SPACOMM 2012

The Fourth International Conference on Advances in Satellite and
Space Communications

ISBN: 978-1-61208-194-6

April 29 - May 4, 2012

Chamonix / Mont Blanc, France

SPACOMM 2012 Editors

Timothy Pham, Jet Propulsion Laboratory, USA

Petre Dini, Concordia University, Canada // China Space Agency Center, China

SPACOMM 2012

Foreword

The Fourth International Conference on Advances in Satellite and Space Communications [SPACOMM 2012], held between April 29th and May 4th, 2012 in Chamonix / Mont Blanc, France, continued a series of events attempting to evaluate the state of the art in academia and industry on the satellite, radar, and antennas based communications bringing together scientists and practitioners with challenging issues, achievements, and lessons learnt.

Significant efforts have been allotted to design and deploy global navigation satellite communications systems. Satellite navigation technologies, applications, and services still experience challenges related to signal processing, security, performance, and accuracy. Theories and practices on system-in-package RF design techniques, filters, passive circuits, microwaves, frequency handling, radars, antennas, and radio communications and radio waves propagation have been implemented. Services based on their use are now available, especially those for global positioning and navigation. For example, it is critical to identify the location of targets or the direction of arrival of any signal for civilians or on-purpose applications; smarts antennas and advanced active filters are playing a crucial role. Also progress has been made for transmission strategies; multiantenna systems can be used to increase the transmission speed without need for more bandwidth or power. Special techniques and strategies have been developed and implemented in electronic warfare target location systems.

We take here the opportunity to warmly thank all the members of the SPACOMM 2012 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to SPACOMM 2012. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the SPACOMM 2012 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that SPACOMM 2012 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of satellite and space communications.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed their stay in the French Alps.

SPACOMM Advisory Committee:

Stelios Papaharalabos, ISARS/National Observatory of Athens, and Athens Information Technology (AIT), Greece
Piotr Tyczka, Poznan University of Technology, Poland
Michael Sauer, Corning Cable Systems, USA
Ling Pei, Finnish Geodetic Institute, Finland

SPACOMM 2012 PROGRAM COMMITTEE

SPACOMM Advisory Committee

Stelios Papaharalabos, ISARS/National Observatory of Athens, and Athens Information Technology (AIT), Greece
Piotr Tyczka, Poznan University of Technology, Poland
Michael Sauer, Corning Cable Systems, USA
Ling Pei, Finnish Geodetic Institute, Finland

SPACOMM 2012 Technical Program Committee

Markos P. Anastasopoulos, National Technical University of Athens, Greece
Loredana Arienzo, Joint Research Centre, European Commission - Varese, Italy
Iva Bacic, University of Zagreb, Croatia
Marco Baldi, Università Politecnica delle Marche - Ancona, Italy
Mark Bentum, University of Twente & ASTRON, The Netherlands
Igor Bisio, University of Genoa - Italy, Italy
Shkelzen Cakaj, Telecom of Kosovo / Prishtina University, Kosovo
Enzo Alberto Candreva, University of Bologna, Italy
Emmanuel Chaput, IRIT-CNRS, France
Bruno Checcucci, Perugia University, Italy
Vittorio Dainelli, Rheinmetall Italia S.p.A. - Rome, Italy
Leonardo Dagui de Oliveira, Escola Politécnica da Universidade de Sao Paulo, Brazil
Francescantonio Della Rosa, Tampere University of Technology, Finland
Felix Flentge, ESA/ESOC HSO-GIB - Darmstadt, Germany
Thierry Gayraud, LAAS-CNRS, Université de Toulouse, France
Mathieu Gineste, Thales Alenia Space, France
Christian Ibars, Centre Tecnologic de Telecomunicacions de Catalunya - CTTC - Castelldefels, Spain
Massimiliano Laddomada, Texas A&M University - Texarkana, USA
Simona Lohan, Tampere University of Technology, Finland
Krešimir Malaric, University of Zagreb, Croatia
Herwig Mannaert, University of Antwerp, Belgium
Emmanouel T. Michailidis, University of Piraeus, Greece
Marina Mondin, Politecnico di Torino, Italy
Brian Niehöfer, Technische Universität Dortmund, Germany
Pasquale Pace, DEIS - University of Calabria, Italy
Stelios Papaharalabos, ISARS/National Observatory of Athens, and Athens Information Technology (AIT), Greece
Ling Pei, Finnish Geodetic Institute, Finland
Cathryn Peoples, University of Ulster - Coleraine, UK
Dionysia K. Petraki, National Technical University of Athens, Greece
Timothy Pham, Jet Propulsion Laboratory / California Institute of Technology, USA
Prashant Pillai, University of Bradford, UK
Luigi Portinale, Università del Piemonte Orientale "A. Avogadro" - Alessandria, Italy
Ronald Raulefs, German Aerospace Center, Germany
Vincent Roca, INRIA Rhone-Alpes, France
Alexandru Rusu-Casandra, Politehnica University of Bucharest, Romania

Heung-Gyoon Ryu, Chungbuk National University, Republic of Korea
Michael Sauer, Corning Cable Systems, USA
Ana Maria Sierra Diaz, Telefónica I+D - Madrid, Spain
Danai Skournetou, Tampere University of Technology, Tampere, Finland
Heidi Steendam, Ghent University, Belgium
Sarang Thombre, Tampere University of Technology, Finland
Piotr Tyczka, Poznan University of Technology, Poland
Zhiwen Zhu, Communications Research Centre Canada, Canada

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Real Time Algorithm for Bandwidth and Time-slot Assignment for Rain Faded DVB-RCS Systems

Mohamed Al-mosawi, Rinat Khusainov, Boris Gremont

School of Engineering, University of Portsmouth

Portsmouth, UK

{mohamed.al-mosawi, rinat.khusainov, boris.gremont}@port.ac.uk

Abstract—Broadband satellite communication networks, operating at Ka band and above, play a vital role in today's worldwide telecommunication infrastructure. The problem, however, is that rain can be the most dominant impairment factor for radio propagation at these frequencies. This paper, addresses bandwidth and time-slot allocation problem for rain faded DVB-RCS satellite networks. We formulate the task as a combinatorial optimization problem and propose a novel algorithm for dynamic bandwidth and time allocation, which works with CRA type of traffic. The algorithm is evaluated using a MATLAB simulation with historical rain data for the UK.

Index Terms—MF-TDMA, time-slot allocation, fade mitigation techniques, combinatorial optimization.

I. INTRODUCTION

VSAT networks provide high-speed multimedia services, including voice, video and Internet for a vast number of subscribers distributed over a very wide area. The DVB-RCS stands for Digital Video Broadcast Return Channel via Satellite, which is a centralized communication system. There is a single station called Network Control Centre (NCC) that controls all communication processes in real time.

European DVB-RCS systems providers use the following access mechanisms recommended by the European Telecommunications Standards Institute (ETSI). TDMA is used on the forward link as the access mechanism, while the return link is shared by earth stations using an MF-TDMA scheme. Typically, Ku band (12-18 GHz) is used for the forward link and Ka band (18-30 GHz) for the return link [1]. Forward link transmissions for user terminals are organized in bursts. The bursts are assumed to be made of a fixed number of time-slots, which are long enough to transmit one fixed-size packet.

Satellite links can be affected severely by rain fading, which can reduce the link capacity. NCC has to make sure that extra time slots or extra bandwidth are available to provide users with the requested quality of service in varying weather conditions. Consequently, efficient Radio Resource Management (RRM) and bandwidth utilization under rainy conditions have become important research topics in satellite telecommunications.

Combinatorial optimization has been used to address the problem of efficient resource allocation in DVB-RCS satellite systems. In [2], resource allocation for rain faded forward links has been modeled as a knapsack problem. In this paper,

we look into resource allocation for DVB-RCS return links. The difference is that here we have to work with MF-TDMA frames instead of TDMA.

The main contributions of this paper are as follows:

- We formulate the task of bandwidth allocation under rainy conditions mathematically as a combinatorial optimization problem.
- We propose a low-cost allocation algorithm to solve it. The main objective of the algorithm is to achieve a better bandwidth utilization while still meeting the Quality of Service (QoS) requirements.
- Finally, we evaluate the algorithm using a MATLAB simulation with real historical rain data for the UK.

The current version of the algorithm deals with the guaranteed throughput services. The guaranteed throughput services are defined as the services, which ensure that the subscriber always gets the bandwidth requested regardless of traffic behavior of other users [1].

The rest of the paper is structured as follows. Section II describes briefly the Bandwidth on Demand process and MF-TDMA frame. The allocation process is mathematically formulated as a combinatorial optimization problem. Section III describes the proposed allocation algorithm. Section IV presents the simulation scenario and results. Finally, section Vdiscusses directions for future work.

II. MF-TDMA FRAME AND BANDWIDTH ON DEMAND

The DVB-RCS NCC sends the general network information to Return Channel via Satellite Terminals (RCSTs) once the power is on. The NCC provides monitoring and control functions and generates the control and timing messages required for operating the satellite network. The messages are sent using Moving Picture Exported Group Transport Stream (MPEG-2TS) via the private data section, which is transmitted over the forward link. There are two types of forward links in the DVB-RCS specification: one for the interaction control and the second for data transmission. Both links can be supported by the same DVB-S transport multiplex [3].

MF-TDMA bandwidth portion is usually defined to be approximately 20 MHz due to the RCST maximum frequency hoping limitation [3][4]. The goal of the Bandwidth on Demand (BoD) process is to determine the resources required

Symbol	Definition
U	The total number of RCSTs supported by the network.
β	Buffer size (in packets) for a given RCST.
M	The total number of packets that can be buffered in the network, which is equal to the size of the RCST buffer multiplied by the total number of RCST supported by the network (all RCSTs have the same buffer size).
K	The number of available ACM modes in the network.
r_i	The rain fading level for RCST i.
N_k	The number of packets suffering rain fading of level k.
$f(x)$	The space required for a time slot with rain fading level of x in the MF-TDMA frame.
p_i	Number of time-slots requested by RCST i.
C	The available MF-TDMA space. The space here stand for a portion of bandwidth given for a duration of time in a single frame for data traffic.
B_k	The bandwidth assigned by BoD to mitigate level k rain effects.
L	The number of available MF-TDMA frames.
T	The duration of an MF-TDMA time-slot.
R_i	Requested data rate by RCST i..

Table I
LIST OF SYMBOLS

to satisfy the rate requested by an RCST. A BoD controller receives RCST data requests, selects an appropriate Adaptive Coding and Modulation (ACM) to match their rain level, and calculates the required bandwidth. The Media Access Control (MAC) then decides whether to accept or reject the request based on a set of rules as explained in [3].

Network traffic can be characterized as a Markov-modulated continuous stream of bits with peak and mean rates. On-off traffic sources with variable bit rates can be modeled as constant flow sources with an equivalent bandwidth [5]. Thus, we work under a simplifying assumption that the network uses the Constant Rate Allocation (CRA) to eliminate traffic complexity from the analysis. Consequently, the real-time bandwidth allocation will be governed purely by rain fading. There are several approaches to determining the equivalent bandwidth for the CRA. The fluid flow approach has been well investigated in [1][5][6]and[7]. In this paper, we use the normal "Gaussian" distribution investigated in [1-8].

III. TIME ALLOCATION PROBLEM

A. Mathematical Formulation

The problem of resource allocation with rain fading can be defined as a combinatorial optimization problem. The goal is to pack optimally packets of different sizes into the MF-TDMA frame space. Table 1 describes the initial parameters for the optimization process.

At any time instant, let us define vector $\mathbf{p}$, as

$$\mathbf{p} = (p_i, ..., p_u \mid p_i \leq \beta \,, i = 1, 2, ..., U), \qquad (1)$$

where p_i is the number of packets buffered by RCST i, which is always smaller or equal to the RCST buffer size for all stations. Similarly, vector$\mathbf{r}$ containing the instantaneous rain fading levels for all RCSTs is defined as

$$\mathbf{r} = (r_i, ..., r_u \mid r_i \leq K \,, i = 1, 2, ..., U), \qquad (2)$$

where r_i is the rain fading level for RCST i. There can be many constraints involved in determining how to pack a given set of time-slots, and there are many objectives to be met, such as fairness, efficiency, QoS, and utilization. In this paper, the main objective is to transmit the packets using the smallest possible portion of the available MF-TDMA frame space. This problem can be described as follows:

$$maximize \quad \frac{\sum_{i=1}^{U} \left(f(r_i) \cdot \sum_{j=1}^{p_i} x_{ij} \right)}{C \cdot L} \qquad (3)$$

$$s.t. \sum_{j=1}^{p_i} x_{ij} = R_i \cdot \frac{T}{bits\,per\,slot}, \; \forall i \qquad (4)$$

$$where, \; x_{ij} = \begin{cases} 1, if\,the\,packet\,is\,selected. \\ 0, otherwise. \end{cases} \qquad (5)$$

The last constraint essentially specifies that each RCST should be able to transmit the requested number of packets within L MF-TDMA frames.

The secondary objective is to minimize the variation in the delays experienced by individual packets, also known as cell delay variation (CDV) [1].

B. The Algorithm

In any given time interval, there is a number of packets suffering K or less levels of rain fading, which are needed to be fitted into L consecutive MF-TDMA frames of capacity c each. For our algorithm, the total demand can be characterized as a $M \times K$ matrix W defined as follows:

$$w_{mk} \in \begin{cases} 1, if\,k=r_i, and\,m\,mod(i*\beta) \leqslant p_i. \\ 0, otherwise. \end{cases} \qquad (6)$$

Each column corresponds to a different level of rain fading k. Each row corresponds to the packet sequence number in an RCST buffer. Packets from RCST 1 correspond to rows $1, 2, ..., \beta$, packets from RCST 2 correspond to rows $\beta+1, \beta+2, ..., 2\beta$. Packets from RCST$i$ correspond to rows $(i-1) \cdot \beta + 1, (i-1) \cdot \beta + 2, ..., 2(i-1) \cdot \beta$. Therefore, if RCST i suffers rain fading level of k, then the corresponding rows in column k of the demand matrix will be equal to 1, otherwise they will be equal to 0. The total number of '1's in each column of the demand matrix is equal to N_k and each row contains at most single '1'. Figure 2 illustrates this idea. The demand matrix can be used by the MAC to perform fitting of packets into MF-TDMA frames.

