyclic shift of a codeword of the set is also a codeword of the set, i.e. $(c_0, \dots, c_{n-1}) \in C \Rightarrow (c_{n-1}, c_0, \dots, c_{n-2}) \in C$. We present d as the minimum distance.

Definition 3 A binary code, C_b , consists of $|C_b|$ words of length n , where every word contains zeros and ones as symbols.

Definition 4 A permutation code, C_p , consists of $|C_p|$ words of length M , where every word contains the M different integers $1, 2, \dots, M$ as symbols.

Definition 5 The symmetric group, S_M , consists of $|S_M| = M!$ words obtained by permuting the symbols $1, 2, \dots, M$ in all the possible ways.

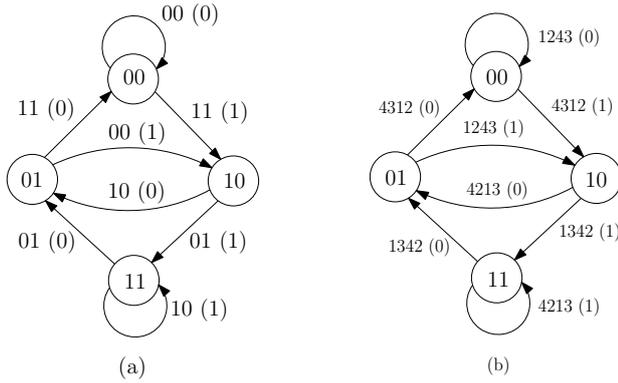


Figure 1: State systems for (a) convolutional base code and (b) permutation trellis code

Definition 6 The Hamming distance $d_H(\mathbf{y}_i, \mathbf{y}_j)$ is defined as the number of positions in which the two sequences \mathbf{y}_i and \mathbf{y}_j differ. The sequences \mathbf{y}_i and \mathbf{y}_j are considered to have same lengths.

3. DISTANCE-PRESERVING MAPPING CODES

3.1 Introduction

Distance-preserving mappings map the outputs of a convolutional code to other codewords from a permutation codebook. The code obtained will have greater error correction capability than the codebook itself.

Ferreira *et al.* [9] made use of distance-preserving mappings from binary sequences to permutation sequences to construct permutation trellis codes by mapping the binary convolutional output sequences to permutation sequences and also made use of the maximum-likelihood Viterbi algorithm for decoding [9, 12, 13]. They have combined the permutation trellis codes with M -FSK modulation for certain applications in power-line communication [13, 14]. Several subsequent studies have been completed in this research area regarding the combination between coding and modulation schemes [14–17].

The mappings considered are from all possible binary sequences of length n in a binary set denoted C_b , with $|C_b| = 2^n$, to a subset of S_M denoted by C_p , where we have $|C_p| = |C_b|$.

Emphasizing the mapping technique in a better way, we present an example, where we use the convolutional code with half rate and constraint length $K = 3$ [18] as a base code. The output of the encoder, which is a set of binary 2-tuple code symbols, can be mapped to a set of permutation M -tuples. The corresponding state systems appear in Fig. 1. Note that in general the information transmission rate of the resulting permutation trellis coded scheme will be bits per channel use.

Applying the definition of the Hamming distance to binary and non-binary sequences, we can denote by $\mathbf{D} = [\mathbf{d}_{ij}]$ the

distance matrix whose entries are the Hamming distances between two binary sequences \mathbf{x}_i and \mathbf{x}_j defined as follows:

$$\mathbf{D} = [\mathbf{d}_{ij}] \text{ with } \mathbf{d}_{ij} = \mathbf{d}_H(\mathbf{x}_i, \mathbf{x}_j). \quad (1)$$

Similarly for permutation sequences we denote by $\mathbf{E} = [\mathbf{e}_{ij}]$ the distance matrix defined as follows:

$$\mathbf{E} = [\mathbf{e}_{ij}] \text{ with } \mathbf{e}_{ij} = \mathbf{d}_H(\mathbf{y}_i, \mathbf{y}_j). \quad (2)$$

The sum of all the distances in \mathbf{E} , which is denoted by $|\mathbf{E}|$ plays a role in the error correcting capabilities, as was shown in [5].

Example 1 Applying (1) and (2) for the mapping presented in Fig. 1, the distance metrics from each state machine could be presented in a matrix form as

$$\mathbf{D} = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \text{ and } \mathbf{E} = \begin{bmatrix} 0 & 2 & 2 & 4 \\ 2 & 0 & 4 & 2 \\ 2 & 4 & 0 & 2 \\ 4 & 2 & 2 & 0 \end{bmatrix}. \quad (3)$$

Taking into consideration the fact that the entries on the main diagonals are all zeros, we have $e_{ij} \geq d_{ij} + \delta$, with $\delta \geq 1$ for all $i \neq j$. The mapping of the outputs of the base code $\{00, 01, 10, 11\}$ to the permutation set $\{1243, 1342, 4213, 4312\}$, guarantees an increase of at least one unit of distance per step between any two unremerged paths in the trellis diagram of the resulting permutation trellis code, when comparing it to the base code.

For the base code, the shortest re-merging paths in the trellis diagram, which are known to determine the free distance [18], have different distances between each pair of branches. These branch distances have been changed with the code obtained after the mapping. It can be seen that the distances have increased, which makes our example represent a distance-increasing mapping.

In general, three types of DPMs [19] can be defined, depending on how the Hamming distance is preserved.

Definition 7 A *distance-conserving mapping* (DCM) is a distance-preserving mapping that only guarantees conservation of the base code's minimum distance. Hence at least one $e_{ij} = d_{ij}$ while for all other i and j , $e_{ij} \geq d_{ij}$, $i \neq j$.

Definition 8 A *distance-increasing mapping* (DIM) is a distance-preserving mapping that guarantees that the resulting code's distance will have some increase above the base code's minimum distance. Hence, $e_{ij} \geq d_{ij} + \delta$ for $i \neq j$, where $\delta \in \{1, 2, \dots\}$, depending on the increase required.

Definition 9 A *distance-reducing mapping* (DRM) is a distance-preserving mapping that yields a mapping of which the distance is lower than that of the base code,

although this decrease is controlled. The decrease is guaranteed to be not more than a fixed amount, determined by the length of the shortest merging paths in the trellis diagram, with $e_{ij} \geq d_{ij} + \delta$, $\delta \in \{-1, -2, \dots\}$ and $i \neq j$.

In Example 1, we have $|\mathbf{E}| = 32$. The value of $|\mathbf{E}|$ depends on the choice of the permutation set on which the construction is based. As we have mentioned before, the Hamming distance contribute in the improvement of the error correcting codes. Thus any construction that contribute in higher values of $|\mathbf{E}|$, is expected to perform better than others in a channel coding scheme. An upper bound on the sum of the Hamming distances was introduced by Swart and Ferreira [5] as a means to compare between different constructions. The upper bound is defined as follows:

$$|\mathbf{E}_{\max}| = \mathbf{M}[2^{2^n} - (2\alpha\beta + \beta + \alpha^2\mathbf{M})], \quad (4)$$

where $\alpha = \lfloor 2^n / M \rfloor$ and $\beta \equiv 2^n \pmod{M}$.

We have $|\mathbf{E}_{\max}| = 48$. Thus we can say that our mapping in that case did not reach the upper bound on the sum of the Hamming distances since we have $|\mathbf{E}| < |\mathbf{E}_{\max}|$. To compare any distance sum of a mapping to that of the upper bound, a distance optimality [5] parameter is defined as follows:

Definition 10 The distance optimality of a mapping is given by

$$\eta = \frac{|\mathbf{E}|}{|\mathbf{E}_{\max}|}, \quad (5)$$

where $|\mathbf{E}|$ is the sum of the Hamming distances for the mapping and $|\mathbf{E}_{\max}|$ is the upper bound on the sum of the Hamming distances for the mapping.

In the case of the mapping in Example 1, we have $\eta = 0,67$.

3.2 DPMs Constructions

Here, a brief summary of some DPMs constructions is presented, in order to show the differences between them.

Prefix construction: Ferreira and Vinck [14] presented a construction based on the use of prefix where permutation sequences of length M can be used to create an $M + 1$ mapping. The idea can be briefly explained as follows.

A set of binary $(n + 1)$ -tuples can be ordered following normal lexicography, i.e. setting up the standard table of $(n + 1)$ -bit binary numbers. Note that the first 2^n $(n + 1)$ -bit binary numbers are obtained by prefixing the set of n -bit binary numbers with a most significant bit 0, and the second 2^n $(n + 1)$ -bit binary numbers by prefixing the set of binary n -bit binary numbers with a most significant bit 1. This table can thus be partitioned according to the prefix bit into two subsets. Within each subset, the intra-distance between elements is determined by the $n \times n$ \mathbf{D} matrix, and

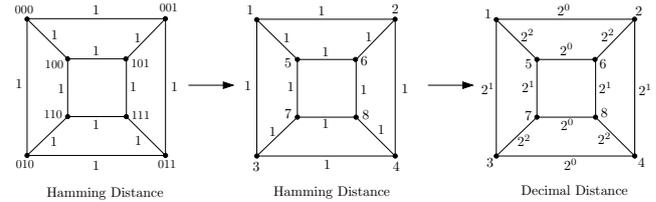


Figure 2: Distance presentation of a 3-cube

stays the same. However, the binary prefixes of 0 and 1 account for an additional one unit of distance between two elements from the two different subsets.

Chang construction: Chang *et al* [20] generalized the prefix construction to all permutation sequences of length M . They have presented two recursive constructions, making use of two mappings, M_1 and M_2 . Then they were combined to create a mapping for $M_1 + M_2$. The other, being similar to that of Ferreira and Vinck's prefix method, made use of an M mapping to create a new one mapping for $M + 1$. This technique can be considered as a generalization of the prefix technique proposed by Ferreira and Vinck.

k-Cube Graph Construction: The idea of the k -cube graph construction, denoted by $Q(M, n, \delta)$, is based on the concept of the cube distance. Fig. 2 shows different ways of presenting distances in a cube. The Hamming distance is not considered in this construction since for binary sequences or integer sequences the Hamming distances are the same. Using the absolute value of the differences between the decimal values of the edges, we found that these distances are in a geometric series of ratio 2. The distances are from 2^0 till 2^{k-1} . Fig. 2 shows that when using the cube distance approach, all parallel edges are equidistant.

We can summarize the properties of a k -cube based on the cube distance as follows:

1. $k = \lceil \log_2 M \rceil$, where M is the length of the permutation sequence;
2. Geometric series with a ratio of 2, from 2^0 till 2^{k-1} ;
3. Any vertex in a k -cube graph has a degree of k and all its distances are in a geometric series from 2^0 till 2^{k-1} ;
4. Parallel edges are equidistant. Diagonal edges are considered as parallel to one another;
5. Distances in a k -cube increase when k increases;
6. $(k - 1)$ -cube graph is a subgraph of the k -cube graph.

The concept of the k -cube construction is based on the grouping of all pairs of symbols having same decimal

Table 1: Comparison of distances for different conserving mappings

M	$ E_{\max} $	Prefix construction [14]	Chang construction 2 [20]	Cube graph construction
4	768	732	768	768
5	4090	3616	3712	3712
6	20472	17072	17536	19456
7	98294	78528	81024	94208
8	458752	355840	367744	458752

distance and the swapping will be based on the these distance groupings.

Tables 1 and 2 present a few results of the sum on the Hamming distances and certain codebooks of the constructed distance preserving mapping codes.

4. CYCLIC-SHIFT PREFIX/SUFFIX PERMUTATION CODES

It is known that cyclic shifts of codewords play an important role in coding theory. We try to present in this section the importance of this technique in the increase of the sum on the Hamming distances.

Definition 11 We denote by F_{sh} the cyclic-shift function as presented in the following:

$$F_{sh} : C_M \rightarrow C_M$$

$$\mathbf{y}^0 = (y_1, y_2, \dots, y_M) \rightarrow \mathbf{y}^1 = \mathbf{F}_{sh}(\mathbf{y}_M) = (y_2, \dots, y_M, y_1)$$

The definition is true for any $M-1$ cyclic shift of the vectors elements. We denote by \mathbf{y}^i the i -times shift of symbols of the vector \mathbf{y}^0 and by M_M the matrix with all elements from C_M , which has the following form:

$$M_M = \begin{matrix} & \mathbf{y}^0 & \mathbf{y}^1 & \mathbf{y}^2 & \dots & \mathbf{y}^{(M-1)} \\ \mathbf{y}^0 & \left[\begin{array}{cccccc} y_1 & y_2 & y_3 & \dots & y_M \\ y_2 & y_3 & y_4 & \dots & y_1 \\ y_3 & y_4 & y_5 & \dots & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_M & y_1 & y_2 & \dots & y_{M-1} \end{array} \right] \\ \mathbf{y}^1 & & & & & \\ \mathbf{y}^2 & & & & & \\ \vdots & & & & & \\ \mathbf{y}^{(M-1)} & & & & & \end{matrix},$$

where each row is the result of a single left-cyclic shift of the row above it, and the column and the row outside the brackets represent the vectors \mathbf{y}^i with $0 \leq i \leq M-1$.