The main idea behind the proposed algorithm is to bundle together packets coming from RCSTs with the same rain fading level (the same ACM mode). This can be done by allocating carriers in MF-TDMA frame separately for each column in the demand matrix according to the actual number of packets in that column. The algorithm is specified in Figure 1.

The constraint in equation 7 is to ensure not to exceed the total frame bandwidth, where B_k defined as the bandwidth assigned by the BoD to mitigate level k rain effects. The packing procedure tries to reduce the CDV without violating any agreed QoS bounds. To achieve this, the procedure schedules

---

Step 1 (Initialization)
  Create the demand matrix W
Step 2 (Iteration)
  for $k = 1...K$ (processing each demand matrix column separately)
    Calculate the required number of carriers C_k:

$$C_k = \left\lceil \frac{\sum_{m=1}^{M} w_{mk}}{(number\ of\ time\ slots\ per\ carrier) \cdot L} \right\rceil$$

$$C_k = \left\lceil \frac{N_k}{10 \cdot L} \right\rceil$$

$$s.t.\ C_k \cdot B_k \leq available\ bandwidth \qquad (7)$$

    Pack packets from column k of the demand matrix into the C_k carriers
    of bandwidth B_k.
    Update available bandwidth (reduce by $C_k \cdot B_k$).
  end for
Step 3
  Delete W

Figure 1.   Allocation algorithm

packets from the RCST with the longest packet delay first. This ensures that the packets with earlier deadlines are scheduled into earlier time slots within the L-frame window. Scheduling packets from RCSTs with the same rain fading level into of carriers of the same bandwidth reduces the amount of unused space in each MF-TDMA frame.

There may still be a situation when a portion of a carrier in a frame is unused because there are not enough packets at the given rain fading level to fill in all the time slot in that carrier. In the proposed algorithm, we introduce a time window to further improve carrier utilization. The time window is defined as the time interval required to transmit L MF-TDMA frames. Within this time period, requests will be stored in RCSTs buffers to be transmitted in the next time window. The algorithm looks at requests after each window. Therefore, instead of generating a Terminal Burst Time Plan (TBTP) for each super frame, there will be one TBTP generated for each time window.

Having the time window helps capturing more packets at a given rain fading level, which can then be transmitted within the carrier of an appropriate bandwidth. This means that less frames may be required to transmit them. The freed up frames in a given time window can be used to carry traffic from users with best effort level of services. For example, instead of transmitting one packet per carrier in 10 frames, the 10 packets can be buffered and transmitted in the next time window within a single frame. That means we will not need to reserve bandwidth for this rain fading level in the other 9 frames, and freed up bandwidth will be available for other services. This improvement in bandwidth utilization comes at the price of an added delay, which can be given in the following formula:

$$delay\ bound = L \cdot frame\ duration \qquad (8)$$

A downside of using this algorithm is that it will delay and queue data even for unfaded satellite links. One possible way to address it is to use a dynamic window size dependent

on the rain fading conditions as well as burstiness of traffic. The algorithm can be illustrated graphically using Figure 2, which shows how traffic from different RCSTs is grouped according to their rain fading levels. Each group of packets then is sequentially scheduled by MAC into the corresponding carriers across the MF-TDMA frames in the time window.
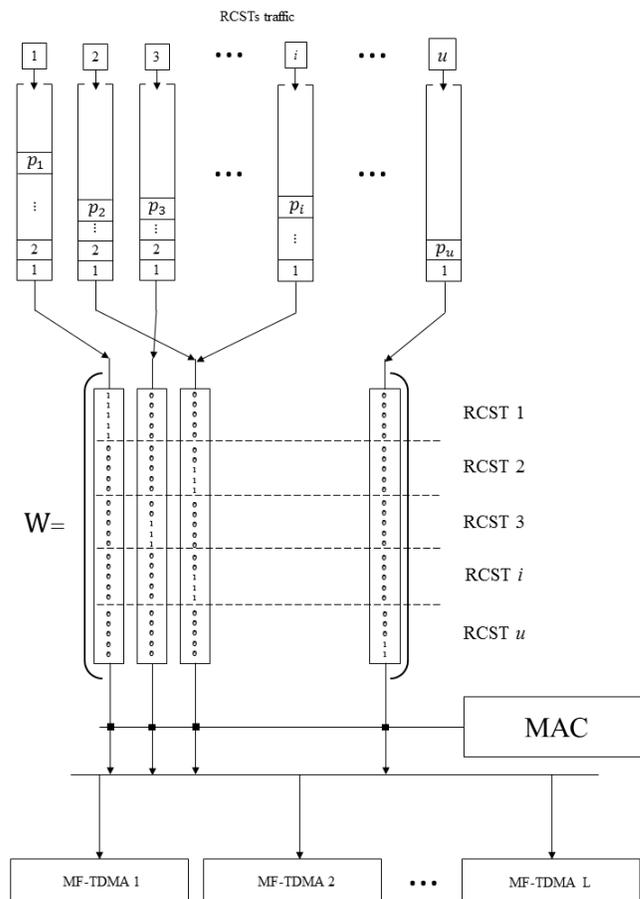
Figure 2.   MAC serving K packets groups into L frames

IV. SIMULATION MODEL

The data rate that has been used is 144 Kb/s (topology A) guaranteed throughput of a low rate video streaming traffic type, with burstiness of 8. In the simulation, we assume the burstiness is measured over a single time window. More specifically, we assume that each RCST transmits at peak rate only for 1/8th of the time window duration. Thus, a total of 10 packets will be received in each time window. The MF-TDMA frame is be sub-divided in 60 standard carries with 10 time slots each, as shown in table 2 [8]. Each time slot has a capacity of 16 Kb/s guaranteed throughput. Therefore, 10 time slots per time window must be allocated to each RCST to satisfy the targeted rate. The above numbers were obtained from calculation carried in [8],[3], and [4]. In this version of the simulator, we considered one data type "video streaming" over one MF-TDMA of 20 MHz bandwidth, with transmission frequency of 20 GHz.

Peak information data rate	Slots per carrier and per frame				Carriers per frame
	traffic slots	CSC/ slots	SYNC/ slots	Total slots	
144 Kb/s	9	1	2	10	60
384 Kb/s	24	2	4	26	23
1024 Kb/s	64	4	8	68	9
2048 Kb/s	128	8	16	136	4

Table II
FEATURES OF THE MF-TDMA RETURN LINK

In the given simulation, 480 RCSTs were used with the following distribution pattern: 75% of the RCSTs were located in 4 different cities: London, Manchester, Dublin, and Glasgow, and the rest was randomly distributed all over the UK with the minimum separation distance of 1 km between RCSTs as illustrated in Figure 3.

The algorithm was applied on version 1 DVB-RCS system with 9 ACM FMT modes available, i.e., $K = 9$ in our case. The simulation uses historical rain data, with a simulation duration of 3 months (January, February and April 2011).
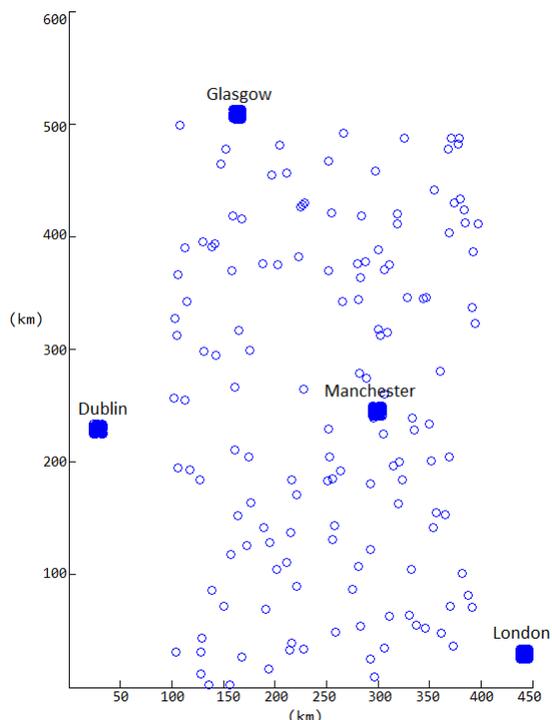
Figure 3.   RCSTs ground distribution

A. Numerical Results and Performance Evaluation

Figure 4 shows the amount of used bandwidth for the given distribution of RCSTs. The bottom level of the used bandwidth corresponds to the minimum rain fading level (i.e. no rain). The overall bandwidth usage statistics can be represented as a CDF of total bandwidth used. This is shown in Figure 5.

As can be seen from the results shown in Figure 4 and 5, only up to 40% of the total allocated frame bandwidth was used at any moment in time. Moreover, for a standard 99.9% network availability only 35.2% of the frame bandwidth was actually used by the RCSTs.
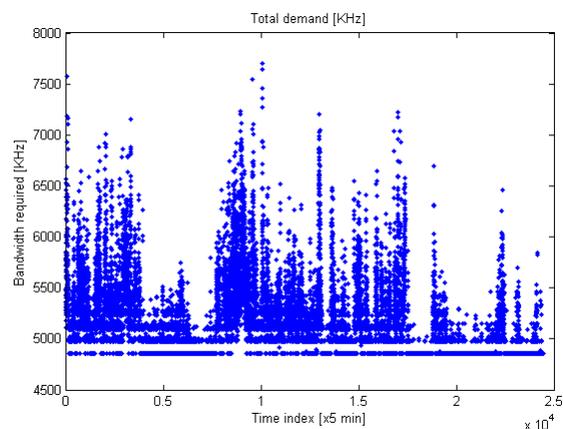
Figure 4.   Total bandwidth used by the algorithm
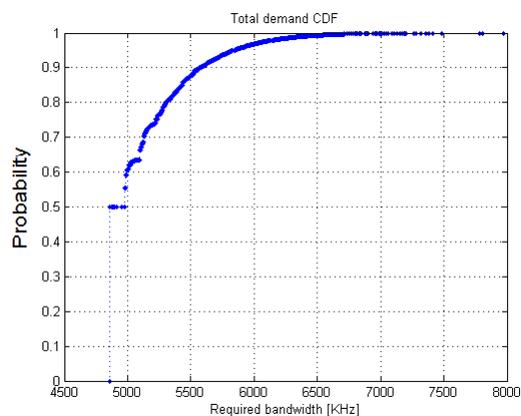
Figure 5.   CDF for the total bandwidth used by the algorithm

V. CONCLUSION

The task of bandwidth and time slot assignment in rain faded DVB-RCS systems has been described and formulated mathematically as a combinatorial optimization problem. Unlike other referenced approaches, where RCSTs are divided into only two groups of rainy and non-rainy terminals, the proposed algorithm bundles traffic requests into multiple groups corresponding to their individual rain fading levels. The algorithm also uses a time window, which consists of multiple consecutive MF-TDMA frames, to improve time slot utilization within each carrier and free up frame bandwidth for other services. Our evaluation of this algorithm using a simulation with historical rain data shows a significant improvement in bandwidth utilization. Only 35.2 to 40% of the available frame bandwidth was used at any given time.

In the future work, we will be studying the effects of geographical distribution of RCSTs on the total demand. The current version of the algorithm can also be expanded to include different types of services and address the delay bound issue as one of the primary objectives.

REFERENCES

[1] Chao, H. J., & Guo, X. (2002). Quality of service control in high-speed networks. New York, John Wiley.

[2] Y. Feng, "Resource allocation in Ka-band satellite systems," Master. dissertation, Dept. of Elecrical and Computer Engineering, University of Maryland, 2001.

[3] E. Noussi, "A study of fade mitigation and resource management of satellite networks in rain condition" Ph.D. dissertation, Dept. of Eelectroic and Computer Enginnering, University of Portsmouth, June. 2008.

[4] Ki-Dong Lee; Kun-Nyeong Chang; , "A real-time algorithm for timeslot assignment in multirate return channels of interactive satellite multimedia networks," Selected Areas in Communications, IEEE Journal on , vol.22, no.3, pp. 518- 528, April 2004 doi: 10.1109/JSAC.2004.823433.

[5] Guerin, R.; Ahmadi, H.; Naghshineh, M.; , "Equivalent capacity and its application to bandwidth allocation in high-speed networks," Selected Areas in Communications, IEEE Journal on , vol.9, no.7, pp.968-981, Sep 1991 doi: 10.1109/49.103545 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=103545& isnumber=3202.

[6] S. Jamin, "A Measurement-based admission control algorithm for integrated services packet networks," Ph.D. dissertation, Dept. of Computer Science, University of Southern California, Aug. 1996.

[7] Genin, D.; Marbukh, V.; , "Bursty fluid approximation of TCP for modeling internet congestion at the flow level," Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on , vol., no., pp.1300-1306, Sept. 30 2009-Oct. 2 2009 doi: 10.1109/ALLERTON.2009.5394525 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5394525& isnumber=5394483.

[8] Pace, P.; Aloi, G.; Marano, S.; , "Multimedia traffic admission schemes comparison for satellite systems," Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th , vol.6, no., pp. 4022- 4026 Vol. 6, 26-29 Sept. 2004 doi: 10.1109/VETECF.2004.1404833 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1404833& isnumber=30415.

[9] T. Kiong, "Connection admission control in multi service networks: simplicity versus efficiency," Ph.D. dissertation, Dept. of Electrical and Electronic Engineering, University of Melbourne, Nov. 2002.

[10] S. Floyd, "Comments on measurement-based admissions control for controlledload service," ftp:rrftp.ee.lbl.govrpapersradmit.ps.Z.