Definition 12 A cyclic permutation code is a permutation code in which a cyclic shift of any codeword is also a codeword.

Proposition 1 For any cyclic code C_M of length M which is a subset of S_M and consists of M codewords, we have $|S_M| = (M-1)!|C_M|$.

Proof: from Definition 5, we have $|S_M| = M! = (M-1)!M$, thus $|S_M| = (M-1)!|C_M|$.

In general, a code C_M of length M is called a cyclic code when it is invariant under any cyclic shift operation defined as follows:

Definition 13 We denote by C_M^t of length M , with $1 \leq t \leq n$, the t -th subset of S_M in a series of n different cyclic permutation codes C_M . The corresponding matrix will be denoted as M_M^t .

Proposition 2 The union of n cyclic codes C_M^t of length M , with $1 \leq t \leq n$ is a subset of S_M .

Proof: We know that $|C_M^t| = M$ and $|S_M| = (M-1)!|C_M|$. Then the union of n different cyclic codes C_M^t of length M , with $1 \leq t \leq n$ is a subset of S_M .

Proposition 3 The upper bound on the sum of the Hamming distances for any set C_M is $M^2(M-1)$.

Proof: As defined previously, \mathbf{y}^i is the i -th times shift of all elements of the vector \mathbf{y} .

$$\mathbf{E}_{M_M} = \begin{matrix} & \mathbf{y}^0 & \mathbf{y}^1 & \mathbf{y}^2 & \dots & \mathbf{y}^{(M-1)} \\ \mathbf{y}^0 & \left[\begin{array}{cccccc} 0 & M & M & \dots & M \\ M & 0 & M & \dots & M \\ M & M & 0 & \dots & M \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M & M & M & \dots & 0 \end{array} \right] \\ \mathbf{y}^1 & & & & & \\ \mathbf{y}^2 & & & & & \\ \vdots & & & & & \\ \mathbf{y}^{(M-1)} & & & & & \end{matrix}$$

\mathbf{E}_{M_M} is a $M \times M$ matrix, with all the Hamming distances $d_H(\mathbf{y}_i, \mathbf{y}_j)$. The first diagonal has M values of zeros corresponding to $d_H(\mathbf{y}_i, \mathbf{y}_j) = 0$, when $\mathbf{y}_i = \mathbf{y}_j$.

$$\begin{aligned} |\mathbf{E}_{M_M}| &= \sum (d_H(\mathbf{y}_i, \mathbf{y}_j)) \\ &= M \times M^2 - M \times M \\ &= M^2(M-1) \end{aligned}$$

Example 2 We take the case of $M = 3$, with the codebook $C_3 = \{123, 231, 312\}$, where all its codewords are cyclically shifted. The sum on the distances is then,

$$\mathbf{E}_{M_3} = \begin{matrix} & 123 & 231 & 312 \\ 123 & \left[\begin{array}{ccc} 0 & 3 & 3 \\ 3 & 0 & 3 \\ 3 & 3 & 0 \end{array} \right] \\ 231 & & & \\ 312 & & & \end{matrix} \Rightarrow |\mathbf{E}_{M_3}| = \mathbf{18},$$

which is equal to $3^2(3-1)$.

Table 2: Some Constructed Permutation Trellis Codes

Construction	M	η	Mapping Codebook
Prefix	4	1	$\left\{ \begin{array}{l} 1234, 1432, 3214, 3412, 1243, 1342, 4213, 4312, 2134, 2431, \\ 3124, 3421, 2143, 2431, 4123, 4321 \end{array} \right\}$
	5	0.9	$\left\{ \begin{array}{l} 12345, 52341, 14325, 54321, 32145, 52143, 34125, 54123, \\ 12435, 52431, 13425, 53421, 42135, 52134, 43125, 53124, \\ 21345, 51342, 24315, 54312, 31245, 51243, 34215, 54213, \\ 21435, 51432, 23415, 53412, 41235, 51234, 43215, 53214 \end{array} \right\}$
Chang	4	1	$\left\{ \begin{array}{l} 1234, 3214, 1243, 4213, 2134, 3124, 2143, 4123 \\ 1432, 3412, 1342, 4312, 2431, 3421, 2341, 4321 \end{array} \right\}$
	5	0.9	$\left\{ \begin{array}{l} 12345, 15342, 32145, 35142, 12435, 15432, 42135, \\ 45132, 21345, 25342, 31245, 35241, 21435, 25431, \\ 41235, 45231, 14325, 15324, 34125, 35124, 13425, \\ 15423, 43125, 45123, 24315, 25314, 34215, 35214, \\ 23415, 25413, 43215, 45213 \end{array} \right\}$
k -Cube	4	1	$\left\{ \begin{array}{l} 1243, 1342, 4213, 4312, 1234, 1432, 3214, 3412, 2143, 2341, \\ 4123, 4321, 2134, 2431, 3124, 3421 \end{array} \right\}$
	5	0.9	$\left\{ \begin{array}{l} 12345, 52341, 14325, 54321, 32145, 52143, 34125, 54123, \\ 12435, 52431, 13425, 53421, 42135, 52134, 43125, 53124, \\ 21345, 51342, 24315, 54312, 31245, 51243, 34215, 54213, \\ 21435, 51432, 23415, 53412, 41235, 51234, 43215, 53214 \end{array} \right\}$
Exhaustive Search	5	1	$\left\{ \begin{array}{l} 12534, 21435, 13254, 24153, 21354, 12345, 23514, \\ 23145, 15243, 51423, 25134, 53241, 41325, 21543, \\ 31524, 35142, 14235, 12453, 34251, 54132, 42513, \\ 32415, 34512, 43152, 54321, 52431, 45231, 35421, \\ 52314, 45312, 43521, 53412 \end{array} \right\}$

Proposition 4 Any set C_M^t is an optimum set and the sum on the Hamming distances attains the upper bound.

Proof: Taking into consideration previous propositions, the proof is almost trivial. We have proven that the sum on the Hamming distances has attained the upper bound on the distances with all the designed cyclic-shift types of sets, which were constructed from cyclic shifts of their elements. Thus all the constructed sets C_M^t are optimum [5]– [8].

Proposition 5 Any set G of length M and cardinality multiple of M , with $G = C_M^1 \cup C_M^2 \cup \dots \cup C_M^n$ and $n = \lfloor |G|/M \rfloor$ is an optimal set where each one of its symbols appears only n times in each column.

Proof: As has been defined, $G = C_M^1 \cup C_M^2 \cup \dots \cup C_M^n$. Under the assumption that all sets are distinct, C_M^t are distinct where $C_M^1 \cap C_M^2 \cap \dots \cap C_M^n = \Phi$ and thus we have $|G| = |C_M^1| + |C_M^2| + \dots + |C_M^n| \Rightarrow |G| = n \times M$.

We have proven that all C_M^t are optimum and their symbols appear only once in each of the columns. So for the set G we have each symbol appears only n times in each column. Thus G is optimum and the sum on the Hamming distances attains the upper bound.

Example 3 $M = 4$, $C_4^1 = \{1234, 2341, 3412, 4123\}$, $C_4^2 = \{1324, 3241, 2413, 4132\}$, $C_4^3 = \{2314, 3142, 1423, 4231\}$, $C_4^4 = \{3214, 2143, 1432, 4321\}$.

$$M_4^1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow |E_{M_4^1}| = 48$$

$$M_4^2 = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 3 & 2 & 4 & 1 \\ 2 & 4 & 1 & 3 \\ 4 & 1 & 3 & 2 \end{bmatrix} \Rightarrow |E_{M_4^2}| = 48$$

$$M_4^3 = \begin{bmatrix} 2 & 3 & 1 & 4 \\ 3 & 1 & 4 & 2 \\ 1 & 4 & 2 & 3 \\ 4 & 2 & 3 & 1 \end{bmatrix} \Rightarrow |E_{M_4^3}| = 48$$

$$M_4^4 = \begin{bmatrix} 3 & 2 & 1 & 4 \\ 2 & 1 & 4 & 3 \\ 1 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix} \Rightarrow |E_{M_4^4}| = 48$$

$$G = C_4^1 \cup C_4^2 \cup C_4^3 \cup C_4^4 \Rightarrow G = \left\{ \begin{array}{l} 1 \ 2 \ 3 \ 4 \\ 2 \ 3 \ 4 \ 1 \\ 3 \ 4 \ 1 \ 2 \\ 4 \ 1 \ 2 \ 3 \\ 1 \ 3 \ 2 \ 4 \\ 3 \ 2 \ 4 \ 1 \\ 2 \ 4 \ 1 \ 3 \\ 4 \ 1 \ 3 \ 2 \\ 2 \ 3 \ 1 \ 4 \\ 3 \ 1 \ 4 \ 2 \\ 1 \ 4 \ 2 \ 3 \\ 4 \ 2 \ 3 \ 1 \\ 3 \ 2 \ 1 \ 4 \\ 2 \ 1 \ 4 \ 3 \\ 1 \ 4 \ 3 \ 2 \\ 4 \ 3 \ 2 \ 1 \end{array} \right\}$$

It is clear that the cardinality of the set G is $|G| = 4 \times 4$. Each symbol appears only four times at each column. Thus the set G is optimum and the sum on the Hamming distances reaches the upper bound.

Applying equation (4) to this case with $M = 4$ gives $\alpha = 1$ and $\beta = 0$, which means that $|E_{\max}| = 768$ and this is equal to the sum of the Hamming distance in the set G .

Proposition 6 *In the general case when $|G| = n_1 \times M + n_2$, with $n_2 < M$, the number of times that each symbol will appear in each column is either $\lfloor (|G|/M) \rfloor$ or $\lceil (|G|/M) \rceil$.*

Proof: We have $|G| = n_1 \times M + n_2 \Rightarrow \lfloor (|G|/M) \rfloor = n_1$ and $\lceil (|G|/M) \rceil = n_1 + 1$

In this case we have $G = C_M^1 \cup C_M^2 \cup \dots \cup C_M^{n_1} \cup C_M^{n_2}$. All sets C_M^i , $1 \leq i \leq n_1$ are with cardinalities equal to M . For $|C_M^{n_2}| < M$ we do not have cyclic shifts of all symbols, which means that some symbols will not appear in some of the positions in certain columns. Those appearing will only appear “once” because of the cyclic shift property presented before. Hence each symbol in the set G will appear only n_1 or $(n_1 + 1)$ times in each position column.

It is clear that all vectors in the set $C_M^{n_2}$ are results of a cyclic-shift from one vector to another, which means that they are all different. Thus the sum on the Hamming distances reaches the upper bound. Since all the sets are distinct, it leads to the fact that the set G is optimum.

Example 4 In the case of $M = 4$, we have three different permutation sets: $C_4^1 = \{1234, 2341, 3412, 4123\}$, $C_4^2 = \{1324, 3241, 2413, 4132\}$, $C_4^3 = \{2314, 3142\}$

$$M_4^1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow |E_{M_4^1}| = 48$$

$$M_4^2 = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 3 & 2 & 4 & 1 \\ 2 & 4 & 1 & 3 \\ 4 & 1 & 3 & 2 \end{bmatrix} \Rightarrow |E_{M_4^2}| = 48$$

$$M_4^3 = \begin{bmatrix} 2 & 3 & 1 & 4 \\ 3 & 1 & 4 & 2 \end{bmatrix} \Rightarrow |E_{M_4^3}| = 8$$

$$G = C_4^1 \cup C_4^2 \cup C_4^3 \Rightarrow G = \left\{ \begin{array}{l} 1 \ 2 \ 3 \ 4 \\ 2 \ 3 \ 4 \ 1 \\ 3 \ 4 \ 1 \ 2 \\ 4 \ 1 \ 2 \ 3 \\ 1 \ 3 \ 2 \ 4 \\ 3 \ 2 \ 4 \ 1 \\ 2 \ 4 \ 1 \ 3 \\ 4 \ 1 \ 3 \ 2 \\ 2 \ 3 \ 1 \ 4 \\ 3 \ 1 \ 4 \ 2 \\ 1 \ 4 \ 2 \ 3 \\ 4 \ 2 \ 3 \ 1 \\ 3 \ 2 \ 1 \ 4 \\ 2 \ 1 \ 4 \ 3 \\ 1 \ 4 \ 3 \ 2 \\ 3 \ 1 \ 4 \ 2 \end{array} \right\}$$

We have $|G| = 2 \times 4 + 2$ and each symbol from G at each column appears only two or three times, “1” appears two times while “2” appears three times. Thus G is an optimum set.

We can now generalize our proposition to cardinalities, which are not multiples of the cardinality of C_M^l .

5. DISTANCE OPTIMALITY FOR DPMS

If we map the outputs of a convolutional code, with a rate of $R = k/n$, onto permutation sequences of length M . We have to choose 2^n codewords from $M!$ codewords from the set S_M .

Let’s call S_M^n the subset of S_M with $|S_M^n| = 2^n$. To guarantee that our subset S_M^n is optimum we need to have $S_M^n = C_M^1 \cup C_M^2 \cup \dots \cup C_M^n$ with $C_M^{t_1} \neq C_M^{t_2}$, $t_1 \neq t_2$.

Proposition 7 *For any set S_M^n with cardinality $|S_M^n| = 2^n$, the number of times that each symbol will appear in each column is $\lfloor (2^n/M) \rfloor$ or $\lceil (2^n/M) \rceil$.*

Proof: special case of proposition 6

Proposition 8 *Any set G , of length M of non-repetitive permutation symbols, which has all of its elements appearing in each column only $\lfloor (|G|/M) \rfloor$ or $\lceil (|G|/M) \rceil$*

Table 3: Sum of number of appearances

Codebook	number of appearances					Sum
	$p^1(1)$	$p^2(1)$	$p^3(1)$	$p^4(1)$	$p^5(1)$	
G_4	4	4	4	4	–	16
G_5	6	6	6	7	7	32
G'_5	8	4	8	8	16	44

times is an optimal set and its sum of the Hamming distances attains the upper bound.

Proof: If each symbol is repeated in each column, then it will be only for $\lfloor(|G|/M)\rfloor$ or $\lceil(|G|/M)\rceil$ times. The resultant codebook, G , can be regrouped as follows:

$G = C_M^1 \cup C_M^2 \cup \dots \cup C_M^{n_1} \cup C_M^{n_2}$, where each of the subsets is an optimum set since all of its elements are in different positions in each column. The sign \cup reflect the cascaded way of putting all C_M^i together to finalize the codebook G .

The set G is then the union of the optimum sets, which leads to the fact that G is an optimum set. This could also be seen from the fact that the sum on the Hamming distances attains the upper bound.

Example 5 We consider the example of three sets designed differently. G_4 and G_5 are optimum and have been designed respectively by using a mapping algorithm and an exhaustive search. G'_5 was not optimum but generated by a mapping algorithm [21, 22].

$$G_4 = \left\{ \begin{array}{l} 1234, 3214, 1243, 4213, 2134, 3124, 2143, 4123 \\ 1432, 3412, 1342, 4312, 2431, 3421, 2341, 4321 \end{array} \right\}$$

$$G_5 = \left\{ \begin{array}{l} 12534, 21435, 13254, 24153, 21354, 12345, 23514, \\ 23145, 15243, 51423, 25134, 53241, 41325, 21543, \\ 31524, 35142, 14235, 12453, 34251, 54132, 42513, \\ 32415, 34512, 43152, 54321, 52431, 45231, 35421, \\ 52314, 45312, 43521, 53412 \end{array} \right\}$$

$$G'_5 = \left\{ \begin{array}{l} 12345, 15342, 32145, 35142, 12435, 15432, 42135, \\ 45132, 21345, 25342, 31245, 35241, 21435, 25431, \\ 41235, 45231, 14325, 15324, 34125, 35124, 13425, \\ 15423, 43125, 45123, 24315, 25314, 34215, 35214, \\ 23415, 25413, 43215, 45213 \end{array} \right\}$$

Table 3 shows the number of appearances of the symbol “1” in each column and also the total number of it appearance in the G_4 , G_5 and G'_5 sets.

Symbol “1” has appeared the same number as the cardinality of the sets G_4 and G_5 . This is expected since

they are optimum on the sum of the Hamming distances. For G'_5 it appears 44 times and this explain the fact that this set is not optimum.

We conclude with a summary of what we have derived from the foregoing discussion.

- Cyclic-shift of data increases the sum on the Hamming distances.
- Any set of M elements cyclically shifted leads to an optimum set.
- Any set obtained as a result of the union of optimum sets is an optimum set.
- Any set whose elements appear only a specific number of times is a an optimum set
- Each element appears as $|G|$ in all columns of the set.

6. CYCLIC-SHIFT PREFIX/SUFFIX TECHNIQUE

We present in this section another technique based on the cyclic shift technique but this time for any cardinality of the subsets. We use the fact that all codewords in the same set are different to reach the upper bound on the Hamming distances and thus we need to guarantee that we do not have repetitive codewords in the same codebook or symbols at the same positions as presented previously.

6.1 The Prefix/Suffix Technique

To present the technique of addition of a suffix or a prefix we take two examples of sets of lengths $M = 3$ and $M = 4$.

Example 6 For the case of $M = 4$, we take two different codewords, $\mathbf{y}_1 = (1, 2, 3, 4)$ and $\mathbf{y}_2 = (3, 4, 1, 2)$. Their corresponding cyclic-shift matrices are presented as follows:

$$M_4^1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow |\mathbf{E}_{M_4^1}| = 48,$$

$$M_4^2 = \begin{bmatrix} 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix} \Rightarrow |\mathbf{E}_{M_4^2}| = 48.$$

Note that although we used two different codewords, we ended up with the same sets $M_4^1 = M_4^2$. We have to take

$$(1, 2, 3, \dots, M-1) + \{M\} \xrightarrow{F_{sh}} M_M^1 = \begin{bmatrix} 1 & 2 & 3 & \dots & M-1 & M \\ 2 & 3 & \dots & M-1 & M & 1 \\ 3 & \dots & M-1 & M & 1 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M & 1 & 2 & 3 & \dots & M-1 \end{bmatrix}, \quad (6)$$

$$(3, 2, M-1, \dots, 1) + \{M\} \xrightarrow{F_{sh}} M_M^2 = \begin{bmatrix} 3 & 2 & M-1 & \dots & 1 & M \\ 2 & M-1 & \dots & 1 & M & 3 \\ M-1 & \dots & 1 & M & 3 & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M & 3 & 2 & M-1 & \dots & 1 \end{bmatrix}. \quad (7)$$

into consideration the repetition of codewords when using the cyclic shift technique. The solution to this is to apply a “suffix” or a “prefix” that can break the sequence of codewords and guarantee non-repetitive sequences. □

The permutation of non-repetitive symbols in a sequence of length M , results in the generation of $M!$ different permutation sequences or codewords. As always we have $M! = (M-1)!M$, the symbol M could play a very important role as a suffix/prefix position for a permutation sequence of length $M-1$.

If we take two sequences of length $M-1$ and then cyclic-shift them, we might end up with two sets of same codewords as seen in Example 6. Thus the technique of adding a suffix/prefix M to the permutation sequence of length $M-1$ will avoid the repetition of codewords as shown in (6). If we take another sequences of length $M-1$ and add a suffix/prefix we will have different sets as presented in (7).

Example 7 For $M = 3$, we have $M! = 6$. For $M = 4$, we have $M! = 6 \times 4$. We note that adding the symbol M as a prefix will lead to the multiplication of the cardinality of the codebook with codewords length $M-1$ by a factor of M .

For the case of $M = 3$, if we take two different codewords, $\mathbf{y}_1 = (1, 2, 3)$ and $\mathbf{y}_2 = (3, 1, 2)$ and then we add a suffix 4 to each one of them and cyclic-shift the resultant codewords $M-1$ times to obtain for each the corresponding M_M^i as shown below:

$$(1, 2, 3) + \{4\} \xrightarrow{F_{sh}} M_4^1 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow |E_{M_4^1}| = 48,$$

$$(3, 1, 2) + \{4\} \xrightarrow{F_{sh}} M_4^2 = \begin{bmatrix} 3 & 1 & 2 & 4 \\ 1 & 2 & 4 & 3 \\ 2 & 4 & 3 & 1 \\ 4 & 3 & 1 & 2 \end{bmatrix} \Rightarrow |E_{M_4^2}| = 48.$$

It is clear that the suffix 4 has prevented the repetition of the sequences in both sets M_4^1 and M_4^2 as was the case in Example 6. The two sets are different and optimum since both have reached the upper bound on the sum of the Hamming distances.

We have shown how the cyclic shifting of symbols increases the sum on the Hamming distances and how the suffix prevents the repetition of codewords without affecting the Hamming distances.

Proposition 9 *The addition of a prefix or a suffix to all codewords of a codebook, of length $M-1$ and cardinality n , and the cyclic shift of all its codewords will lead to an optimum codebook on the sum of the distances of length M and cardinality nM .*

Proof: As it was mentioned before, the cyclic-shift of a codeword of length M , generates an optimum codebook of cardinality M . If we have a codebook of length $M-1$ and cardinality n , then by adding a suffix to each of its codewords and cyclic-shift them we obtain M new codewords from each of its codewords, which form separately optimum codebooks. Thus we have in total nM new codewords as the cardinality of the new codebook, which is considered to optimum on the sum of the distances since it is the union of n optimum codebooks.

6.2 Codebooks with Cardinalities of 2^n

In general a permutation codebook should have a cardinality equal to 2^n to map the outputs of a convolutional base code with a rate of $R = k/n$. Although we are not interested in this paper in the mapping technique but just to study the conditions that make any permutation codebook reaches the upper bound on the sum of the Hamming distances.

To explain the idea simply, we take the case of mappings with the same lengths between the convolutional outputs and the permutation sequences. We take the case of $M = n$, where we need to choose only $\lceil 2^n/M \rceil$ from $(M-1)!$ permutation codebook. We add the suffix and use the cyclic-shift technique of the codewords.

Example 8 We take the case of $M = 4$, we choose randomly $2^4/4 = 4$ codewords from the $(M - 1)! = 3! = 6$ codebook.

$$\left\{ \begin{matrix} 1 & 2 & 3 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \\ 3 & 2 & 1 \end{matrix} \right\} \Rightarrow \left\{ \begin{matrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{matrix} \right\}$$

After selecting the subset from the $(M - 1)!$ codebook (in our case 3!), we add the prefix to achieve the desired length of the codeword,

$$\left\{ \begin{matrix} 1 & 3 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{matrix} \right\} + \left\{ \begin{matrix} 4 \\ 4 \\ 4 \\ 4 \end{matrix} \right\} \Rightarrow \left\{ \begin{matrix} 1 & 3 & 2 & (4) \\ 2 & 1 & 3 & (4) \\ 2 & 3 & 1 & (4) \\ 3 & 1 & 2 & (4) \end{matrix} \right\}.$$

Now we are sure that by cyclic shifting our codewords we will not generate repetitive codewords and this will lead to a set with 2^4 codewords for the number required for our mapping,

$$\left\{ \begin{matrix} 1 & 3 & 2 & (4) \\ 2 & 1 & 3 & (4) \\ 2 & 3 & 1 & (4) \\ 3 & 1 & 2 & (4) \end{matrix} \right\} \xrightarrow{F_{sh}} \left\{ \begin{matrix} 132(4), 32(4)1, 2(4)13, (4)132, \\ 213(4), 13(4)2, 3(4)21, (4)213, \\ 231(4), 31(4)2, 1(4)23, (4)231, \\ 312(4), 12(4)3, 2(4)31, (4)312 \end{matrix} \right\}.$$

In the following example we study the case of codebooks with cardinalities which are not multiples of M .

Example 9 In the case of $M = 5$. To design a codebook with 2^5 codewords we have to get $\lceil 2^5/5 \rceil = 7$. This which means that we need 7 subsets of C_M as described in the previous sections. Thus we only choose the first 2^5 codewords.

7. CONCLUSION

The purpose of this work is to analyze and find the reasons behind the optimality of any distance-preserving mappings construction. Using the technique of the cyclic-shift of a prefix, the upper bound on the sum of the Hamming distances is investigated and the condition on the position of any symbol in a permutation sequence that should lead to an optimum permutation set is explained.

The cyclic-shift prefix/suffix codebooks that we have designed have all reached the upper bound on the sum of the Hamming distances. For Distance-preserving mapping codes, it is required to have codes which can reach the upper bound on the sum of the Hamming distances since it can guarantee the increase of the minimum distance and thus the number of corrected errors. Therefore the error correction capability will improve.