[11] De Rango, F.; Tropea, M.; Fazio, P.; Marano, S.; , "Call admission control with statistical multiplexing for aggregate MPEG traffic in a DVB-RCS satellite network," Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE , vol.6, no., pp.6 pp.-3236, 2-2 Dec. 2005 doi: 10.1109/GLOCOM.2005.1578372 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1578372& isnumber=33290.

[12] De Rango, F.; Tropea, M.; Fazio, P.; Marano, S.; , "Call Admission Control for Aggregate MPEG-2 Traffic Over Multimedia Geo-Satellite Networks," Broadcasting, IEEE Transactions on , vol.54, no.3, pp.612-622, Sept. 2008 doi: 10.1109/TBC.2008.2002716 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4604798& isnumber=4604792.

A Practical Simulation Tool for Predicting Spectral Regrowth
in Communication Satellite Receivers

Veli Yanikgonul, Lokman Kuzu, Ahmet F. Yagli,
Mesut Gokten

R & D - Satellite Design Dept.
Turksat Satellite Communications AS
Ankara, TURKEY
vyanikgonul@turksat.com.tr, lkuzu@turksat.com.tr,
afyagli@turksat.com.tr, mgokten@turksat.com.tr

Murat H. Sazlı

Electronics Engineering Dept.
Ankara University
Ankara, TURKEY
sazli@eng.ankara.edu.tr

Abstract— **The intermodulation distortion phenomena caused by nonlinear characteristic of active RF equipments is a significant source of system impairment in satellite communications. Telecommunications signals are usually composed of more than one carrier modulated by information signal. When number of carriers passing through a nonlinear medium increases, which is the case in communications satellite receivers, the prediction of the intermodulation distortion becomes more and more complex. Considering the scarcity of frequency spectrum and other resources in orbit, satellite channel performances should delicately be examined in order to ensure efficient and uninterrupted communications through all channels and avoid any channel failure due to intermodulation phenomena. This paper presents a practical tool for calculating and simulating the effect of intermodulation signal distortion and C/I performances of the broadband satellite receivers placed at the communication satellite payload input section. With this tool, the effect of a collocated satellite system operating in an adjacent frequency band can also be easily calculated.**

Keywords- Intermodulation distortion; spectral regrowth; satellite communications.

I. INTRODUCTION

A significant amount of communications traffic is handled by communications satellites and therefore these systems fulfill a huge amount of the communication needs in today's world. Broadband receivers are used on-board satellites in order to operate over a frequency range of a few hundred MHz usually, serving a maximum number of carriers to achieve high spectral efficiency.

Multiple signals simultaneously arriving at the receiver generate intermodulation noise due to the nonlinear characteristics of the receiver itself. This noise is generated by a large number of intermodulation products which interfere with the useful carriers. Due to the existence of these products, the amplitudes of the useful carriers are limited and the system efficiency is decreased. In the frame of these arguments, simulation and prediction of intermodulation distortion is a crucial issue for arranging frequency plan of satellite communication systems.

There are several manifestations on the algorithms and techniques developed for modeling and simulating intermodulation products [1]-[6]. Fast Counting Algorithm (FCA) assumes a number of identical carriers distributed over a frequency range with equally spaced frequency positions, and the distortion is dominated by the third-order intermodulation products of $f_i+f_j-f_k$ type, neglecting the ones of $2f_i-f_j$ type [1]. FCA employs FFT algorithm together with indicator polynomials modeling the carrier configuration. Higher order intermodulation products and harmonics are also computed by extending the FCA algorithm [2]. In [3], the third-order intermodulation spectrum is calculated by convolution of the input power spectrum using the FFT method. Another alternative method suggests noise band representation replacing identical contiguous narrow band carriers with an equivalent band of thermal noise [4]. In this method, intermodulation products are restricted to $f_i+f_j-f_k$ type. Another popular technique is modeling with Volterra series. Discrete third-order Volterra model is used in [5] and [6] to compute third-order products. In [7], Volterra-Wiener technique is adopted for this purpose.

In this work, our first goal is to model and simulate intermodulation distortion taking place in communication satellite receivers with a fair accuracy and complexity. To achieve this goal, polynomial model is employed together with a counting algorithm, and only two types of intermodulation products, namely $f_i+f_j-f_k$ and $2f_i-f_j$ types, are taken into account. This is because these two types of products are dominating the in-band intermodulation distortion [8]. Secondly, we are aiming to design this simulation tool in a way that it will be useful for satellite communications systems' designers in the industry. In this respect, the inputs of the simulation tool are the uplink frequency plan of the satellite and some characteristics of the equipments. The outputs are the total third-order intermodulation noise power and carrier to intermodulation distortion power ratio (C/I), given for each predefined satellite channel, which are going to assist the designers to figure out the system linearity performances. Considering the importance of the risk management in the communication satellite industry, such tools can have a critical role in the AIT (Assembly, Integration and Test) phase.

II. MODELING THIRD-ORDER IMPs

A. Polynomial Model

Polynomials have been used extensively to model nonlinearities for many applications. The polynomial model, also called as power series representation in literature, is a simple and highly effective technique for modeling nonlinear behavior. It provides freedom of design in choosing a higher-degree polynomial for higher accuracy, or a lower-degree polynomial for less computational complexity.

The output $y(t)$ of a nonlinear system with respect to an input signal $x(t)$ can be modeled by a polynomial function as

$$y(t) = \sum_{i=1}^{\infty} a_i x^i(t) = a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \cdots \quad (1)$$

where a_i represents nonlinearity coefficients that characterizes the system's behavior. The second, third and higher order terms of (1) explain the spectral regrowth.

B. Two-Tone Input Model

Assume an input signal composed of two carriers with the same amplitude and at different frequencies, as shown in (2). The phase values are taken to be zero for simplicity. In the output spectrum, there are elements originated from interaction of these two tones, which are called "intermodulation products", as well as the "harmonics" of the fundamental frequencies.

$$x_{2\_tone}(t) = A\cos(\omega_1 t) + A\cos(\omega_2 t) \quad (2)$$

Substituting (2) in the polynomial model which is truncated to third order gives:

$$y(t) = \left(a_1 A + \frac{9}{4}a_3 A^3\right)\cos(\omega_1 t) + \left(a_1 A + \frac{9}{4}a_3 A^3\right)\cos(\omega_2 t) \quad (3)$$
$$+ a_2 A^2 \cos((\omega_2 \pm \omega_1)t) + \frac{1}{2}a_2 A^2 \cos(2\omega_1 t)$$
$$+ \frac{1}{2}a_2 A^2 \cos(2\omega_2 t) + \frac{3}{4}a_3 A^3 \cos((2\omega_1 \pm \omega_2)t),$$
$$+ \frac{3}{4}a_3 A^3 \cos((2\omega_2 \pm \omega_1)t) + \frac{1}{4}a_3 A^3 \cos(3\omega_1 t)$$
$$+ \frac{1}{4}a_3 A^3 \cos(3\omega_2 t) + a_2 A^2$$

The 3rd order intermodulation products have an in-band distortion effect whereas the harmonics gather at distinct frequencies constituting out-of-band distortion. In fact, all odd order intermodulation products appear exactly over, or very close to the output fundamental frequencies.

Signal to third-order intermodulation distortion ratio (C/I_3) is an important figure of merit of a system's nonlinearity described by two-tone characterization method.

It is defined as the ratio between one of the output fundamental's power and the power of the nearest third-order intermodulation product, given as

$$\frac{C}{I_3} = \frac{P(\omega_1)}{P(2\omega_1 - \omega_2)} = \frac{P(\omega_2)}{P(2\omega_2 - \omega_1)} = \frac{\left(a_1 A + \frac{9}{4}a_3 A^3\right)^2}{\left(\frac{3}{4}a_3 A^3\right)^2} \quad (4)$$

C. Three-Tone Input Model

Three-tone response should also be investigated to be able to compare both $f_i + f_j - f_k$ and $2f_i - f_j$ types of products. With the same assumptions made for two-tone response, the input signal is written as

$$x_{3\_tone}(t) = A\cos(\omega_1 t) + A\cos(\omega_2 t) + A\cos(\omega_3 t) \quad (5)$$

Injecting (5) in the polynomial model, the resultant third-order products' amplitudes are given below:

$$IM_{2fi-fj} = \frac{3}{4}a_3 A^3 \quad (6)$$

$$IM_{fi+fj-fk} = \frac{3}{2}a_3 A^3 \quad (7)$$

Thus, the power of each $f_i + f_j - f_k$ type product is approximately 6 dB higher than that of $2f_i - f_j$ types. Having this information, the power of both types of individual products can be calculated since the C/I_3 specification of the receiver reveals the power of $2f_i - f_j$ type products. Therefore, the parameters values in (1) are not necessarily calculated since the relationship between the powers of both types of products are known.
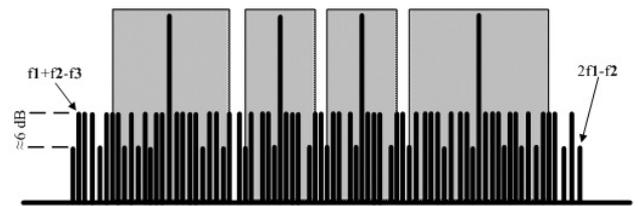
Figure 1. A portion of a satellite spectrum depicting channels, carriers and IMPs.

III. SIMULATION TOOL

A tool has been developed for modeling and simulation of intermodulation distortion caused by the broadband receivers placed at the input section of communication satellite payloads. It can be helpful for both satellite manufacturers and operators to predict intermodulation distortion within the utilized bandwidth. This is because satellite manufacturers need to predict accurately this kind of signal distortion to ensure communication system

performances, whereas the operators need to decide on the configuration of the satellite frequency plan for minimizing the risk of interruptions and providing the best to their customers.

A. Assumptions and Inputs

In order to simulate this kind of distortion, it is assumed that the $2f_i$-f_j and f_i+f_j-f_k types of third-order intermodulation products dominate the in-band intermodulation noise and the higher order terms can be neglected. Unlike many other studies, the carriers are not necessarily being equally spaced. What is assumed in this tool is that there exists one single-level carrier at the center frequency of each satellite channel. Thus, the carrier configuration reflects exactly the frequency plan of the satellite.

Under these assumptions, the tool can be initialized with the following inputs as shown in Fig. 2:

a) *Input power / Carrier (dBm):* Carrier levels at the receiver input. It can be extracted from uplink budget.

b) *C/I₃ (dB):* Signal to third-order intermodulation distortion ratio. C/I_3 is a parameter from receiver datasheets.

c) *Satellite frequency plan (MHz):* Min. and max. frequencies of each channel.

d) If there is a collocated satellite using an adjacent band:
- *Collocated satellite frequency plan (MHz):* Min. and max. frequencies of each channel.
- *Main satellite's input filter rejection ratio for the adjacent band (dB).*

Figure 2. Simulation program user interface

B. Outputs

The algorithm is composed of five main steps. First step is the initialization. After providing the necessary data, program parameters are loaded. Some preliminary calculations are also performed at this stage. In the second and third steps, the frequencies and powers of all $2f_i$-f_j and f_i+f_j-f_k types of products emerging from the given carrier configuration are computed and stored, respectively. Then, all third-order products are sorted and counted with respect to the satellite channel frequencies in order to figure out C/I performances of each channel. Finally, the results are printed on screen in tabular and graphical forms.

The program provides the following results for each satellite channel:

a) Total number and power of $2f_i$-f_j type products
b) Total number and power of f_i+f_j-f_k type products
c) Cumulative third-order intermodulation power
d) C/I, carrier to intermodulation noise ratio

IV. CASE STUDIES

Two cases have been considered and simulated. In the first case a single satellite in the orbit is considered, whereas in the second case, a collocated satellite is assumed to be operating at the same orbital location.

A. First Scenario

We consider a hypothetical communication satellite "Sat_1" that has 15 Ku-band transponders with 36 and 72 MHz bandwidth channels totaling 920 MHz bandwidth. 4 MHz guard bands exist between each channel. The intermodulation products are calculated for the vertical polarization. This calculation should be done for each polarization. In our case, frequency plan is exactly the same for both polarizations. Based on the equipment specifications, the input parameters are given as follows:

a) Input power / Carrier = -60 dBm
b) C/I_3 = 55 dBc
c) Sat_1 Frequency Plan = as seen in Fig. 2.

When the simulation is performed under the assumptions of this scenario, the results are summarized in Table 1. C/I values for each channel are shown graphically in Fig. 3. The number of IMPs of each type generated within each channel is shown graphically in Fig. 4.

B. Second Scenario

In this case, we assume that there exists a collocated satellite "Sat_2" sharing the same orbital position with Sat_1. Sat_2 utilizes a 600 MHz bandwidth in the adjacent band which is comprised of 15 transponders with 36 MHz bandwidths. The input parameters in reference to the equipment specifications are given as follows:

a) Sat_1 Input Filter Rejection @ Sat_2 Freq. = 20 dB
b) Sat_2 Frequency Plan = as seen in Fig. 2.

The simulation results are given in Table 2. The C/I values and the number of IMPs for each channel are also shown in graphical form in Fig. 5 and 6, respectively.