From the study conducted in this paper, it was confirmed

that the k -cube construction is optimum where the sum on the Hamming distances reached the upper bound.

It is important to mention that this research focused only on the reasons that make a code reach the upper bound on the sum of the Hamming distances. But it is possible to further investigate these types of codes and derive other decoding algorithms inspired from their fundamental properties, which is the cyclic-shifting property.

REFERENCES

- [1] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, Cambridge, MA: MIT Press, 1972.
- [2] E. N. Gilbert, "Cyclically permutable error-correcting codes," *IEEE Trans. Inf. Theory*, vol. IT-9, no. 4, pp. 175–182, Jul. 1963.
- [3] D. E. Maracle and C. T. Wolverton, "Generating cyclically permutable codes," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 554–555, Jul. 1974.
- [4] M. Kuribayashi and H. Tanaka, "How to Generate Cyclically Permutable Codes from Cyclic Codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4660–4663, oct. 2006.
- [5] T. G. Swart and H. C. Ferreira, "A generalized upper bound and a multilevel construction for distance-preserving mappings," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3685–3695, Aug. 2006.
- [6] K. Lee, "Cyclic constructions of distance-preserving maps," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4392–4396, Dec. 2005.
- [7] T. Wadayama and A. J. H. Vinck, "A multilevel construction of permutation codes," *IIEICE Transactions on Fundamentals*, vol. E84-A, no. 10, pp. 2518–2522, Oct. 2001.
- [8] K. Ouahada and H. C. Ferreira, "A k -Cube Construction mapping binary vector to permutation," in *Proceedings of the International Symposium on Information Theory*, South Korea, Seoul.
- [9] H. C. Ferreira, A. J. H. Vinck, T. G. Swart and I. de Beer, "Permutation trellis codes," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1782–1789, Nov. 2005.
- [10] H. C. Ferreira and D. A. Wright and A. L. Nel, "Hamming distance preserving mappings and trellis codes with constrained binary symbols," *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 1098–1103, Sept. 1989.
- [11] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland, 1977.

- [12] G. D. Forney Jr., "The Viterbi algorithm", *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–277, Mar. 1973.
- [13] T. G. Swart, I. de Beer, H. C. Ferreira and A. J. H. Vinck, "Simulation results for permutation trellis codes using M-ary FSK," in *Proc. Int. Symp. on Power Line Com. and its Applications*, Vancouver, BC, Canada, Apr. 6–8, 2005, pp. 317–321.
- [14] H. C. Ferreira and A. J. H. Vinck, "Interference cancellation with permutation trellis codes," in *Proceedings of the IEEE Vehicular Technology Conference*, Boston, MA, USA, Sept. 24–28, 2000, pp. 2401–2407.
- [15] A. J. H. Vinck and J. Haering, "Coding and Modulation for Power Line Communications," in *Proceedings of the International Symposium on Power-Line Communications and its Applications*, Limerick, Ireland, Apr. 5–7, 2000, pp. 265–272.
- [16] A. J. H. Vinck and J. Haering and T. Wadayama, "Coded M-FSK for power line communications," in *Proceedings of the IEEE International Symposium on Information Theory*, Sorrento, Italy, Jun. 25–30, 2000, pp. 137.
- [17] K. W. Shum, "Permutation coding and M-FSK modulation for frequency selective channel," in *IEEE Personal, Indoor and Mobile Radio Communications*, Sorrento, Italy, vol. 13, no. 5, Sept. 25–30, 2002, pp. 2063–2066.
- [18] A. Viterbi and J. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill Kogakusha LTD, Tokyo Japan, 1979.
- [19] H. C. Ferreira and I. de Beer and A. J. H. Vinck, "Distance preserving mappings onto convolutional codes revisited," in *Proceedings of the IEEE Information Theory Workshop*, Breisach, Germany, Jun. 26–29, 2002, pp. 23–26.
- [20] J.-C. Chang and R.-J. Chen and T. Kløve and S.-C. Tsai, "Distance-preserving mappings from binary vectors to permutations," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 1054–1059, Apr. 2003.
- [21] J.-C. Chang, "Distance-increasing mappings from binary vectors to permutations," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 359–363, Jan. 2005.
- [22] J.-C. Chang, "Distance-increasing mappings from binary vectors to permutations that increase Hamming distances by at least two," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1683–1689, Apr. 2006.

MODIFIED EARLIEST DECODING IN NETWORKS THAT IMPLEMENT RANDOM LINEAR NETWORK CODING

S. von Solms* and A.S.J. Helberg*

* School of Electric, Electronic and Computer Engineering, North-West University, Potchefstroom Campus, Potchefstroom. E-mail: {sune.vonsolms, albert.helberg}@nwu.ac.za

Abstract: In this paper we present a formalised description of Modified Earliest Decoding. We simulate the performance of the method in comparison with Earliest Decoding in networks that implement random linear network coding. We show that Modified Earliest Decoding has a smaller decoding complexity than Earliest Decoding and Gaussian Elimination as well as a smaller decoding delay.

Keywords: Earliest Decoding, Gaussian Elimination decoding, Network Coding, Random Linear Network Coding

1. INTRODUCTION

The decentralised approach to network coding namely random linear network coding (RLNC) allows for a more practical approach to network coding [1]. Random linear network coding employs random and independent coding of received packets at intermediate network nodes.

Due to the employment of RLNC in networks, randomly encoded packets are received at the sink nodes. The sink nodes need to employ a decoding method to successfully obtain the source information. There exist several decoding methods in the literature that can be successfully implemented with RLNC, but generally the sink node may not be able to decode the source packets until an entire block of encoded packets are received [2]. This leads to a decoding delay which is not favourable for delay sensitive networks [2, 3]. Decoding delay can be seen as the elapsed time between the reception of a packet at a receiver node and the decoding thereof [2]. The challenge, therefore, is to find a decoding method with a small decoding delay as well as low decoding complexity.

Gaussian Elimination (GE) is a possible decoding method, but is computationally complex due to the use of matrix inversion and it has a decoding delay equal to the length of time needed by the receiver to collect encoded packets of full rank [4, 5].

Earliest Decoding (ED) is a method developed to decrease the decoding delay of GE. This method entails the use of GE on linearly independent packets of sufficient rank as soon as they are collected by a receiver node. The decoding delay of ED is approximately constant and independent of the number of transmitted source packets, but still employs computationally complex matrix inversion [4 - 6].

In this paper we look at an improvement on ED for the implementation in a network that uses RLNC, called Modified Earliest Decoding (MED). This method is based on ED and shows an improvement on the low decoding delay of ED. Modified Earliest Decoding also reduces the decoding complexity by significantly reducing the use of matrix inversion for decoding. The method of MED was proposed in [7], but it was neither formalised nor analysed. We present a formal algorithm of MED and an analysis through simulation of the decoding delay and complexity of MED.

2. NETWORK ENVIRONMENT

A typical network environment where RLNC can be implemented successfully is that of wireless ad-hoc and sensor networks [8]. Wireless sensor networks include scenarios where a block of data or file needs to be transmitted from a single source to a receiver where the intermediate nodes do not require the file. A node in a wireless sensor network is connected to another node in the network when one node is in the coverage of the other node's signal.

It is shown that the random geometric graph (RGG) is a realistic model for a wireless sensor network as it considers the communication distances of nodes [9]. We adopt the notation used in [1, 5] and the graph construction of [10, 11].

Consider an acyclic network which implements RLNC as a random geometric graph $\mathcal{G} = (N, l) = (\mathcal{V}, \mathcal{E})$ with $N = |\mathcal{V}|$. The graph is formed by placing N nodes uniformly at random on a unit square with communication radius of l . An edge $e = (u, v) \in \mathcal{E}$ exists between two nodes $u, v \in \mathcal{V}$ when the Euclidean distance between u and v is $d(u, v) \leq l$, where the value of l corresponds to the broadcast radius of a node in the wireless network. We assume a symmetric case where all

the network nodes have equal transmission power and thus an identical connectivity radius l . The probability p that two nodes u, v are connected is bounded by:

$$\frac{1}{4}(\pi l^2) \leq p \leq \pi l^2. \quad (1)$$

The lower bound is due to the fact that a node can be situated in one of the corners in the unit square. The upper bound is the direct consequence of the communication radius of a node [12].

The wireless sensor network consists of a single source node $s \in \mathcal{V}$ and a set of sink nodes $Z = \{z_1, \dots, z_{|Z|}\}$, $Z \subset \mathcal{V}$ with $\min\text{-cut}(s, z) \geq n$. The data present at the source node, s , is divided into n packets and denoted by

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \quad (2)$$

where \mathbf{x}_i represents the i th source packet from a finite field \mathcal{F} of size q . These source packets are multicast sequentially over the edges $e \in \mathcal{E}$ of network \mathcal{G} to synchronised intermediate nodes $v \subset \mathcal{V}$. This means that the source node does not transmit encoded packets but single source packets one by one.

Each intermediate network node randomly and linearly combines the packets received from its incoming edges e' to form a new encoded packet to be transmitted on its outgoing edges e . A coding vector of length n is included in the header of each outgoing packet. It describes the source packets that have been linearly combined in the transmitted packet.

Each receiver node $z \in Z$ collects a set of $N \geq n$ encoded packets from the network, $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$, where the j th encoded packet is of the form

$$\mathbf{y}_j = \sum_{i=1}^n g_{ij} \mathbf{x}_i, j = 1, 2, \dots, N \quad (3)$$

where the coefficients $\{g_{ij}\}$ are randomly generated from a finite field \mathcal{F}_2 and \mathbf{g}_j forms the global coding vector of packet \mathbf{y}_j . These coding vectors can be represented as the column vectors of a $n \times N$ matrix \mathbf{G} [4, 13] where

$$\mathbf{X} \times \mathbf{G} = \mathbf{Y}. \quad (4)$$

The solution of the linear system of equations in (4) decodes the source packets \mathbf{X} .

3. RELATED WORK

Next we discuss three known decoding methods that can be implemented in a network that implements RLNC. Earliest Decoding and belief propagation (BP) decoding require alterations to the encoding procedure of the network, where Gaussian Elimination can be implemented without any alterations. The choice of decoding method influences the decoding delay as well as the computational resources at the receiver nodes.

3.1 Gaussian Elimination

Gaussian Elimination is an efficient method for solving a system of linear equations as described in (4). Gaussian Elimination can be performed only when \mathbf{G} is of full rank n . Thus the decoding delay of GE equals the time the receiver has to wait in order to collect n linearly independent packets which is proportional to the size of n . Thus the decoding delay of GE increases linearly with the increase of source packets. Gaussian Elimination requires $\mathcal{O}(n^3)$ operations for decoding via matrix inversion which is computationally complex [4, 5]. In a situation of a small number of source messages, GE is an efficient decoding method, but decreases in efficiency as n becomes large.

3.2 Earliest Decoding [4]-[6]

Earliest Decoding performs the same decoding steps as GE but does not require \mathbf{G} being of full rank n . Earliest Decoding allows a receiver to perform decoding on a subset of source packets as soon as sufficient information is received, even though the decoding matrix is incomplete. This decoding algorithm is run every time an innovative packet is obtained at the receiver. An *innovative packet* is defined as a packet that increases the rank of \mathbf{G} . This method enables a receiver to decode a subset of source packets, $\mathbf{X}' \subseteq \mathbf{X}$ when the global coding vectors in matrix $\mathbf{G}' \subseteq \mathbf{G}$ can be successfully inverted.

The sequential multicasting of source packets over the network results in \mathbf{G} likely to be lower triangular. This means that $\mathbf{X}' = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$, $m \leq n$ can be decoded with high probability after the collection of m or $m + \delta$ packets, where δ is small in relation to m [5]. This makes ED practical in a RLNC environment.

Earliest Decoding yields a smaller decoding delay than GE as packets can be decoded by the receiver while still obtaining innovative packets. The decoding delay stays approximately constant and independent of n [4]. The inherent divide and conquer approach also leads to a faster decoding time, but still requires computationally complex matrix inversion.

3.1 Belief propagation decoding

The decoding method employed for Luby Transform (LT) codes [14] is an iterative process where a sink node first has to determine the degree of each received packet \mathbf{y}_j .

Definition 1: The *degree* of a packet indicates the number of source packets linearly combined in the packet or can be seen as the number of non-zero entries in the packet's global encoding vector \mathbf{g}_j .

The BP decoding process can be described by the following steps [14]:

1. Find an encoded packet, $\mathbf{y}_j, 1 \leq j \leq N$, which only contains a single source packet, $\mathbf{x}_i, 1 \leq i \leq n$ (i.e. native packet or packet of degree one).
2. Set source packet $\mathbf{x}_i = \mathbf{y}_j$ and delete \mathbf{y}_j .
3. Subtract the value of \mathbf{x}_i from all the other encoded packets $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_{j-1}, \mathbf{y}_{j+1}, \dots, \mathbf{y}_N]$, $N \geq n$ that contains source packet \mathbf{x}_i , reducing their degrees.

The reduction of the packet's degrees produces a new native packet with high probability. Repeat process from (1) until all source packets $\mathbf{x}_i, 1 \leq i \leq n$ are determined.

To ensure the presence of a native packet each time the process iterates, all packets are encoded according to the Robust Soliton (RS) degree distribution [14].

4. MODIFIED EARLIEST DECODING

Earliest Decoding is successful when m linearly independent packets are present to decode all m source packets. Modified Earliest Decoding [7] allows for an iterative approach to the decoding of source packets.

Modified Earliest Decoding applies the low decoding delay concept of ED but reduces the decoding complexity by significantly reducing the use of matrix inversion. As with ED, MED runs the decoding algorithm every time a new innovative packet is obtained at a receiver node.

The sequential transmission of source packets leads to the scenario where the m th received packet \mathbf{y}_m tends to be a linear combination of the first m transmitted source packets $\mathbf{X}' = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m], m \leq n$. Due to this lower triangular structure of \mathbf{G} it is possible to decode the source packets through a method adopted from the low complexity belief propagation (BP) decoding algorithm of LT codes.

4.1 Decoding in a RLNC environment

In a RLNC network scenario packets are encoded

randomly and thus employing a low complexity decoding method, like BP, which requires packets to be from the RS distribution, can be complicated. The lower triangular structure of \mathbf{G} , however, largely consists of encoded packets where the packets following each other only contain a single additional source packet. A source packet can be decoded when receiving two packets with coding vectors have a Hamming distance of $d_H(\mathbf{g}_p, \mathbf{g}_q) = 1$.

Definition 2 [15]: The *Hamming distance* between two vectors $\mathbf{g}_p, \mathbf{g}_q$ is defined as the number of coordinates that they differ and is denoted by

$$d_H(\mathbf{g}_p, \mathbf{g}_q). \quad (5)$$

Definition 3: The *Hamming weight* of a vector \mathbf{g}_p is defined as the number of non-zero coordinates and is denoted by

$$w_H(\mathbf{g}_p). \quad (6)$$

In the context of this paper, the Hamming weight of a coding vector is equivalent to the degree of the packet.

The linear combination of two packets with $d_H = 1$ produces a native packet with a coding vector $\mathbf{g}_{x_i} = \mathbf{g}_p \oplus \mathbf{g}_q$ of degree one. This packet contains information regarding a single source packet \mathbf{x}_i which is equivalent to a decoded packet. This native packet can now be used for decoding in a similar way as with BP decoding. By linearly combining \mathbf{x}_i with other received packets containing \mathbf{x}_i , the degrees of these packets are reduced. This process is iterated until all the other source packets $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n]$ are decoded.

This MED decoding algorithm is summarised in Algorithm 1.

Algorithm 1: Modified Earliest Decoding

```

1: Initialise  $\mathbf{g}_0 = \mathbf{0}, \mathbf{g}_0 \in \mathbf{G}$ , and  $\mathbf{y}_0 = \mathbf{0}, \mathbf{y}_0 \in \mathbf{Y}$ 
2: while not all  $\{\mathbf{x}_i\}_{i=1}^n$  are decoded do
3:   Collect  $\{\mathbf{g}_m\}_{m>0} \in \mathbf{G}$  of received packet  $\{\mathbf{y}_m\}_{m>0} \in \mathbf{Y}$ 
4:   while  $\mathbf{g}_m$  contains a native (unit) coding vector  $\mathbf{g}_{x_i}$  do
5:      $\mathbf{g}_m \leftarrow \mathbf{g}_m \oplus \mathbf{g}_{x_i}$ 
6:     update  $\mathbf{G}, \mathbf{Y}$ 
7:   end while
8:   determine  $d_H(\mathbf{g}_p, \mathbf{g}_q), \{\mathbf{g}_p, \mathbf{g}_q\} \in \mathbf{G}, q > p$ 
9:   if  $d_H(\mathbf{g}_p, \mathbf{g}_q) = 0$  do
10:    remove  $\mathbf{g}_q$  from  $\mathbf{G}$ 
11:    update  $\mathbf{Y}$ 
12:   end if
13:   if  $d_H(\mathbf{g}_p, \mathbf{g}_q) = 1$  do
14:      $\mathbf{g}_{x_i} \leftarrow \mathbf{g}_p \oplus \mathbf{g}_q$ 
15:     determine  $w_H(\mathbf{g}_p), w_H(\mathbf{g}_q)$ 
16:     remove  $\mathbf{g}_q, w_H(\mathbf{g}_p) < w_H(\mathbf{g}_q)$ 
17:   else remove  $\mathbf{g}_p, w_H(\mathbf{g}_q) < w_H(\mathbf{g}_p)$ 
18:     set  $\mathbf{x}_i \in \mathbf{X}$  equal to  $\mathbf{y}' \leftarrow \mathbf{y}_p \oplus \mathbf{y}_q$ 
19:     mark  $\mathbf{g}_{x_i}$  as a native (unit) coding vector
20:   update  $\mathbf{Y}$ 

```

```

21:   if any  $\{g_m\}_{m \in N}$  in  $G$  contains native  $g_{x_i}$  do
22:      $g_m \leftarrow g_m \oplus g_{x_i}$ 
23:     update  $G, Y$ 
24:   end if
25: end if
26: if no MED is possible, perform ED if possible
27: end while

```

Although the structure of G tends to be lower triangular allowing MED to function successfully, the random encoding of packets do not always guarantee a Hamming distance of 1 between packets. In this instance ED is performed on $G' \subseteq G$ to decode source packets allowing decoding to continue.

4.2 Example

The MED process is illustrated in a small example in Fig. 1. Assume that $n = 4$ and the receiver has obtained 4 encoded packets $Y = [y_1, y_2, y_3, y_4]$ from the network that implements RLNC.

I	II
$\begin{bmatrix} y_1 & y_2 & y_3 & y_4 \\ x_1 & x_2 & x_3 & x_4 \end{bmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} g_0 & g_1 & g_2 & g_3 & g_4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$
III	IV
$\begin{pmatrix} g_0 & g_1 & g_2 & g_3 & g_4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $w_H(g_1) = 2$ $w_H(g_3) = 3$ $d_H(g_1, g_3) = 1$	$\begin{pmatrix} g_0 & g_1 & g_2 & g_{x3} & g_4 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $g_{x3} = g_1 \oplus g_3$ $w_H(g_2) = 2$ $w_H(g_{x3}) = 1$ $d_H(g_2, g_{x3}) = 1$
V	VI
$\begin{pmatrix} g_0 & g_1 & g_{x1} & g_{x3} & g_4 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $g_{x1} = g_2 \oplus g_{x3}$ $w_H(g_1) = 2$ $w_H(g_{x1}) = 1$ $d_H(g_1, g_{x1}) = 1$	$\begin{pmatrix} g_0 & g_{x2} & g_{x1} & g_{x3} & g_4 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $g_{x2} = g_1 \oplus g_{x1}$ $w_H(g_{x1}) = 1$ $w_H(g_4) = 2$ $d_H(g_{x1}, g_4) = 1$
VII	VIII
$\begin{pmatrix} g_0 & g_{x2} & g_{x1} & g_{x3} & g_{x4} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ $g_{x4} = g_{x1} \oplus g_4$	$G = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Figure 1. Modified Earliest Decoding example

Frame (I): The global encoding vectors of the received packets are shown. In this example sufficient encoded packets are received before the decoding process starts. This is for clarity purposes for the example. Practically decoding can commence as soon as two packets with a Hamming distance of one are received.

Frame (II): By default a receiver adds the zero-vector as

coding vector g_0 to matrix G and then adds the received coding vectors $\{g_i\}_{i=1}^4$ to G .

Frame (III): Through evaluation it can be seen that $d_H(g_1, g_3) = 1$. The degrees of the packets are evaluated where $w_H(g_1) = 2$ and $w_H(g_3) = 3$.

Frame (IV): Because $w_H(g_1) < w_H(g_3)$, coding vector g_3 is removed and replaced by native vector $g_{x3} = g_1 \oplus g_3$. The process starts again where $d_H(g_2, g_{x3}) = 1$ and $w_H(g_{x3}) < w_H(g_2)$.

Frame (V): Coding vector g_2 is replaced by native vector $g_{x1} = g_2 \oplus g_{x3}$. The next iteration shows that $d_H(g_1, g_{x1}) = 1$ where $w_H(g_{x1}) < w_H(g_1)$.

Frame (VI): Coding vector g_1 is removed and replaced by native vector g_{x2} .

Frame (VI)-(VII): The last undecoded vector is g_4 that can be replaced by native vector g_{x4} because $d_H(g_{x1}, g_4) = 1$ and $w_H(g_{x1}) < w_H(g_4)$.

Frame (VIII): A permuted identity matrix can be seen which shows that all the source packets have been determined and the transmitted data successfully decoded.

Note: In this example only the linear operations performed on the coding vectors are shown. The same operations are performed on the corresponding data packets Y .

From the example it can be seen that MED is a modified version of ED, where MED also forms a sub-matrix from the global encoding matrix to decode. Modified Earliest Decoding, however, eliminates the presence of source packets in the encoded packets that are mutual. The above example cannot be decoded via ED as no sub matrix of G exist that can be inverted. Gaussian Elimination would have been performed after the reception of all the packets.

5. SIMULATION AND RESULTS

In this section we evaluate the decoding performance in the RLNC network environment, as described in Section II, when different decoding methods are implemented at the receiver nodes. We evaluate the decoding delay and decoding complexity for MED in comparison to ED.

5.1 Simulation setup

The network topology is based on that of [9] where the network $\mathcal{G} = (N, l) = (\mathcal{V}, \mathcal{E})$ with a single source s and single receiver z for simplicity. The wireless sensor network is modelled by a random geometric graph (RGG) formed by placing the nodes uniformly at random on a unit square with communication radius of l as described in Section 2. Let $r(s, z)$ be the achievable rate at which s can multicast the source packets reliably to the receiver z .

From the min-cut max-flow theorem, the value of $\min\text{-cut}(s, z)$ is the upper bound on $r(s, z)$ for z [16]. For a single source single receiver network, the expected value of $\min\text{-cut}(s, z)$ can be calculated by

$$(N - 1)p - \sqrt{(N - 1)p(1 - p)/\pi} \quad (7)$$

where the value of p is shown in (1) [10]. In order to ensure the successful transmission of n source packets from s to the receiver node z , the min-cut of the network must be equal to $\min\text{-cut}(s, z) \geq n$. From (1) and (7) the connectivity radius l and the number of network nodes N are chosen to accommodate the required min-cut value $\min\text{-cut}(s, z) \geq n$, for varying values of n .

The communication radius l is chosen specifically in each simulation set to ensure a minimum cut between source and receiver node of $\min\text{-cut}(s, z) \geq n$. If a constructed RGG has a min-cut smaller than the required n , the graph is discarded and a new RGG is generated.

We followed the method of *independent replications* from [17] in order to obtain results which are not affected by different network scenarios. For the simulations we generated 40 random geometric graphs with different seeds. From each random graph we run 5 instances with a different sources and receivers which are randomly chosen. Finally, for each of the sub-instances we ran the simulation 5 times with different seeds. This equates to 1000 Monte-Carlo simulations for each value of n .

The data transmitted by s to the receiver consists of approximately 10000 packets in the finite field \mathcal{F}_2^8 . These packets are divided into n transmission packets $\{x_i\}_{i=1}^n$ of size m . A coding vector of length n from finite field \mathcal{F}_2 is included in the header of each packet, describing included source packets.

5.2 Decoding delay

As described in Section I, the decoding delay at a receiver is defined as the elapsed time between the reception of an encoded packet and the decoding thereof.

We denote t as the timestep of the simulation when z obtains a new packet from the network. We denote the global rank of the network as R_n , which is equal to the number of source packets n . The rank present at receiver node z at time t is defined as $R_z(t)$. The source packets decodable by node z are defined as *effective packets* and the total number of effective packets at z up to time t is denoted as $E_z(t)$ [18]. The number of effective packets decodable by node z , $E_z(t)$, is upper bounded by the rank present at node z , $R_z(t)$. The value of $R_z(t)$ is in turn upper bounded by the number of packets received by z up until time t .