RECEIVER INTERMODULATION SIMULATOR										
Receiver:		Input Power/Carrier : -60 dBm				INTERMODULATION PRODUCTS				
		C/I3	: 55 dBc							
Input Filter:		Adj. Band Rejection : N/A				Analysis Results				
Interfering Carriers		**Analyzed Channels**			(2fi–fj) Type		(fi+fj–fk) Type		∑ Power	C / I
Channels	Carriers	Fmin	Fmax	BW	∑ Number	∑ Power	∑ Number	∑ Power		(dBc)
Sat_1 Ch.1	13558	13522	13594	72	8	-105.97	51	-91.90	-91.74	31.74
Sat_1 Ch.2	13634	13598	13670	72	8	-105.97	67	-90.72	-90.59	30.59
Sat_1 Ch.3	13710	13674	13746	72	8	-105.97	70	-90.53	-90.41	30.41
Sat_1 Ch.4	13768	13750	13786	36	5	-108.01	51	-91.90	-91.80	31.80
Sat_1 Ch.5	13808	13790	13826	36	4	-108.98	27	-94.67	-94.51	34.51
Sat_1 Ch.6	13848	13830	13866	36	5	-108.01	55	-91.58	-91.48	31.48
Sat_1 Ch.7	13906	13870	13942	72	8	-105.97	82	-89.84	-89.74	29.74
Sat_1 Ch.8	13982	13946	14018	72	10	-105.00	89	-89.49	-89.37	29.37
Sat_1 Ch.9	14058	14022	14094	72	8	-105.97	82	-89.84	-89.74	29.74
Sat_1 Ch.10	14116	14098	14134	36	5	-108.01	55	-91.58	-91.48	31.48
Sat_1 Ch.11	14156	14138	14174	36	4	-108.98	27	-94.67	-94.51	34.51
Sat_1 Ch.12	14196	14178	14214	36	5	-108.01	51	-91.90	-91.80	31.80
Sat_1 Ch.13	14254	14218	14290	72	8	-105.97	70	-90.53	-90.41	30.41
Sat_1 Ch.14	14330	14294	14366	72	8	-105.97	67	-90.72	-90.59	30.59
Sat_1 Ch.15	14406	14370	14442	72	8	-105.97	51	-91.90	-91.74	31.74

Table 1. Simulation results for the first scenario. (On the left hand side of the table, inputs and other information derived from the inputs, such as channel bandwidth, are listed. On the right hand side, total number and power of both types of third-order intermodulation products, cumulative intermodulation noise power and *C/I* values are given for each channel. Frequencies are given in MHz and the power values are in dBm.)

RECEIVER INTERMODULATION SIMULATOR										
Receiver:		Input Power/Carrier : -60 dBm				INTERMODULATION PRODUCTS				
		C/I3	: 55 dBc							
Input Filter:		Adj. Band Rejection : 20 dB				Analysis Results				
Interfering Carriers		**Analyzed Channels**			(2fi–fj) Type		(fi+fj–fk) Type		∑ Power	C / I
Channels	Carriers	Fmin	Fmax	BW	∑	∑	∑	∑		(dBc)
Sat_1 Ch.1	13558	13522	13594	72	18	-105.92	258	-91.84	-91.67	31.67
Sat_1 Ch.2	13634	13598	13670	72	16	-105.93	290	-90.68	-90.55	30.55
Sat_1 Ch.3	13710	13674	13746	72	18	-105.92	323	-90.49	-90.37	30.37
Sat_1 Ch.4	13768	13750	13786	36	10	-107.97	187	-91.88	-91.78	31.78
Sat_1 Ch.5	13808	13790	13826	36	9	-108.93	171	-94.62	-94.46	34.46
Sat_1 Ch.6	13848	13830	13866	36	9	-107.98	200	-91.56	-91.46	31.46
Sat_1 Ch.7	13906	13870	13942	72	15	-105.93	379	-89.82	-89.71	29.71
Sat_1 Ch.8	13982	13946	14018	72	17	-104.98	398	-89.46	-89.34	29.34
Sat_1 Ch.9	14058	14022	14094	72	16	-105.95	403	-89.81	-89.70	29.70
Sat_1 Ch.10	14116	14098	14134	36	10	-107.99	222	-91.54	-91.45	31.45
Sat_1 Ch.11	14156	14138	14174	36	9	-108.96	201	-94.59	-94.43	34.43
Sat_1 Ch.12	14196	14178	14214	36	10	-108.00	223	-91.86	-91.76	31.76
Sat_1 Ch.13	14254	14218	14290	72	18	-105.96	417	-90.46	-90.34	30.34
Sat_1 Ch.14	14330	14294	14366	72	20	-105.96	412	-90.64	-90.51	30.51
Sat_1 Ch.15	14406	14370	14442	72	21	-105.97	424	-91.76	-91.60	31.60

Table 2. Simulation results for the second scenario.

C. Discussion

For the first scenario, the maximum intermodulation distortion occurs in the 8th channel (ch.8) with -89.37 dBm, and the minimum in 5th and 11th channels with a value of -94.51 dBm. Due to the symmetric configuration of the frequency plan, the results are also symmetric around ch8. In the second scenario, the maximum distortion takes place in the 8th channel again with -89.34 dBm, whereas the minimum is in the 5th channel with -94.46 dBm. Since there is no symmetry in this case, the existence of the collocated satellite channels affects the distortion scheme. According to these results, we can conclude the followings:

a) Resultant C/I values of all channels are at least 20 dB less than the C/I_3 specification of the receiver itself i.e., the effect of third-order intermodulation distortion is quite significant.

b) The C/I values of the channels vary up to 5 dB. Thus, intermodulation distortion affects the channels differently depending on the carrier configuration.

c) $f_i+f_j-f_k$ type of intermodulation products are much greater than $2f_i-f_j$ type both in number and power. i.e., $f_i+f_j-f_k$ type products are much more dominant in determining the cumulative distortion power.

d) The existence of a collocated satellite system operating in an adjacent band affects the C/I performances. However, this effect is not significant in this case because the adjacent satellite carriers are suppressed by 20 dB by the input filter before arriving at the receiver of the first satellite. Depending upon the input filter rejection, C/I values will be different. The more rejection means much better C/I values.
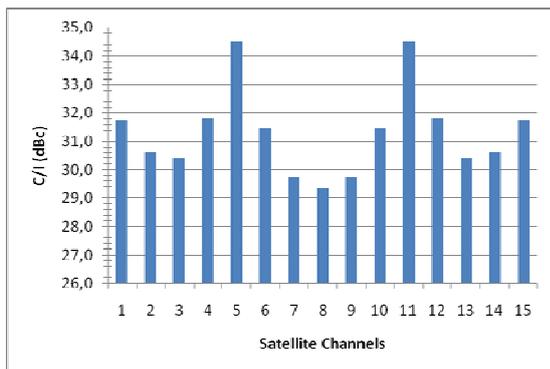
Figure 2. C/I at receiver output for each channel (first scenario)
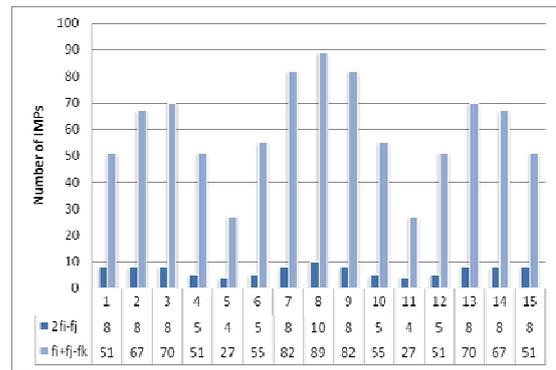
Figure 3. Number of IMPs for each channel (first scenario)
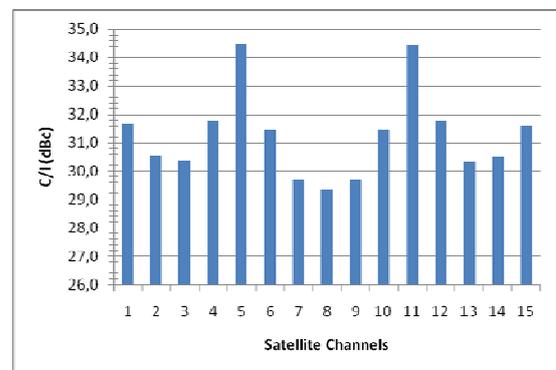
Figure 4. C/I at receiver output for each channel (second scenario)
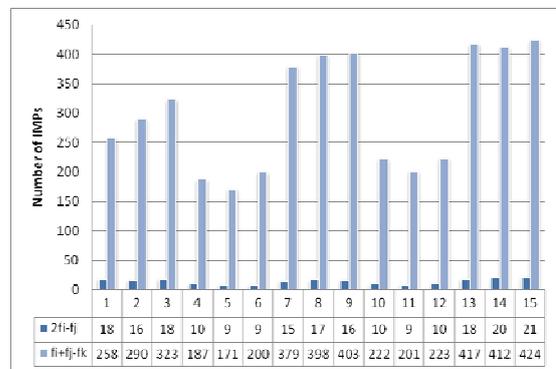
Figure 5. Number of IMPs for each channel (second scenario)

V.   CONCLUSION

The two case studies demonstrate that significant degradation in the C/I performance may occur due to intermodulation phenomena. Besides, the level of degradation varies among the satellite transponders. Therefore, the C/I performances of each channel should carefully be analyzed. In this work, we have presented a practical simulation tool to compute the effect of third order intermodulation distortion caused by the nonlinear characteristics of communication satellite receivers. It presents channel by channel results and helps optimizing

satellite frequency plan. Besides, the effect of collocated satellite system on the linearity performance of the payload input section can also be simulated. In order to achieve a good balance between computational complexity and accuracy, third order intermodulation products are taken into account, namely $2f_i\text{-}f_j$ and $f_i\text{+}f_j\text{-}f_k$ types. The tool is configured for the use of satellite communications experts.

REFERENCES

[1] J. M. P. Fortes and R. Sampaio-Neto, "A fast algorithm for sorting and counting third-order intermodulation products," *IEEE Trans. Commun.*, vol. COM-34, no. 12, pp. 1266-1272, Dec. 1986.

[2] G. Baruffa and G. Reali, "A fast algorithm to find generic odd and even order intermodulation products," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3749-3759, Oct. 2007.

[3] Y. H. Lau and D. J. Skellern, "Calculation of third order intermodulation spectrum," *IEEE Electron. Lett.*, vol. 26, no. 21, pp. 1776–1778, Oct. 1990.

[4] O. A. Diogo and J. P. A. Albuquerque, "An indirect procedure for counting intermodulation products among narrow-band carriers," *IEEE Trans. Commun.*, vol. 40, no. 2, pp. 345-350, Feb. 1992.

[5] S. Nam and E. J. Powers, "On the selective counting of third-order products," *IEEE Trans. Commun.*, vol. 43, no. 8, pp. 2405-2413, Aug. 1995.

[6] B. Baytekin and R. G. Meyer, "Analysis and simulation of spectral regrowth in radio frequency power amplifiers," *IEEE J. Solid-State Circuits*, vol. 40, no. 2, pp. 370–381, Feb. 2005.

[7] J. C. Pedro and N. B. Carvalho, "On the use of multitone techniques for assessing RF components' intermodulation distortion," *IEEE Trans. Microwave Theory Tech.,* vol. 47, no. 12, pp. 2393-2402, Dec. 1999.

[8] F. E. Bond and H. F. Meyer, "Intermodulation effects in limiter amplifier repeaters," *IEEE Trans. Commun. Technol.*, vol. COM-18, no. 2, pp. 127-135, Apr. 1970.

Assessment of Weather Effects on DSN Antenna Tracking at Ka-band

Timothy Pham

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California, United States
e-mail: Timothy.Pham@jpl.nasa.gov

Abstract— **This paper presents the findings of weather statistics at the three sites of the NASA Deep Space Network (DSN) and the effect of weather on antenna tracking performance at Ka-band frequency, based on operational data. Both aspects of weather effects are examined: the distribution statistics of wind patterns and the antenna pointing performance (and hence received signal level) in the presence of high winds. Due to smaller beamwidth inherent in the large aperture of DSN antennas (34-m diameter for Ka-band reception), a small offset in antenna pointing caused by winds could significantly reduce the received signal power and potentially result in a data outage. Some past models indicated a strong wind impact. Winds, even at a moderately low speed of 16 km/hr (10 mph), can distort the antenna's main reflector and cause signal degradation by as much as 2.9 dB at Ka-band frequency. Establishing the actual performance that is based on operational data is very important for mission operation planning, especially for missions that are power-constrained and thus need to conserve the link margin. Finding an optimal link margin setting that is neither wasteful nor insufficient would enable maximum data return for missions. The findings presented in this paper show that there are significant differences in the weather among the three DSN sites; thus, one possible strategy to maximize data return is to allocate more tracking at sites with low winds. Another surprising finding is that the wind effect is not as large as previously thought. Only small signal degradation (0.8 dB) was observed for wind speeds up to 50 km/hr.**

Keywords- Deep Space Network, Ka-band tracking, weather effect, antenna pointing, performance analysis.

I. INTRODUCTION

The communications between NASA mission operations team and their respective spacecraft venturing into deep space is supported by the NASA Deep Space Network (DSN). Over time, the S- and X-band spectrum allocated for deep space communications has become more and more crowded. By going to a higher operating frequency at Ka-band (32 GHz), missions would be less exposed to potential radio frequency interference (RFI). At the same time, they can realize a better performance with higher antenna gain [1].

One potential drawback however is the higher sensitivity to weather. Winds can affect antenna pointing, pushing it away from the maximum gain position. At Ka-band, the antenna beamwidth is about four times smaller than that at X-band. A small antenna pointing error at Ka-band could

significantly reduce the received signal power and potentially result in more data outages. Some past models indicated a strong wind impact on the DSN 34-m antennas. They predicted that even at moderately low winds of 16 km/hr (10 mph), the signal degradation could be as much as 2.9 dB due to antenna main reflector distortion [2]. This posed a great concern to mission operations planning, especially for missions that have limited link margin; thus, prompted the interest to validate such model with actual operational data. Knowing the true operational performance would allow missions to set an appropriate link margin that is neither unnecessarily large nor insufficient. This in turn enables maximum data return.

The following sections describe the way data are processed and present the results of analysis. Section II discusses the data used for the analysis. Section III shows the statistics of wind measurements at the three main DSN tracking sites. Section IV presents the wind effect on received signal, followed by the effect on antenna pointing error in Section V. The final section provides a summary on the findings.