We ran the MED and ED algorithms at the receiver node for each simulation instance in order to get a fair comparison in decoding delay. Fig. 2 shows the normalised $E_z(t)/R_n$ decoding curves for the MED and ED for $n = 10$, where $N = 250$ and $l = 0.18$. The case for a small n is chosen for Fig. 2 as it clearly illustrates the difference in decoding delay. The curve $R_z(t)/R_n$ shows the normalised value of the rank available at z , which expresses the total number of source packets possibly decodable at time t . This curve gives the upper limit of decoding for any system at time t .

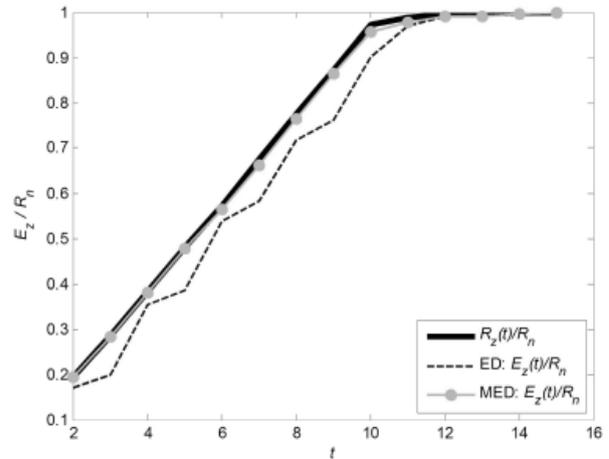


Figure 2: Normalised decoding delay of MED and ED for $n = 10$

It can be seen in Fig. 2 that MED produces a larger number of effective packets at z when innovative packets are still being received than compared to ED. This means that the MED method is able to decode more source packets at time t than the ED method, resulting in a smaller decoding delay. The graph further shows variable decoding delay for ED and constant decoding delay MED which may be of advantage for certain applications.

The decoding delay of ED is independent of n and remains approximately constant [4, 5]. From simulation in [4] it was shown that the algorithmic decoding delay of ED is often only in the order of a few source packets, much smaller than n . Therefore, the decoding delay observed in Fig. 2 would continue to be in the order of a few source packets even for larger values of n .

The same observation can be made for MED as this method also decodes subsets of a few source packets much smaller than n . The decoding method of MED for larger values of n continues to decode small subsets of source packets, independent of n , therefore producing a decoding delay in the order of a few source packets.

Thus for large values of n , the decoding delay of MED remains approximately constant and an improvement on that of ED, as can be seen in Fig. 2.

The decoding delay of MED is also upper bounded by that of ED, because when no packets with $d_H = 1$ are available the MED algorithm reverts to ED.

5.3 Decoding complexity

In addition to decoding delay another important characteristic of a decoding method is its decoding complexity. We determined the decoding complexity of ED and MED for $10 \leq n$ by calculating the number of arithmetic operations for both decoding methods.

Earliest Decoding consist of two steps namely *forward elimination* and *backward substitution*. The number of arithmetic operations for the forward elimination step is approximately

$$\sum_{i=1}^{u-1} (u-i) + \sum_{i=1}^{u-1} (u-i)^2 \quad (8)$$

divisions and multiplications/additions respectively, where $u < 1 \leq n$ is the size of the subset of source packets decoded in each step. The backward substitution step requires

$$\sum_{i=1}^{u-1} (u-i) \quad (9)$$

multiplications/additions.

For the MED algorithm the zero vector is added to \mathbf{G} , which leads to $u+1$ packets per subset. The $u+1$ packets are compared to each other which can require a maximum of

$$u + (u-1) + \dots + 1 = \frac{u(u+1)}{2} \quad (10)$$

arithmetic operations, where $u \leq 1 \leq n$ is the size of the subset of source packets decoded in each step. After these comparisons a single source packet is decoded and eliminated from the other packets in the block, which requires a maximum of

$$u-1 \quad (11)$$

operations per decoded packet.

We use the abovementioned formulae to determine the decoding complexities of MED in comparison to ED through simulation. We define the normalised decoding complexity as the ratio of the number of operations for successful MED to the number of operations for successful ED [19].

Fig. 3 shows the decoding complexity advantage of MED over ED for varying values of n . It can be seen that the decoding complexity of MED is lower than that of ED. It is clear from the analysis of arithmetic operations in (8) – (11) that both methods are of the same order of complexity. There, however, is an improvement of approximately 30% on the number of arithmetic operations required for successful decoding for MED over ED, as shown in Fig. 3.

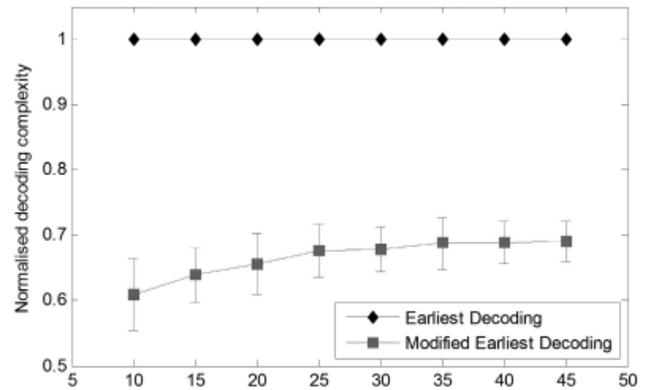


Figure 3: Normalised decoding complexity of MED over ED

From the results obtained it can be seen that MED renders favourable results of an improved decoding complexity. MED also renders a smaller decoding delay when compared to ED.

6. CONCLUSION

The Modified Earliest Decoding algorithm is based on the low complexity belief propagation decoding concept of LT codes and low decoding delay concept of Earliest Decoding.

In this paper we formalised the MED process in Algorithm 1. This algorithm shows how the Hamming distances between coding vectors can be used to obtain native packets for successful decoding.

The performed simulations show a lower decoding delay as well as a lower decoding complexity of MED over ED. When no coding vectors of $d_H = 1$ are present in \mathbf{G} and MED cannot be performed, ED is used and therefore the decoding delay of MED is upper bounded by that of ED for large n .

Earliest Decoding showed a significant improvement on decoding delay and complexity in comparison to Gaussian Elimination [4, 5] when implemented in a RLNC network environment. Our improvement over ED further improves the performance over GE which is the predominant decoding method in RLNC network environment.

7. REFERENCES

- [1] T. Ho, R. Koetter, M. Médard, D. R. Karger and M. Effros, "The benefits of coding over routing in a randomized setting," in Proc. IEEE Int. Symp. on Inform. Theory (ISIT), Yokohama, pp. 442, Jul. 2003.
- [2] J. K. Sundararajan, D. Shah and M. Medard, "Online network coding for optimal throughput and delay – the three receiver case," in Proc. Int. Symp. on Inform. Theory and its Applic., Auckland, New Zealand, 2008.
- [3] R. A. Costa, D. Munaretto, J. Widmer and J. Barros, "Informed network coding for minimum decoding delay," in Proc. IEEE Int. Conf. on Mobile Ad-hoc and Sensor Systems, Atlanta, Georgia, Sept 2008.
- [4] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," in Proc. Allerton Conference on Communication, Control and Computing, Monticello, IL, October 2003.
- [5] P. A. Chou, Y. Wu, "Network coding for the internet and wireless networks," IEEE Signal Processing Magazine, pp. 77-85, Sept 2003.
- [6] H. Wang, J. Goseling, J. H. Weber, "Network coded flooding," M.S. thesis, Dept. of Telecommunications, Delft University of Technology, Delft, June 2009.
- [7] S. von Solms and A. S. J. Helberg, "Modified Earliest Decoding for Random Network Codes," in Proc. The 2011 Int Symp. on Network Coding (NetCod), Beijing, China, 2011.
- [8] D. Petrovic, K. Ramchandran and J. M. Rabaey, "Overcoming untuned radios in wireless networks with network coding," IEEE Trans. on Inf. Theory, vol.52, no. 6, pp. 2649-2657, 2006.
- [9] A. Ramamoorthy, J. Shi, and R. Wesel, "On the capacity of network coding for random networks," IEEE Trans. on Inf. Theory, vol. 51, no. 8, pp. 2878–2885, 2005.
- [10] H. Wang, P. Fan and K. B. Letaief, "Maximum Flow and Network Capacity of Network Coding for Ad-hoc Networks," IEEE Trans. on Wireless Commun. vol. 6, no. 12, pp. 4193-4198, 2007
- [11] C. Avin, "Random Geometric Graphs: An Algorithmic Perspective," Ph.D dissertation, University of California, Los Angeles, 2006
- [12] S. A. Aly, V. Kapoor, J. Meng, and A. Klappenecker, "Bounds on the network coding capacity for wireless random networks," in Proc. 3rd Workshop on Network Coding, Theory, and Applications, San Diego, CA, U.S.A., 2007.
- [13] T. Ho, "Networking from a network coding perspective," PhD Thesis, Massachusetts Institute of Technology, Dept. of EECS, May 2004.
- [14] M. Luby, "LT codes", in Proc. The 43rd Annual IEEE Symp. on Foundations of Computer Science, pp. 271-280, 2002.
- [15] F. Fu, V. K. Wei, R. W. Yeung, "On the minimum average distance of binary codes: linear programming approach" Journal on Discrete Applied Mathematics, pp. 263–281, 2001.
- [16] R. Koetter and M. Medard, "Beyond routing: an algebraic approach to network coding," in Proc. Twenty-First Annual Joint Conf. of the IEEE Computer and Communications Societies, vol. 1, pp. 122-130, 2002.
- [17] D. Goldsman and G. Tokol, "Output analysis: output analysis procedures for computer simulations," in Proc. The 32nd Conf. on Winter simulation (WSC), pp. 39-45, 2000
- [18] L. Lima, D. Ferreira, and J. Barros, "Topology matters in network coding," Springer Telecommunications systems, pp. 1-11, 2011.
- [19] S. Kim, K. Ko and S. Chung, "Incremental Gaussian Elimination Decoding of Raptor Codes over BEC," IEEE Commun. Letters, vol. 14, no. 4, April 2008.

INVESTIGATING THE EFFECTS OF SUBSIDY ON NETWORK RESOURCE UTILIZATION IN MULTI-TIER COMMUNITIES

M. Sumbwanyambe* and A. L. Nel†

* Department of Mechanical Engineering Science, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa, e-mail:sumbwam@gmail.com

† Department of Mechanical Engineering Science, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa, e-mail:andren@uj.ac.za

Abstract: One of the main aims of telecommunication subsidies in developing countries is to extend information and communication services to the information “have nots” through subsidized communication services. However, subsidies may have an impact on network resource utilization, Quality of Service (QoS) and the number of users subscribing to an Internet Service Provider (ISP). For example, subsidies may lead to low QoS and high resource utilization while in some instances unsubsidized services may lead to high QoS and low utilization of resources. This *see-saw effect* may eventually lead to market failure and it may, now and then, destroy market efficiency. Thus far, this phenomenon calls for a combined study, in which the relationship between subsidy, price, QoS and resource utilization is investigated. In this paper, the impact of subsidies on QoS and resource utilization in multi-tier communities is investigated. We try to find a middle ground between implementation of a subsidy discount rate policy for socio-economic development and its effects on QoS and resource utilization in a network.

Key words: ICTs, Pricing policy, Quality of service, Resource utilization, Subsidy

1. INTRODUCTION

Fundamental to some multi-tier communities or developing countries (countries with significant variations in income distributions) is that subsidies should be an important component of their economic system. This is especially true if such communities want to promote Information and Communications Technology (ICT) “access for all” agendas. However, developing countries, especially in sub-Saharan Africa, face greater challenges than most countries in promoting subsidy-driven-ICT-usage while at the same time trying to maximize the social economic welfare of its citizen [1] [2] [3] [4].

These challenges emanate from, and may include, but not be limited to: the size or the group of people receiving subsidies, the subsidy effect on the network QoS and resource utilization. More importantly, however, is the difficulty for institutional frameworks, such as the Universal Service and Access Agency of South Africa (USAASA) and the Zambia Information and Communications Technology Authority (ZICTA), in administering subsidies in developing countries. It is worth noting that, USAASA and ZICTA have struggled with the administration of the Universal Service and Access Fund (USAF) for social and economic growth in unserved areas as noted in the following statements respectively:

“Untill now, USAASA has used very little of the money in the USAF. It must justify its spending plans each year before it is given access to any of the funds” [5].

“Universal access of ICT services with particular reference to rural areas is an issue that has been on the agenda

of most countries in the whole world today. You are aware that while some countries are currently effectively extending services to rural areas using various means, others have found this challenging and have ended up with unutilized funds in the banks. The challenges have emanated from various angles ranging from legal, regulatory to socio-economic dynamics. In Zambia, we have grappled with this issue due to the legal framework” [6].