II. DATA PROCESSING

Data collected and used for this analysis are taken from the DSN performance dashboards, as reported in [3]. The dashboards capture performance monitor data for every spacecraft tracking pass conducted by the DSN antennas. Specific monitor data used in this analysis include wind speeds, antenna pointing errors, and received signal to noise ratio (SNR). Data collected from the tracking of Kepler spacecraft, which employs Ka-band for its downlink telemetry, are used to correlate the SNR fluctuation with antenna pointing error and wind speed. Observed wind speed data collected over a period of January - June 2011 are used to compute the exceedance probability distribution function at each of the three DSN sites.

III. WEATHER STATISTICS

Figure 1 presents a plot of the average wind speed observed for each tracking pass from January to June, 2011. From this data set, the exceedance probability distribution of wind speed are derived and presented in Figure 2. One can see distinctively different statistics for Goldstone compared to those at Canberra and Madrid. Goldstone is significantly windier than the other two sites. The median wind speed at Goldstone is 15 km/hr, more than twice the median wind

speed of 7 km/hr at Canberra and Madrid. Five percent of the time, the wind speed at Goldstone is greater than 40 km/hr, but at Canberra and Madrid it exceeds only 15-18 km/hr for the same percentage of time.

Given this highly unbalanced distribution, one possible strategy in optimizing the mission data return would be to schedule most of the tracking passes over Canberra and Madrid complexes, if the view periods at all sites are equal. However, the weather (clouds and rain) at Canberra and Madrid is significantly worse than at Goldstone [4], so that somewhat offsets the benefits of having lower wind speeds.
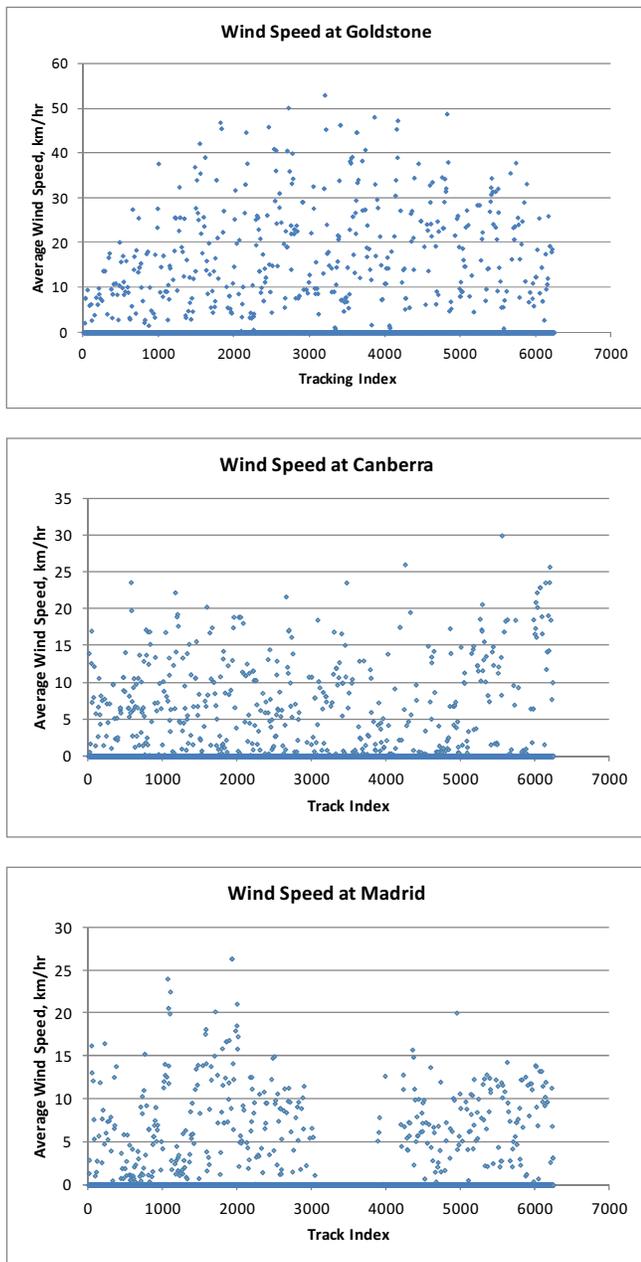
Figure 1. Wind speed distribution at three tracking complexes – Goldstone, Canberra, and Madrid.
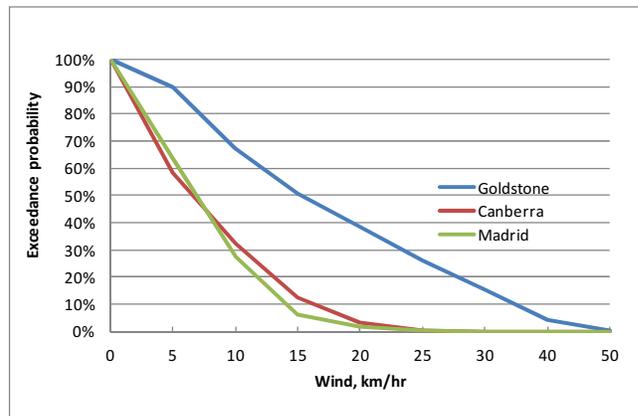
Figure 2. Exceedance probability distribution of wind speed at three tracking complexes.

IV. WIND EFFECT ON SIGNAL RECEPTION

In the current mission set that the DSN supports, Kepler is the only mission that uses Ka-band (32 GHz) for telemetry downlink. The spacecraft, launched in March 2009, spends most of its time searching for exoplanets (planets orbiting other stars in our galaxy) and only downlinks the data to Earth roughly once a month. Thus, Ka-band telemetry tracks are rather limited, making the observed passes with high winds much less common. Figure 3 presents the measurement of received signal to noise ratio (SNR) for both carrier and symbol data, along with wind speeds during the track for about three hours on day-of-year (DOY) 241/2011. On this particular day, wind speed averaged 18 km/hr, with relatively large gusts of up to 50 km/hr. Note also that there were several step-like changes in the carrier SNR profile, but that phenomenon was not reflected in the symbol SNR. Unfortunately, the cause for such jumps is not well understood. Because of these jumps, the carrier SNR data were not used for correlation analysis.

Over the three hours of tracking, the antenna elevation changed from 40 to 70 degrees. This caused a change in the system noise temperature of about 10 kelvins and resulted in a nearly 1-dB increase in SNR from the beginning to the end of the track. The antenna gain also changed slightly over this elevation range. Both changes were removed in the correlation analysis so that their effects would not mask the impact of the wind.

Figure 4 shows the correlation of received symbol SNR relative to the wind speed. Within the inherent fluctuation of 0.8 dB standard deviation in the symbol SNR measurements, a linear fit of the data indicates the degradation caused by winds at 20 km/hr is about 0.3 dB. For winds up to 50 km/hr, the degradation is within 0.8 dB, much less than predicted in past models.

Combining this observation with the exceedance distribution function in Figure 2 where the 50 km/hr wind speed threshold is at the end of the distribution curves for all three sites, it seems that a very high percentage of tracking

passes (~99%) should only experience at most 0.8 dB degradation from the winds, from all causes including main reflector distortion and pointing errors.
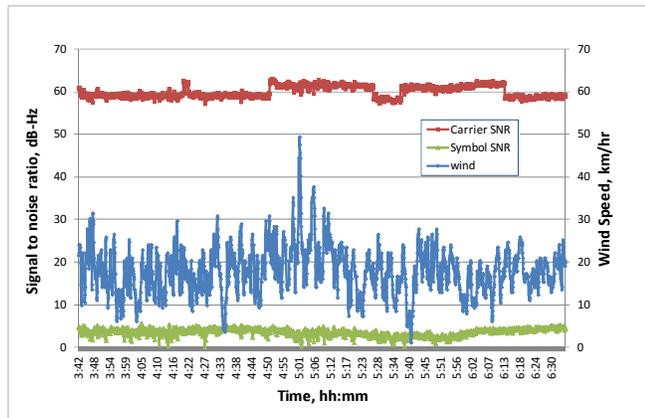
Figure 3.   Signal SNR and wind speed measurements on DOY 241/2011 track with Kepler spacecraft at Ka-band.
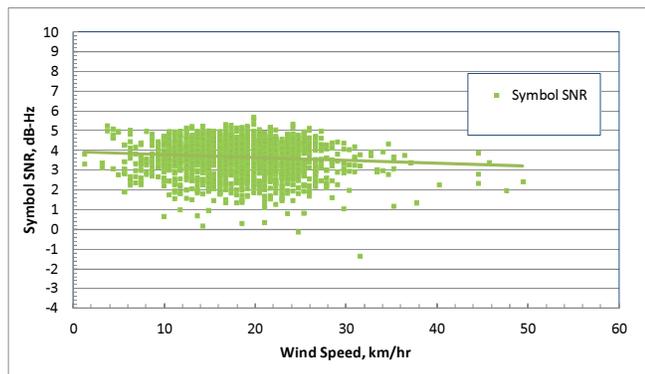
Figure 4.   Correlation of symbol SNR against wind speed.

V.   WIND EFFECT ON ANTENNA POINTING ERROR

To examine the wind effect on antenna pointing, we look at the correlation of wind speed and pointing error. The pointing error is derived from the conical scanning offset. Conical scanning, often referred to as conscan, is a method to keep the antenna on point by continuously circling the antenna around the predicted position. As the antenna traces this conscan circle, the received signal power varies as a function of the radial distance between the instantaneous antenna position and the spacecraft's true position. This variation in signal power allows for determination of pointing offset, and enables correction at the next scan cycle. The conscan offset measurements are used to infer the wind's effect on antenna pointing.

In the earlier data set on DOY241/2011 as discussed in Section IV, conscan was not enabled; thus, we could not analyze the correlation of pointing errors on that particular windy date. However, on the DOY 177/2011 pass, winds up to 20 km/hr were observed and conscan offset data were available for analysis, as shown in Figure 5.
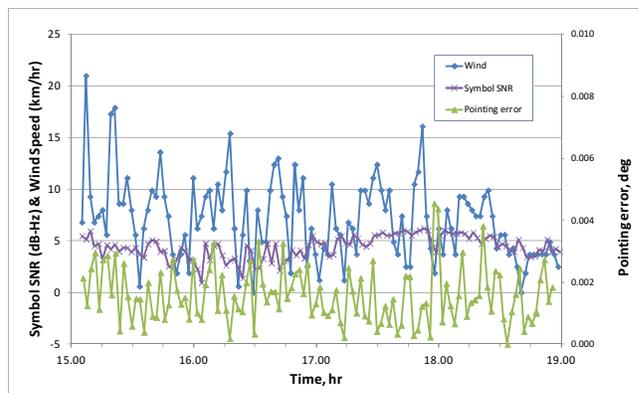
Figure 5.   SNR, pointing error and wind speed on pass DOY 177/2011.

Figure 6 shows a correlation of antenna pointing error and received SNR versus wind speed. The linear fit for the pointing errors is almost a constant, indicating a weak correlation. The change in pointing error is less than 0.0002 deg for wind speed ranges from zero to 20 km/hr. This error would have translated to a very small additional pointing loss of only 0.05 dB around the operating point for the beamwidth of a 34-m antenna at Ka-band.

Note that the linear fit for the symbol SNR seems to be affected by the high noise level of symbol SNR measurements, and does not accurately reflect the expected behavior. It shows an increase in SNR for greater wind speeds and slightly increased pointing errors. As such, the indicated change of 0.6 dB from 0 to 20 km/hr wind does not appear to be real. One conclusion we can draw from the SNR data is that there seems to be no degradation observed.
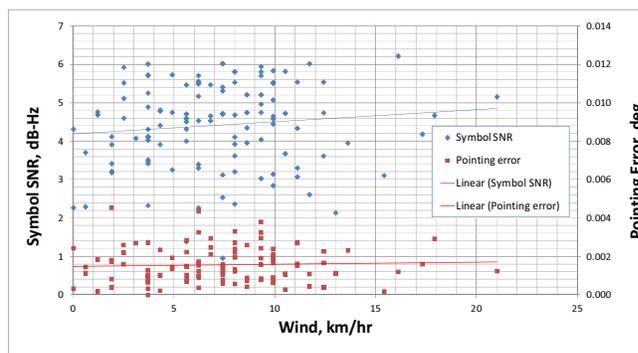
Figure 6.   Correlation of received symbol signal to noise ratio and pointing error to wind speed.

VI.   CONCLUSION

From the analysis of weather data collected at the three main DSN sites, it is established that tracking antennas at Goldstone are subject to significantly windier conditions than those at Canberra and Madrid. Missions, especially those that operate at low link margin, can optimize data return by having more tracking scheduled at Canberra and Madrid, assuming that the view period is equal at all three sites.

Further analysis of the wind effect on symbol SNR and pointing error indicates a minimum effect. Only a small increase in pointing error, 0.0002 deg, was observed with wind speeds up to 20 km/hr. Symbol SNR was estimated to degrade by no more than 0.8 dB for winds up to 50 km/hr, and possibly not at all at wind speeds up to 20 km/hr. These estimates, however, are subject to a comparable SNR measurement uncertainty of 0.8 dB standard deviation.

The finding shows a much more robust Ka-band operation under windy condition than previously believed. This is a testament to the inherent mechanical stability and excellent tracking performance of the large NASA DSN antennas.

ACKNOWLEDGMENT

REFERENCES

[1] Peng, Ted K., "High Rate Telemetry Should Use Ka-band – A Spectrum Management Perspective,", Jet Propulsion Laboratory, Pasadena, CA, April 28, 2001

[2] DSN Telecommunications Link Design Handbook 810-005, 104, Rev. F, 34-m BWG Stations Telecommunications Interface, Jet Propulsion Laboratory, Pasadena, California, June 2010, http://deepspace.jpl.nasa.gov/dsndocs/810-005/104/104F.pdf, accessed: December 2011.

[3] Pham, T. and Liao, J., "DSN Performance Dashboards", SpaceOps 2008, Heidelberg, May 2008.

[4] DSN Telecommunications Link Design Handbook 810-005, Atmospheric and Environmental Effects, Jet Propulsion Laboratory, Pasadena, California, September 2009, http://deepspace.jpl.nasa.gov/dsndocs/810-005/105/105D.pdf, accessed: December 2011.