More often than not, the motivations for subsidies have been criticized for a variety of reasons. For example, the “social welfare function” argument would call for selectivity in subsidization so that only the group that really need it pay the reduced rate. Instead, most countries have applied subsidy schemes where all users gain from the subsidy discount rate, even when they would not need it [7]. Notwithstanding the fact that underserved regions of developing countries are likely to be more expensive to service, it is necessary for us to point out that the extra subsidy may be justified as a means of stimulating social and economic growth in some of these regions. Moreover, the observed variation between the optimal and actual prices emphasizes the importance of differential subsidization among customers particularly in areas where ICT penetration is low.

In a naturally monopolistic market pricing of ICT services at marginal cost would result in a financial deficit. Unless there is a government subsidy to offset this deficit, the ISP must at least break even in order to maintain financial viability. Therefore, subsidization of Internet usage in a heterogeneous community is a necessary and complex

matter that should involve appropriate and well executed economic decisions involving the government, ISPs and consumers [7] [8] [9].

Even though it would be appropriate for the government to provide free Internet access to its underprivileged populace [1] [4], such a provision may, sometimes, result in overutilization of resources which can result in reduced QoS. For instance, when there is substantial subsidization of the price paid by the information “have nots”, users may face no marginal usage costs, resulting in the Internet being treated much like a public good [10]. In instances where a public resource is over-provisioned and the QoS is not affected with use, government intervention or control is not necessary. However, this is clearly not the case with the Internet and we should, therefore, expect it to exhibit similar problems that face other finite resources.

Sumbwanyambe and Nel [3] have shown that the correct subsidy mechanism is a very effective tool in controlling the usage of resources in multi-tier communities. They argue that a well applied subsidy mechanism can provide a resourceful way of ensuring equal distribution of network service, QoS guarantees in networks and the regulation of network resource usage. Yet, an ill conceived subsidy may occasionally affect the allocation of other economic resources especially in multi-tier communities.

In this paper we investigate the outcome of a poorly implemented subsidy discount rate on QoS and resource utilization in network service provision. Furthermore, we try as far as possible to show that a subsidy policy void of other considerations (such as consumer perception of services and reservation price) may not be applicable in multi-tier communities.

This paper is organized as follows: in Section 2 we propose our pricing model describing all the necessary mathematical concepts and formulas; and in Section 3 we evaluate how subsidies affect usage in multi-tier communities; in Section 4 we evaluate the effect of subsidy on QoS and resource utilization in multi-tier communities. Simulation results are presented in Section 5 of this paper and these results show that the subsidy discount rate, which is used mainly in developing countries to promote social and economic growth in underserved areas, may affect the QoS in service provision. Section 6 concludes our study by emphasizing the effects of a subsidy discount rate.

2. THE PRELIMINARIES

Our pricing model reflects a multi-tier community setting, with heterogeneous consumers; the information “haves” (N_2) and “have-nots” (N_1), who have an observable income disparity. Since consumers, especially in multi-tier communities, have differing economic requirements [11] it is necessary for us to adopt a pricing model that will reflect the differing needs of such a community. In [3] Sumbwanyambe and Nel presented a subsidy driven pricing model that incorporates the perceived utility of consumers towards a set price in heterogeneous

communities.

We adopt this model, with no modification, to reflect our price setting and the number of users. It is observed, from the model, that when the price is zero the number of users increases significantly and when the price is very high the number of users decreases significantly. Since the number of users define the number of packets produced in a network, equivalently, we assume the following equation:

$$T_{pkts} = \lambda \left(\left(\frac{\beta p_2}{p_{th1}} \right)^{-\alpha} - 1 \right) + \lambda \left(\left(\frac{p_2}{p_{th2}} \right)^{-\alpha} - 1 \right) \quad (1)$$

Within the current setting, and from equation (1), it is assumed that the N_1 consumers are offered a subsidy so as to promote social and economic growth, while the N_2 users pay the price that is set by the ISP. In this model p_1 and p_2 are the prices payable to the ISP by the heterogeneous consumers N_1 and N_2 respectively, such that $p_1 = \beta p_2$. Furthermore, β is a subsidy discount rate and is in the range of $0 < \beta < 1$ where $p_2 \in (0, \infty)$. In equation (1), λ is the number of packets submitted to the network by the consumers per second, and p_{th1} and p_{th2} are the reservation prices* for the N_1 and N_2 users respectively. In equation (1), α can be described as the perceived utility of consumers given a certain price, where $0 < \alpha \leq 1$. The perceived utility increases with an increasing value of α .

3. UNDERSTANDING THE CONSUMER DECISION PROCEDURE IN NETWORK UTILIZATION

The decision of any consumer, in a multi-tier network, to subscribe to any ISP is dependent on two things: firstly, on the set price and the reservation price [12] of the consumer and; secondly, on the subsidy level provided by the government to the information “have nots”. In any economic market, each customer has a maximum price they are willing to pay for a given product which equals the product’s value to the customer. In this model, we assume that consumers are able to perfectly predict equilibrium prices, meaning that they will only subscribe to an ISP if the price of the network resource is lower than the perceived value of the network resource. This perceived value is the consumer’s reservation price for the product. The consumer compares his reservation price for each product with the purchase price and chooses the network service that offers a positive utility. This model, assumes that raising the price above the reservation price of a consumer will lead to a risk of having no customers subscribing to the ISP. In that case, the ISP will incur a loss since no consumer will be admitted to the network. In all our simulations we assume that the price used is below the reservation price. We provide the following theorem to support our claim:

*The reservation price is the highest price that a given person will accept and still purchase the good. In other words, a person’s reservation price is the price at which he or she is indifferent to purchasing or not purchasing the good [12].

Theorem 1: Suppose the decision of any customer to subscribe to an ISP for any given price p_i for $i \in 1, 2$ is zero for price greater than p_{th1} and p_{th2} , then there exist another price βp_2 and p_2 where the consumer will subscribe to the ISP.

Proof: We shall assume that the decision to subscribe to an ISP by any individual is continuous in the range $p_i \in (0, \infty)$. Considering equation (1), and if we assume that different prices are charged with a probability density function $f_{(p)}$ for each group of users (note that the number of users define the number of packets) then the number of packets in a network can be estimated as:

$$T_{pkts} = \int_0^{\infty} \lambda \left(\left(\frac{\beta p_2}{p_{th1}} \right)^{-\alpha} - 1 + \left(\frac{p_2}{p_{th2}} \right)^{-\alpha} - 1 \right) f_{(p)} dp \quad (2)$$

which can be written as:

$$T_{pkts} = \int_0^{p_{th1}} \lambda \left(\left(\frac{\beta p_2}{p_{th1}} \right)^{-\alpha} - 1 + \left(\frac{p_2}{p_{th2}} \right)^{-\alpha} - 1 \right) f_{(p)} dp + \int_{p_{th1}}^{\infty} \lambda \left(\left(\frac{\beta p_2}{p_{th1}} \right)^{-\alpha} - 1 + \left(\frac{p_2}{p_{th2}} \right)^{-\alpha} - 1 \right) f_{(p)} dp \quad (3)$$

In equation (3), we can see that the second integral is always equal to zero by definition, since no consumer will subscribe to the ISP when the price is equal to or greater than the reservation price. Therefore, the number of packets in the network is zero. The first integral defines p_i , for $p_i \in 1, 2$, to be continuous in the interval $(0, p_{th1})$, such that within the given range the consumer will subscribe to the ISP, resulting in some packet production.

From theorem 1 we can see that if the subsidy discount rate (β) is within the range of $0 < \beta \leq 1$ and the price is within the range of $0 \leq p_1 \leq p_{th1}$, the number of information “have-nots” in such a network can be determined by the value of the subsidy discount rate set by the government and the price set by the ISP. Actually, β also determines the maximum profit that the ISP can receive under no network capacity constraints i.e. the assumption of infinite available bandwidth.

4. SUBSIDY, QOS AND RESOURCE UTILIZATION IN MULTI-TIER COMMUNITIES

Depending on the subsidy discount rate, the price and the reservation price, the total number of information “have nots” will increase or decrease leading either to low or high resource utilization. Usually, resource utilization on any finite resource, such as the network, has a significant impact on the QoS experienced by the consumer [10] [13].

To understand the impact of subsidies on resource utilization and QoS on finite resources, such as in bandwidth constrained networks, one has to consider how

subsidies and price affect the number of users subscribing to an ISP [14]. As an example, consider a group of users, such as the information “have nots” utilizing a finite resource. If we assume that the subsidy discount rate from the government is considerably high ($\beta \approx 0$), then the number of information “have nots” subscribing to the ISP will increase, leading to network congestion and low QoS. The opposite is true if the information “have nots” are not subsidized ($\beta = 1$), and the ISP charges a price close to or beyond the reservation price of the information “have nots” [6].

Generally speaking, if the price of the network resource is above the reservation price then we expect high QoS i.e. the network resource is underutilized [13]. If, on the other hand, the price is below the reservation price, a large number of consumers, the information “haves” and “have nots”, will heavily subscribe to the ISP resulting in an increase in the number of packets produced. This increase in the number of packets produced will then lead to overutilization of the bandwidth resource, which in turn will lower down the QoS in such a network. Evidently, one can see that there is an inverse relationship between resource utilization and QoS, and it creates a *see-saw effect* between the QoS and resource utilization.

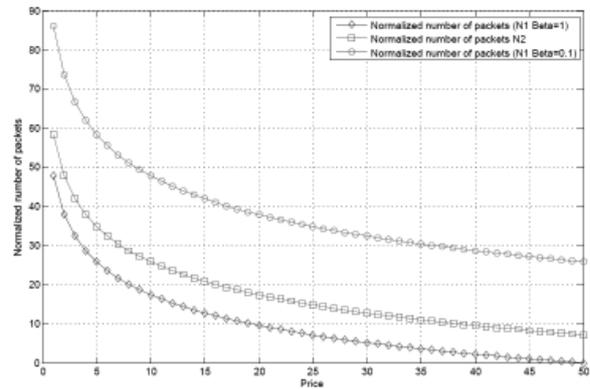


Figure 1: Number of packets versus the price at different subsidy discount rate where $\lambda = 100$, $p_{th1} = 40$ and $p_{th2} = 100$

For example, Figure 1 shows a relationship between packets produced and price, given a government subsidy discount rate of $\beta = 0.1$, $\beta = 1$ and perceived utility $\alpha = 0.1$.

Analysis of the figure will reveal that the number of packets produced fluctuates with an increasing or decreasing subsidy discount rate (β). For instance, for a subsidy discount rate of $\beta = 0.1$, as in Figure 1, the number of packets produced by users will increase significantly due to an increase in the number of users, the opposite is true if there is no subsidy discount rate ($\beta = 1$). We can therefore conclude that, there is an almost hyperbolic relationship between price and the number of users.

From Figure 2, Figure 3 and Figure 4 it can be seen that as α approaches 0 and β approaches 1, the number of packets produced reduces significantly. An analysis of the same

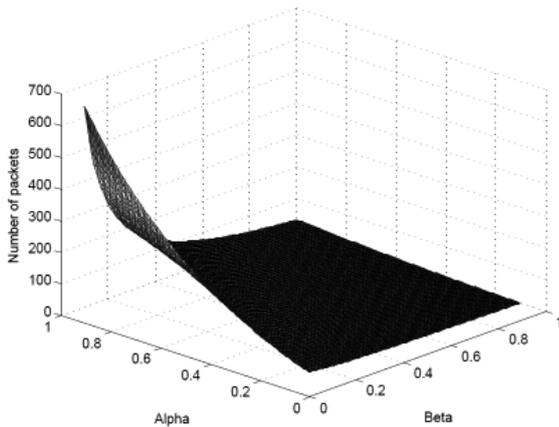


Figure 2: Relationship between expected number of packets versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 60$, $\lambda = 100$

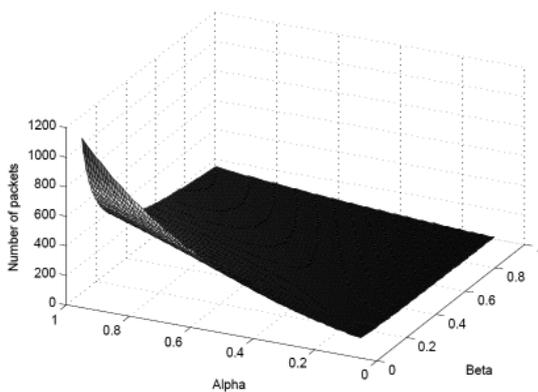


Figure 3: Relationship between expected number of packets and α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 40$, $\lambda = 100$

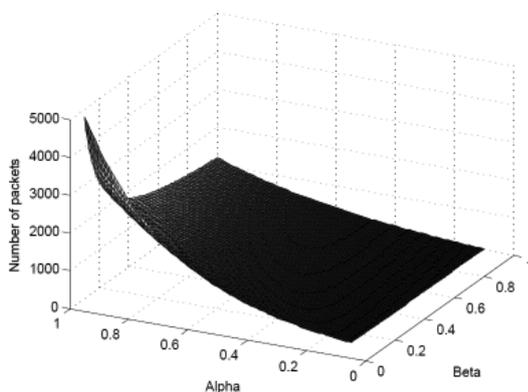


Figure 4: Relationship between expected number of packets versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 10$, $\lambda = 100$

figures will also reveal that as the subsidy discount rate approaches 0 i.e. $\beta \approx 0$ and α approaches 1, the number of packets produced by the users increases considerably.