Nonlinearity Compensation Technique for Common Band Satellite Channel

Takehiro Ishiguro, Takao Hara, and Minoru Okada

Graduate School of Information Science, Nara Institute of Science and Technology,

8916-5, Takayama-cho, Ikoma, Nara, 630-0192, Japan

Email:{takehiro-i,takao-ha,mokada}@is.naist.jp

Abstract—**For effective use of frequency band, carrier superposing (common band) technique has been introduced to satellite communication systems. Meanwhile the satellite's Traveling Wave Tube Amplifier (TWTA) is preferred to be operated near saturation level for power efficiency. However nonlinearity characteristics of TWTA around that level become the cause of interference for carrier superposing systems. Therefore, in this paper, a post compensation technique for TWTA nonlinear distortion is introduced and verified for practical use of carrier superposed Point to Point satellite communication system which adopts interference canceller. Result of the simulation shows it is possible to drastically reduce the BER degradation at all range, specially at nonlinear operating point. This paper also reveals the effect of operating point mismatch that can be caused by rainfall and other environmental changes.**

Index Terms—**satellite, common band, nonlinearity compensation, TWTA**

I. INTRODUCTION

Recently, larger transmission capacity for satellite communication is demanded due to the growing usage like digital broadcasting and internet. From this, common band technique, where two carriers are superposed is becoming a considerable technique as a solution to this problem. In carrier superposing communication system, outbound and inbound signals that are sent on two different frequency bands are superposed and sent in the same frequency band. It is possible to retrieve wanted signals by cancelling out the unwanted signals, signals sent from itself, by subtracting replica signals before demodulating superposed signals. From this, frequency utilization efficiency of two way communication channel composed of outbound and inbound signals can be twice as before at maximum. Thus, many researchers, including authors, have investigated possibilities of carrier superposing technique and is now being in practical use [1]–[6].

However, carrier superposing technique is highly affected by channel nonlinearity. From our previous work it is known that when TWTA (Traveling Wave Tube Amplifier) on satellite is operated near saturation level, signals are degraded very seriously and effective frequency usage achieved by carrier superposing drops down [7]–[9]. To avoid this, authors have proposed a scheme which adds a nonlinear distortion to replica signal to emulate TWTA's nonlinear characteristics for compensation [8], [10], [11]. The methods shown in these references were proposed to give same nonlinearity on the replica signal by digital signal processing at the canceller

as the interference signal is received by the satellite TWTA. Though this method can only be used in point to multi-point VSAT system which has an obvious level difference between outbound and inbound signals, and cannot adopt to more general type of channel.

Therefore, in this paper, a scheme which linearizes the received signal by reversely distort received signal based on AM/AM and AM/PM characteristics of satellite's TWTA is proposed. By using this scheme, this paper also proposes a nonlinearity compensation scheme which is adoptable to general type bidirectional satellite communication systems with any kind of level difference between inbound and outbound signals. For nonlinearity compensation of satellite TWTA, there are other types of method like pre-compensation by adding reverse nonlinear distortion at transmitter. However this could cause spectrum spreading in outbound signal and occupies extra band regions. Moreover, it is quite meanless in the common band network to pre-compensate two individual signals at the transmitter side of each earth station because nonlinear distortion is affected to superposed signal. In the case of nonlinearity compensation at modulator for earth station HPA, the pre-compensation method is valid [12].

From these reasons, we adopt post-compensation of satellite TWTA nonlinearity at receiver. Post-compensation method has not been researched by others. In this research, a case of adopting carrier superpose technique to Point to Point bidirectional channel using QPSK modulation is concerned. A simulation model that is close to the actual system is built to evaluate the performance of interference canceller and demodulator.

Our research reveals the effectiveness of proposed post-compensation scheme. This paper also reveals the effect on the post compensation when rain fall occurs in down path of the link or other environment changes since the proposed scheme may be affected by the change of relation of signal operating level (power) between satellite nonlinearity and its compensator in the receiver. This paper has improved our previous paper [13] by adding detail investigations and simulating in more realistic transmission path model which includes filters between satellite TWTA and receiver.

This paper is organized as follows. Section II explains the detail of nonlinearity compensation algorithm and simulation model of carrier superposed Point to Point satellite communication system. Section III describes the simulation result,

evaluate those results, and shows the effectiveness of proposed scheme. Finally Section IV concludes this paper.

II. CARRIER SUPERPOSED NETWORK WITH NONLINEARITY COMPENSATION
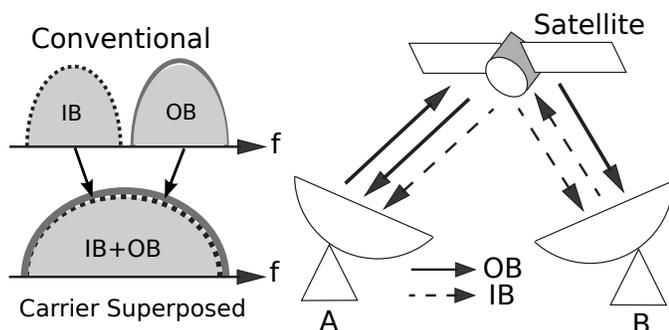
Fig. 1.   P to P equal level carrier superposed satellite network

Fig. 2.   Concept of interference canceller for superposed transmission.

Figure 1 shows P-to-P satellite communication system model used in this research which two opposite carriers have the same level. In this figure, the channel is setup between two earth station A and B. Here, we state the signal carrier from station A to station B an outbound (OB) signal and signal carrier from station B to station A an inbound (IB) signal.

In conventional satellite communication systems, OB and IB signals are transmitted on different frequency band. In contrary, in carrier superposed system, those signals are superposed and transmitted on same frequency band. From this, either the outbound or inbound signal's frequency band will be free and both carrier frequency can be twice as wide as before, which means doubling the frequency band usage efficiency. To retrieve wanted signal from received signal, unwanted signal, in this case self transmitted signal, which is superposed on wanted signal must be cancelled out. This is done by subtracting the replica of self transmitted signal from received signal.

A. Interference Canceller

Interference is cancelled by generating replica of unwanted signal [4], [6]. In Figure 2, two earth stations transmit the signals, and the signals to the left and right are denoted by $S_1(t)$ and $S_2(t)$ respectively.

Let us assume that the cancellation is performed at the earth station to the left in the figure. Since both earth stations transmit the signal at the same frequency, the received signal is the sum of $S_1(t)$ and $S_2(t)$. The received signal at the station is given by:

$$r(t) = \alpha S_1(t - \tau) + \beta S_2(t - \zeta) + z(t) \qquad (1)$$

where α and τ are the round-trip propagation path loss and round-trip delay between the station and the satellite, and β and ζ are the propagation path loss and trip time between two earth stations via satellite. Further, $z(t)$ is an additive Gaussian noise (AWGN) component.

At the canceller, the transmitted signal, $S_1(t)$, is applied to the delay block whose delay time is set to be the same
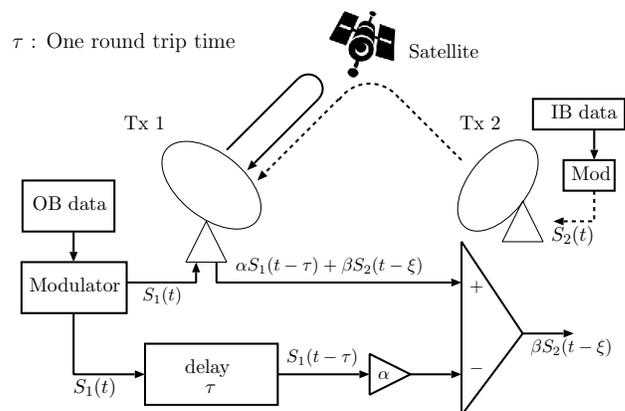
as the round-trip delay, τ. The output of the delay block $S_1(t - \tau)$ is then fed to the variable gain amplifier to adjust amplitude of the replica and the interference signal. If the adjustment is perfect, the gain of the variable gain amplifier is set to be α. The output of the variable gain amplifier is then subtracted from the received signal. The output of the interference canceller is given by

$$u(t) = r(t) - \alpha S_1(t - \tau) = \beta S_2(t - \zeta) + z(t). \qquad (2)$$

From this equation, we can find that the interference signal $S_1(t)$ is successfully cancelled. However, to cancel the interference from the received signal, it is necessary to estimate the round-trip delay of the received signal. Here the accuracy of the delay estimation affects interference suppression performance of the canceller [4], [6].

It is also required to adjust the phase and amplitude of both OB signal and its replica signal. Our previous paper handles the solution to these requirements and has shown that they are performed by applying extended matched filter with phaser locked loop.

B. Nonlinearity Compensation Algorithm

In satellite communications, transponder TWTA is preferable to be operated near the saturated region in order to use its power efficiently. However if used in the saturated region, the signals amplified there suffer from the effects of intermodulation and distortion due to AM/AM and AM/PM properties. Figure 3 shows a typical nonlinear characteristics of the TWTA used in ordinary communication satellite. In the case of carrier superposed system, two or more carriers are amplified together. Figure 4 shows an example of vector diagram of input and output signals through TWTA. Composite signal (a) of outbound (OB) and inbound (IB) is provided to TWTA and converted to TWTA output signal (b).

If the power of OB is large enough compared with IB, the output of TWTA will be almost the same as the TWTA output when only OB signal is input so to cancel the interference, nonlinear distorted OB must be fed to the canceller. On the
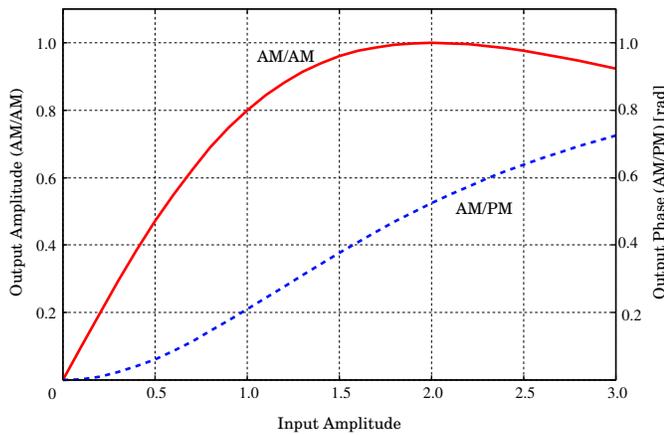
Fig. 3.   Nonlinear characteristics of TWTA
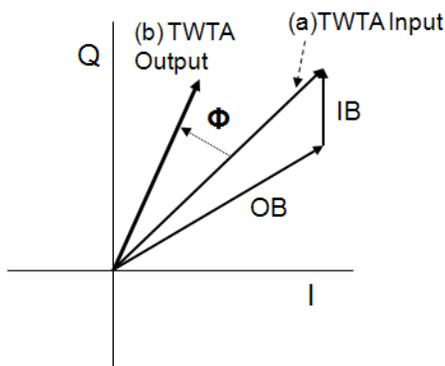
Fig. 4.   Signal vectors of TWTA input and output

other hand, if the power of IB is not negligible compared with OB, TWTA output will be different from nonlinear distorted OB signal and this difference causes interference if simply fed to the canceller. To avoid this interference, we adopt post-compensation scheme to compensate nonlinear effects of TWTA when operated in saturated region.

Amplitude of input vector is distorted according to AM/AM characteristics [14]

$$f(r) = \frac{\alpha_x r}{1 + \beta_x r^2} \qquad (3)$$

and phase of input vector is distorted according to AM/PM characteristics

$$g(r) = \frac{\alpha_\phi r^2}{1 + \beta_\phi r^2} \qquad (4)$$

where r is the amplitude of input vector and α, β are TWTA parameters.

In post-compensation, first, the original TWTA input signal is calculated by the following equation

$$f^{-1}(y) = -\sqrt{\frac{\alpha_x^2}{4\beta_x^2 y^2} - \frac{1}{\beta_x^2}} + \frac{\alpha_x}{2\beta_x y} \qquad (5)$$

where y is the amplitude of received signal vector. This equation can be derived from above equation 3. After received

signal's amplitude is fixed with equation 5, the signal's phase is rotated inversely according to equation 4. These processes of nonlinear compensator are shown in Figure 5. z in the figure represents input signal of nonlinear compensator. Linearization by these methods is actually performed by digital signal processing at receiver side.
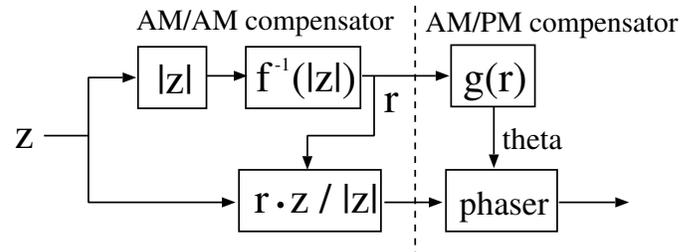
Fig. 5.   Block diagram of nonlinear compensator.

Because of the thermal noise added to the receiver's input signal and filters before nonlinearity compensation, the theory of linearization mentioned above will not be satisfied and thus the signals output from compensator does not exactly match to the satellite's input signal. Therefore, overcompensation and undercompensation might occur and this causes remaining interference after passing the canceller. The simulator explained next is used to explore the effect of thermal noise and filter on nonlinearity compensation.