5. SIMULATION RESULTS

In this section we provide simulation results and evaluate the effect of the subsidy discount rate (β) and perceived utility (α) on resource utilization and QoS in multi-tier communities. We use Matlab to demonstrate such effects.

5.1 Effect of subsidy discount rate (β) and perceived utility (α) on resource utilization

In the field of communications networks, QoS is important if customer satisfaction is to be guaranteed. In general terms, QoS may be defined as the ability to provide a certain level of satisfaction to users given different applications and data flows, or to guarantee a certain level of performance to a packet or data flow. For example, QoS may require that packet rate, delay, jitter, packet dropping probability and/or bit error rate is assured to consumers. Specifically, packet dropping and delay may occur if the total number of packets that are sent over any node in a network exceeds the node capacity or bandwidth of the node.

In our regime of interest, delay and packet dropping may occur if there are too many people sending packets over the Internet creating congestion at many nodes in the network. Such a surge in the increase of consumers in this model could be caused by lower prices or heavily subsidized users of network services in multi-tier communities. In the section that follows, we look at resource utilization and QoS; and analyze how these two are impacted on by the subsidy discount rate and the consumer's perception about the service. Resource utilization (RU) is defined as follows:

$$RU = \frac{\text{Total number of packets arriving at a node}}{\text{Total number of packets that can be transmitted}}$$

Where the total number of packets arriving at the node is taken as shown in equation (1). In all our simulations we assume that the total number of packets that can be transmitted by a node or the bandwidth of the node is 45 units.

From Figure 5, Figure 6 and Figure 7, we can clearly see that as the perceived utility (α) approaches 1, and β approaches 0, the resource utilization in the network increases. It is important to note that the resource utilization in a network is also determined by the price which the ISP charges as displayed in Figures 5, 6 and 7.

Consider, for example, Figure 5 and Figure 6. A close analysis of the figures show that as the price of the network service in multi-tier communities increases, the resource utilization decreases dramatically. Figure 7 displays the situation where the price (p_2) charged by the ISP is greater than the reservation price of the information "have nots". As indicated in Figure 7, the resource utilization of the network decreases significantly compared to Figure 5 when the price (p_2) is less than the reservation price. We

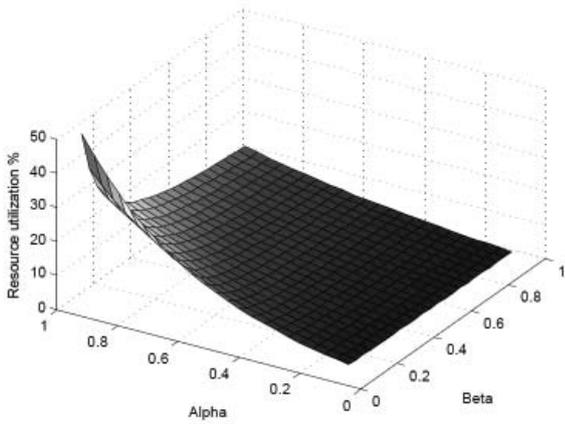


Figure 5: Resource utilization versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 10$, $\lambda = 100$

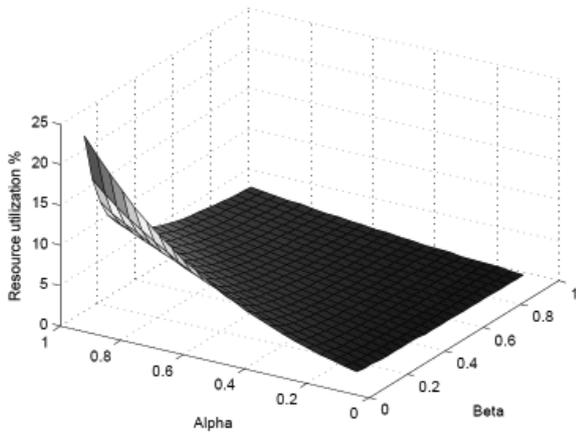


Figure 6: Resource utilization versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 20$, $\lambda = 100$

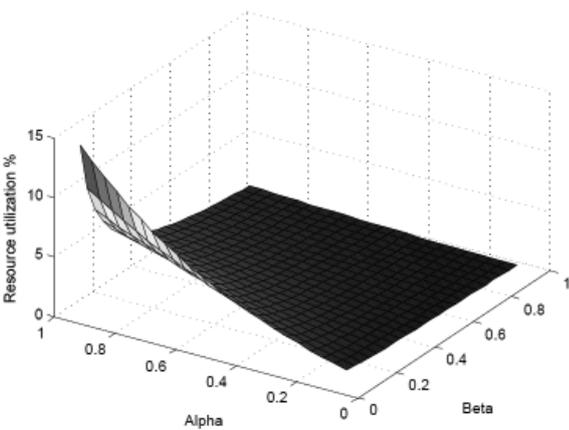


Figure 7: Resource utilization versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 60$, $\lambda = 100$

can thus conclude, from Figure 5, Figure 6 and Figure 7 that the subsidy discount rate (β) and perceived utility (α) have a significant impact on resource utilization developed in the network.

5.2 Effect of subsidy discount rate (β) and perceived utility (α) on QoS

QoS in this paper is defined as the measure of Traffic Load (TL) and it indicates the overall degree of congestion at any particular node in the network. We define TL, an inverse of resource utilization, as follows:

$$TL = \frac{1}{RU}$$

therefore

$$TL = \frac{\text{Total number of packets that can be transmitted}}{\text{Total number of packets arriving at a node}}$$

From Figure 8 to Figure 10 we evaluate the impact of the subsidy discount rate on the QoS in a network.

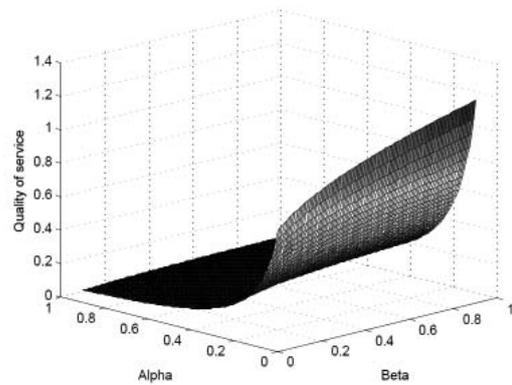


Figure 8: QoS versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 10.7$, $\lambda = 100$

As observed from the figures, we see that the QoS actually decreases with an increase in subsidy discount rate (β) and a decrease in perceived utility (α). Figure 8, for example, shows that at a price of 10.7 units the QoS decreases considerably. The reason being that, at a lower price a large number of users will be willing to pay for Internet services without any government intervention. When the price of network services is actually lower than the reservation price, usage can increase without any subsidy discount rate. Note that as the price increases, however, as shown in Figure 9 and Figure 10, the number of information “have nots” decreases slowly which in turn will result in an increase in QoS, because very few people will be using network services.

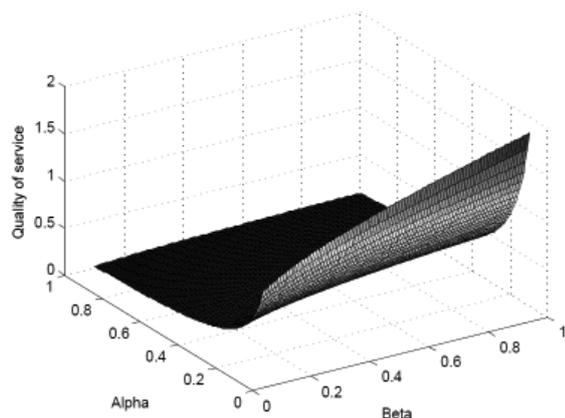


Figure 9: QoS versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 20$, $\lambda = 100$

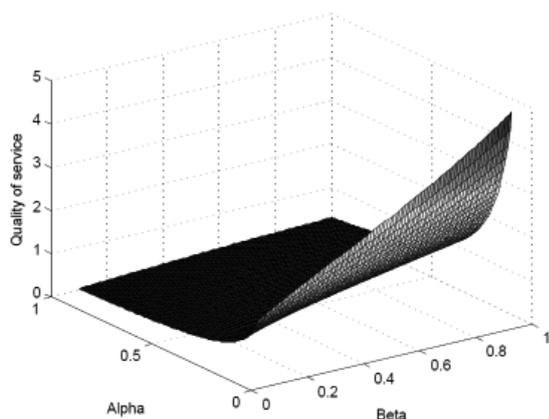


Figure 10: QoS versus α and β : when $p_{th1} = 40$, $p_{th2} = 100$, $p_2 = 40$, $\lambda = 100$

6. CONCLUSION

In this paper, we have analyzed how a subsidy discount rate affects the number of packets produced by users and we have shown that such a relationship has a direct effect on network resource usage. Furthermore, if implemented properly, subsidies can enhance network QoS and resource utilization and can advance social and economic growth in developing countries. In fact, developing countries must consider the tradeoff between resourceful subsidy implementation and its effect on finite resources such as communication networks. In conclusion, the correct usage of subsidies in promoting social and economic targets, can result in the efficient usage of network resources whilst simultaneously guaranteeing QoS and efficient resource utilization to all network users.

REFERENCES

[1] J. Alleman, P. Rappoport and A. Banerjee, "Universal service: A new definition?", in *Telecommunication*

Policy Journal Vol. 34, 2010, pp. 86-91.

- [2] J. Mariscal, "Digital divide in a developing country", in *Telecommunications Policy Journal* Volume 29, Issues 5-6, 2005, pp. 409-428.
- [3] M. Sumbwanyambe and A. L. Nel, "Subsidy and Revenue Maximization in Developing Countries", in *Lecture Notes in Engineering and Computer Science: IMECS*, 2012, pp. 1568-1573.
- [4] B. Ramos, K. Saeed and O. Pavlov, "The impact of Universal Service Obligations and international cross-subsidies on the dispersion of telephone services in developing countries", in *Socio-Economic Planning Sciences* (2009), pp. 1-16.
- [5] Tech Central, "Clarity soon on R1bn telecoms fund", Online document, <http://www.techcentral.co.za/clarity-soon-on-r1bn-telecoms-fund/15674/>.
- [6] M. Sumbwanyambe, "A pricing model for sustainable ICT development in a heterogeneous environment", PhD thesis, Dept. of Mechanical Engineering Science, University of Johannesburg, 2012.
- [7] S. L. Levin, "Universal service and targeted support in a competitive telecommunication environment", in *Telecommunications Policy Journal* Volume 34, 2010, pp. 92-97.
- [8] M. Sumbwanyambe, A. L. Nel and W. A. Clarke, "Challenges and Proposed Solutions Towards Telecentre Sustainability: A Southern Africa Case Study", in *proceedings of IST-Africa*, 2011, Gaborone Botswana.
- [9] C. Garbacz and H. G. Thompson, "Demand for telecommunication services in developing countries", in *Telecommunications Policy Journal*. Vol. 31, 2007, pp. 276-289.
- [10] G. Harding, "The Tragedy of the Commons", in *Journal of Science*, vol. 168, 1968, pp. 1246-1248.
- [11] M. Mandjes, "Pricing strategies under heterogeneous service requirements", in *proceedings of 22nd IEEE Computer and Communications (INFOCOM)*, 2003.
- [12] H. R. Varian, "Intermediate Microeconomics: A Modern Approach", W. W. Norton and Company, New York, London, 6th edition.
- [13] M. A. Heller, "The tragedy of the anti-commons: property in the transition from Marx to markets", in *proceedings of Harvard Law Review*, vol. 3, 1997.
- [14] M. Sumbwanyambe, A. L. Nel and W. A. Clarke, "Optimal pricing of telecomm services in a developing country: A game theoretical approach", in *proceedings of AFRICON*, 2011, pp. 1-6 Livingstone, Zambia.