C. Simulation Model
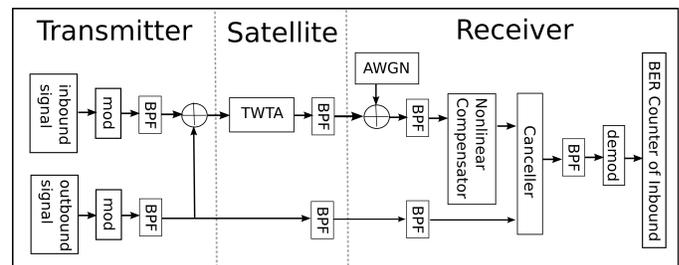
Fig. 6.   Block diagram of the simulator

Figure 6 shows the block diagram of simulator used in this research. The purpose of this simulator is to verify the performance of nonlinearity compensator. Therefore the synchronization mechanism is not implemented in this simulation model. The upper path in the figure represents satellite signal path and lower path represents the replica signal path inside the earth station. To make the transfer function of two paths equal, filters used in satellite path are also inserted in replica path. Nonlinearity compensator is inserted before the interference canceller to fix the distortion that signal receives in TWTA. Root cosine roll-off filter is inserted after transmitter's modulator, TWTA, and before nonlinearity compensator, receiver's demodulator. These filters are for suppression of sideband signal and noises, and also for timing adjustment that is in replica signal path.

III. SIMULATION RESULTS AND CONSIDERATIONS

Simulation is done using software simulator, based on the model described in previous section. The parameters for this simulation is shown in Table I and the parameters for TWTA nonlinearity is shown in Table II.

TABLE I
SIMULATION PARAMETERS

System Parameter	Modulation Method	QPSK
	Modulator Transmit Filter	BW = 1.0, α = 0.35
	Satellite Transmit Filter	BW = 1.2, α = 0.35
	Earth Station Receive Filter	BW = 1.2, α = 0.35
	Demodulator Receive Filter	BW = 1.0, α = 0.35
	Number of Symbol	200,000
	Over Sampling Factor	4
Transmission Channel	DUR (IB/OB ratio)	0dB
	Channel	AWGN

TABLE II
NONLINEARITY PARAMETERS

α_x	1.0
β_x	0.25
α_ϕ	$\pi/12$
β_ϕ	0.25
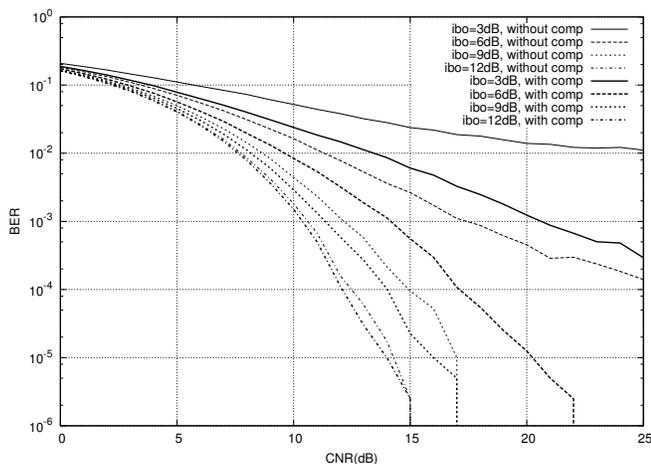
Fig. 7.   Effect of changing input back-off on bit error rate

Figure 7 shows bit error rate (BER) curve when changing backoff. Eight lines on the figure represents the BER at input back-off (IBO) = 3,6,9,12dB each with compensation and without compensation. It is shown on the figure that when IBO is small, which means TWTA operation point is more in nonlinear region, the difference of BER between compensated signal and non-compensated signal becomes larger. Moreover, our compensator recovers BER in more linear region like IBO = 12dB. TWTA's property shown in Figure 3 has slight nonlinearity in amplitude gain and certain amount of phase rotation even if large amount of input back-off is taken. Therefore the compensator is valid on linear region.
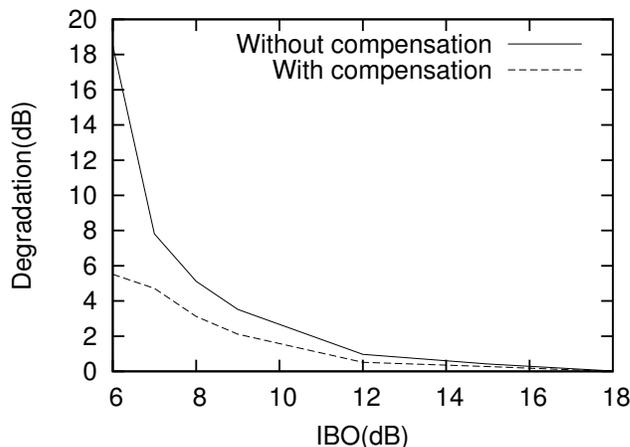
Fig. 8.   Degradation due to TWTA nonlinearity with and without compensation

Figure 8 shows CNR degradation of BER for each IBO with compensation and without compensation. From this figure it is possible to read the amount of BER improvement at same IBO by using nonlinearity compensator. For example, at IBO = 8dB, CNR improvement will be about 3dB which means signal power can be reduced to half to achieve same BER using compensator. Another thing that can be read from this figure is the IBO difference between when nonlinearity is compensated or not at the same CNR. For example, when degradation is 2dB, IBO difference will be about 1.8dB which means amplitude of TWTA operating point can be shifted to 1.2 times higher. This means the TWTA can be used at higher operating level by 2dB.
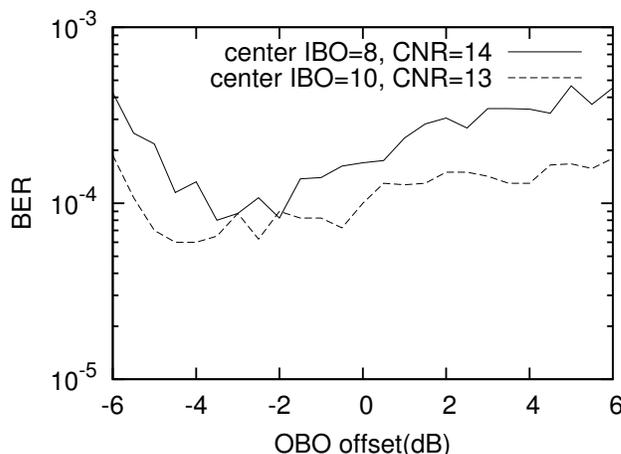
Fig. 9.   BER difference when rain attenuation level changes.

Next the effect of rainfall attenuation is evaluated by intentionally simulating attenuation between satellite and receiver. Figure 9 shows the change of BER when OBO is offseted from ideal value. Two lines in the figure represents the BER change when centers are IBO = 8dB, CNR = 14.0dB and IBO = 10dB, CNR = 13.0dB. This result shows that when

environment changes and channel attenuation varies for a few dB, BER varies in the order of 10^{-4}. This means that CNR degradation will be within about $1 \sim 2\mathrm{dB}$, even in the rain attenuation of $\pm 6\mathrm{dB}$. If an automatic gain control circuit is used at the receiver input, the change of compensator input level due to the rainfall can be compressed within 1dB or so. Then the value of OBO offset in Figure 9 can be reduced to 1dB order.

IV. CONCLUSION

In this paper, a nonlinearity compensation scheme for carrier superposed point to point (P-P) satellite communication system using interference canceller is proposed. To show the effectiveness of proposed scheme, a basic software simulation is performed. The result of simulation shows that it is possible to significantly reduce the degradation caused by nonlinearity of satellite TWTA. This scheme reduces the distortion of received signal vector and from this it is adoptable not only to superposed signals but also to single carrier communication systems and conventional broadcasting signals. Previously, it was very challenging to run this scheme, reverse-nonlinear operation of received signal, on real-time with analog circuit but now the digital signal processing technologies have made it much easier. The result also showed that BER degradation due to operating point mismatch when environment condition changes.

For future work, we are planning to verify the performance of proposed nonlinearity compensator with multi-level modulation such as 32APSK (Amplitude Phase Shift Keying) for more practical use in current satellite systems. In general, multi-level modulation is operated in high CNR condition, which means it is more preferable to nonlinearity compensator. When CNR is large, the difference between compensated signal and original signal becomes smaller. Authors are also planning to verify the effect of nonlinearity compensator from other veiw points like signal constellation, not only the BER performance shown here. Moreover, authors are planning hardware experimentation for complete verification of compensator.

REFERENCES

[1] N. Ishida, "Common-band satellite communication system," *Technical report of IEICE. SAT*, vol. 97, no. 556, pp. 107–114, 1998-02-20. [Online]. Available: http://ci.nii.ac.jp/naid/110003193193/en/

[2] H. Toshinaga, K. Kobayashi, K. Ohata, and H. Kazama, "Interference cancellation for multimedia satellite communication systems employing superposed transmission," *Technical report of IEICE. SAT*, vol. 100, no. 484, pp. 5–10, 2000-12-01.

[3] M. D. Dankberg, M. J. Miller, and M. G. Mulligan, "Self-interference cancellation for two-party relayed communication," U.S. Patent 5 596 439, Jan 21, 1997.

[4] T. Hara, M. Ichikawa, M. Okada, and H. Yamamoto, "Canceller design for carrier super-positioning frequency re-use of vsat satellite communications," *The transactions of the Institute of Electronics, Information and Communication Engineers. B*, vol. 88, no. 7, pp. 1300–1309, 2005-07-01.

[5] M. Osato, H. Kobashi, R. Y. Omaki, T. Hara, and M. Okada, "Development of signal canceller in the carrier super-positioning satellite networks," in *Proceedings of the Third International Conference on Wireless and Mobile Communications*, ser. ICWMC '07, 2007, pp. 35–35.

[6] S. Kuroda, S. Tanaka, S. Naoi, Y. Takeda, R. Miyamoto, T. Hara, and M. Okada, "Development of an interference canceller in satellite communications using a multi-level modulation with superposed transmission." *IEICE Transactions*, vol. E92-B, no. 11, pp. 3354–3364, 2009-11.

[7] T. Uratani, R. Miyamoto, T. Hara, and M. Okada, "Performance of an interference canceller for p-mp satellite networks with nonlinear twta," in *International Symposium on Communications, Control and Signal Processing (ISCCSP'08)*, March 2008.

[8] K. Kubo, R. Miyamoto, T. Hara, and M. Okada, "Nonlinear compensations for interference canceller of super-positioning satellite system," *Technical report of IEICE. SAT*, vol. 109, no. 254, pp. 35–40, 2009-10-22. [Online]. Available: http://ci.nii.ac.jp/naid/110007483048/en/

[9] K. Kubo, H. Matsuda, K. Osawa, T. Shimada, S. Naoi, Y. Takeda, R. Miyamoto, T. Hara, and M. Okada, "Experimental test of frequency reuse interference canceller in the path containing nonlinear twta," *Technical report of IEICE. SAT*, vol. 109, no. 254, pp. 23–28, 2009-10-22.

[10] T. Hara, K. Kubo, and M. Okada, "Adaptive back-off nonlinearity compensation for interference canceller in carrier super-positioning satellite communications," *IEICE Transactions on Communications*, vol. 94, no. 7, pp. 2022–2031, 2011.

[11] T. Hara, H. Matsuda, K. Kubo, and M. Okada, "Performance improvement of interference canceller for carrier super-positioning by the nonlinearity compensation in satellite communications," in *The Sixth International Conference on Wireless and Mobile Communications (ICWMC2010)*, September 2010.

[12] M. Kojima, Y. Suzuki, A. Hashimoto, H. Sujikai, S. Tanaka, T. Kimura, and K. Shogen, "Non-linear compensation technique based on signal point error estimation on satellite transmission system," *Technical report of IEICE. SAT*, vol. 108, no. 440, pp. 37–42, 2009-02-19.

[13] H. Matsuda, T. Ishiguro, T. Hara, and M. Okada, "Nonlinearity compensation for super-positioning satellite system with interference canceller," in *The Seventh Advanced International Conference on Telecommunications (AICT2011)*, March 2011.

[14] G. Santella and F. Mazzenga, "A model for performance evaluation in m-qam-ofdm schemes in prescence of nonlinear distortions," in *Vehicular Technology Conference, 1995 IEEE 45th*, vol. 2, Jul 1995, pp. 830 –834 vol.2.

DVB-S2 Model Based on Reed-Solomon Outer Encoder

Bačić Iva, Malarić Krešimir, Dumić Emil

Department for wireless communication

Faculty of Electrical Engineering and Computing

Zagreb, Croatia

iva.bacic@fer.hr, kresimir.malaric@fer.hr, emil.dumic@fer.hr

Abstract—**This paper analyzes the performance of a new DVB-S2 system. The standard outer Bose-Chaudhuri-Hoquenghem encoder is substituted with a Reed-Solomon (204, 188) encoder in order to analyze and possibly improve the system performances. Simulation for analyzing the communication quality is designed via Matlab and Simulink. The results are compared with results from the existing DVB-S2 model. The new model with the Reed-Solomon outer encoder shows better performances: when noise reaches a level where the existing model gains no quasi-error free performance, the new DVB-S2 model shows a quasi-error-free communication.**

Keywords-DVB-S2; Reed-Solomon; outer encoder; communication system design.

I. INTRODUCTION

DVB-S2 is the second generation specification for satellite broadcasting and part of the European standard, the EN 302 307 [1]. The new combination of a sophisticated coding scheme based on the low-density parity-check, LDPC inner codes concatenated with the Bose-Chaudhuri-Hoquenghem (BCH) outer codes allows a quasi-error-free operation very near the Shannon limit [2], spectral capacity increase by 30-35% over the DVB-S standard for a given symbol rate and carrier to noise ratio is accomplished with modulation and coding schemes of higher order [3]. The DVB-S2 system was designed for high definition television, interactive applications like Internet navigation etc. providing them the best transmission performance, total flexibility and receiver with reasonable complexity [4].

In an earlier work [5], we designed simple communication models designed for analyzing influence of different model parameters on communication quality. These simulations showed that models perform better when the message is coded with the Reed-Solomon code over the BCH code. Therefore, the idea is to substitute the existing BCH code in the DVB-S2 model with the Reed-Solomon code in order to test the communication quality after substitution. Simulation is made using Matlab/Simulink.

This paper is organized as follows. Section II describes the bases of Reed-Solomon encoder and its implementation in the new DVB-S2 model based on the Reed-Solomon outer encoder. In Section III, the simulation models of both, the standard and the new DVB-S2 systems are given. Simulation results are presented in Section IV and Section V concludes the paper.

II. THEORETICAL BACKGROUND

A simplified model of the DVB-S2 transmitter is given in Figure 1. The model is based on two codes: the outer BCH code and the inner LDPC code. The binary message of length k' (generated with the bit generator) is encoded into a BCH codeword of length n'. The BCH code is a multilevel cyclic variable-length digital error-correcting code used to correct multiple random error patterns [4].

The BCH codeword is sent to the LDPC outer encoder and treated as a message of length $k=n'$ to be encoded into a n-bit-long systematic LDPC codeword. The LDPC codes are type of error-correcting codes presented with a large sparse parity-check matrix H of the size $M \times N$. The sparse matrix is defined with the number of ones, $\omega_r << N$ in every row and the number of ones in every columns, $\omega_c << M$ [6], [7]. The parity-check matrix may be written as:

$$H = \begin{bmatrix} H_1 & H_2 \end{bmatrix} \qquad (1)$$

$$H_2 = \begin{bmatrix} 1 & & & & & 0 \\ 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & \cdot & \cdot & \cdot & \\ 0 & & & & 1 & 1 \end{bmatrix} \qquad (2)$$

where H_1 and H_2 are matrixes of size $(N-K) \times K$ and $(N-K) \times (N-K)$. N is the block length and K is the number of message bits. So defined, *(N-K)* are the parity bits. H_1 is a sparse matrix and H_2 is defined with equation (2). Since the low-density parity-check codes are mostly extended irregular repeat-accumulate codes (eIRA) [8], the encoding can be done via the parity-check matrix with linear complexity by noting:

Bit generator → BCH encoder → LDPC encoder → modulation → channel

Figure 1. Simplified DVB-S2 transmitter model.

Q

10

00

$\rho=1$

$\phi=\pi/4$

I

11

01

Figure 2. The Gray coded QPSK mapping.

| Bit generator | → | RS encoder | → | LDPC encoder | → | modulation |

modulation → channel

Figure 3. Simplified block scheme of new transmitter model

204 bytes

| 188 bytes Data | 16 parity bytes |

Figure 4. A Reed-Solomon codeword.

$$Hc^T = 0 \, , \qquad (3)$$

where c is the codeword [9]. This has to be solved for parity bits recursively [10].

According to the DVB-S2 standard [1], the codeword may be of length $n= 64800$ or $n=16200$ referencing the normal and the short frame messages. The ratio k/n at the LDPC encoder defines the code rate of system. The DVB-S2 standard defines eleven different code rates for the normal frame messages: 9/10, 8/9, 5/6, 4/5, 3/4, 2/3, 3/5, 1/2, 2/5, 1/3 and 1/4. Short frames are suitable for all the same code rates, except for rate 9/10 [1].

The coded message is modulated onto a carrier via the QPSK, 8PSK, 16APSK or the 32APSK modulation depending on application area. Gray mapping is used for QPSK (Figure 2) and 8PSK modulation schemes [1].

Reed-Solomon codes showed to be suitable when mentioned burst errors occur. This paper presents a new DVB-S2 system where the BCH encoder is substituted with the Reed-Solomon coder (Figure 3).

With the Reed-Solomon code, the bits are grouped into symbols, instead of being encoded directly in bits. The data words and code words are encoded in these symbols.__

The Reed-Solomon code is specified as RS (n', k') and every Reed-Solomon symbol has s bits. The code takes k' original-message-symbols and encodes it into n' symbols. The number of redundant bits is $n'-k'$ and they are not part of the message itself but are sent through the channel to protect the useful message. The number of possible codewords is $2N$, but only $2K$ contain data words. The rest of the codewords are not contributing to the message and are redundant.

If errors occur, there is a high probability that they will convert the permissible code words into one of the redundant words that the decoder at the receiver is designed to recognize as an error [11]. The number of symbol errors that may be corrected, t is defined:

$$2t = n' - k' \, . \qquad (5)$$
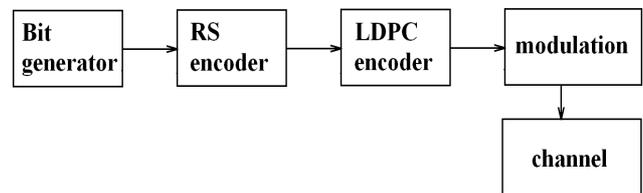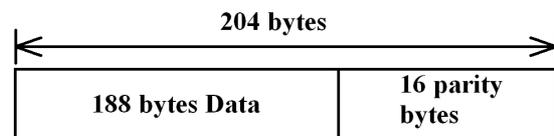
The Reed-Solomon code may detect and correct a corrupted symbol with a single bit errors as it can a symbol with all its bits in error, meaning that the Reed-Solomon code is suitable for correcting burst errors [4].

The DVB code is the Reed-Solomon (204, 188) code using 8-bit symbols [12]. The Reed-Solomon (204, 188) is the shortened code, from the original Reed-Solomon (255, 239) code and it will be applied to each randomized transport packet (of 188 bytes) to generate an error protected packet. The Reed-Solomon code adds the parity bytes after the information bytes of the transport packets (Figure 4). In DVB standards, the Reed-Solomon code can correct up to 8 erroneous bytes per packet.

III. SIMULATION MODEL

Our simulation model is based on implemented DVB-S2 simulation in Matlab; command 'commdvbs2' (Figure 5).

Before simulating the existing model, simulation and system parameters will be chosen. The user has to define the code rate as well as the modulation scheme prior starting simulation. The simulation takes the 188 bytes long random binary sequence generated by using the Bernoulli generator (packet sourced) as its input. The message is sent through an AWGN channel and received at the receiver, demodulated and decoded. The bit error rate, BER is calculated after the low-density parity-check code block and the packet error, PER is calculated after the outer Reed-Solomon code block. Additionally, the simulation provides the constellation diagram for particular communication model.

Our DVB-S2 model is presented in Figure 3 and its Matlab realization in Figure 5. The simulation used only the quadrature phase shift keying modulation scheme (Figure 2). The code rates may be set to any code rate supported by the DVB-S2 standard, as listed in Section II.
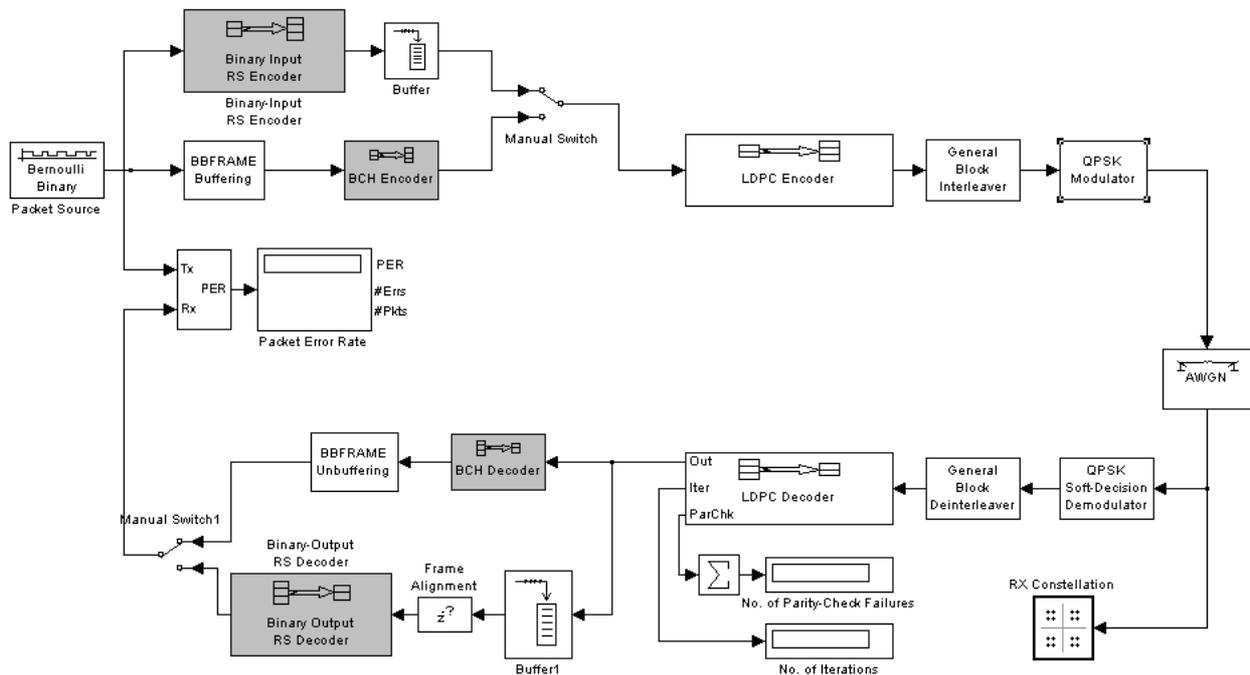
Figure 5.   The DVB-S2 Matlab simulation models with BCH or Reed-Solomon outer encoder.

The message to be sent is generated via the random bit generator. The generated bits are encoded by the Reed-Solomon (204, 188) encoder. The Reed-Solomon encoder adds 16 parity bytes after the information bits (Figure 4) of the transport packets. The transport packet after the Reed-Solomon encoding is 204 bytes long. According to (5), the described Reed-Solomon code may correct up to 8 bytes in one packet. The Reed-Solomon coded message is led to the low-density parity-check code block to be encoded.

The maximum number of iterations is set to 50 [1], meaning that the simulation will stop after 50 iterations if the encoded message is not calculated.

The codeword is defined to be of length $n=64800$ [1].

The encoded message is sent through the Gaussian channel with added white noise meaning that there is only one main path, without any delays. The channel should be quasi-error-free (with BER<10-7) after the outer encoder; after the BCH for the DVB-S2 and the RS outer encoder for the new DVB-S2. In one case, BER falls below 10e-7 and one burst of erroneous bits passed the outer encoder.

At the receiver's side, the received message is demodulated via the QPSK demodulator.

After demodulation, the information is sent to the low-density parity-check decoder. The decoding is based on the concept of belief propagation. It is an iterative process where the information of the received bits is refined iteration by iteration [13]. The maximal number of iteration is set to 50 [1].

The Reed-Solomon decoder performs the outer decoding of message and results in information bits generated at the binary generator (with or without any errors).

For every simulation, the constellation graph is plotted. The bit error rate is calculated as well as the packet error rate. The communication quality may be analyzed by comparing the gained bit error rates and packet error rates for different signal to noise ratios.

IV.   SIMULATION RESULTS

Simulation results are gained with simulation models developed with Matlab and Simulink and described before.

Simulation results are calculated for the implemented DVB-S2 Simulink model and for the presented new DVB-S2 model based on the Reed-Solomon outer coding. The type of chosen Reed-Solomon encoder is the Reed-Solomon (204, 188) encoder meaning that the data words of 188 symbols are mapped into code words with length of 204 symbols (Figure 4). The code is designed to correct up to $t = 8$ symbol errors. The QPSK modulation is chosen. The FEC rates are variable for each simulation. The results are calculated for different noise levels. The first is the noise level where the bit error rate has a quasi-error-free value at the implemented DVB-S2 model. The second is the nearest noise level where bit error rate has no quasi-error-free value; the bit error rate begins to be less than the minimum allowed value. The third noise level is the noise level where the modified DVB-S2 model shows no quasi-error-free value.

Tables 1 – 5 contain representative results from five different simulations.

Table 1 shows that the implemented model of the DVB-S2 system performs quasi-error-free until E_s/N_0 reaches the level of $E_s/N_0 = -2.53$ dB. When $E_s/N_0 = -2.54$ dB the bit error rate is no more quasi-error-free and the system is not performing well. The third row in Table 1 presents the performance of the new DVB-S2 model at this SNR. It is shown that the new DVB-S2 model performs quasi-error-free even at $E_s/N_0 = -2.54$ dB. Therefore, at same SNR the new DVB-S2 model shows better performances. The last row in Table 1 shows that substitution of BCH encoder with Reed-Solomon encoder does not improve the system significantly since the new DVB-S2 model gained a non quasi-error-free performance at noise level of $E_s/N_0 = -2.653$ dB.

TABLE I. E_s/N_0 PEROFORMANCE AT QUASI-ERROR-FREE, 10E-7 AWGN CHANNEL FOR QPSK 1/4

QPSK 1/4	E_s/N_0 [dB]	Number of simulated bits	LDPC BER
DVB-S2	-2.53	1.62 e+5	0
	-2.54	1.62 e+5	6.173e-6
new DVB-S2	-2.54	1.782 e+5	0
	-2.65	1.782 e+5	5.612e-6

TABLE II. E_s/N_0 PEROFORMANCE AT QUASI-ERROR-FREE, 10E-7 AWGN CHANNEL FOR QPSK 1/3

QPSK 1/3	E_s/N_0 [dB]	Number of simulated bits	LDPC BER
DVB-S2	-1.40	2.16 e+5	0
	-1.41	2.16 e+5	4.64e-6
new DVB-S2	-1.41	2.376 e+5	0
	-1.50	2.376 e+5	3.15e-6

TABLE III. E_s/N_0 PEROFORMANCE AT QUASI-ERROR-FREE, 10E-7 AWGN CHANNEL FOR QPSK 1/2

QPSK 1/2	E_s/N_0 [dB]	Number of simulated bits	LDPC BER
DVB-S2	0.85	3.24 e+5	0
	0.84	3.24 e+5	2.16e-5
new DVB-S2	0.84	3.564 e+5	0
	0.80	3.564 e+5	2.806e-6

TABLE IV. E_s/N_0 PEROFORMANCE AT QUASI-ERROR-FREE, 10E-7 AWGN CHANNEL FOR QPSK 2/3

QPSK 2/3	E_s/N_0 [dB]	Number of simulated bits	LDPC BER
DVB-S2	2.98	4.32 e+5	0
	2.95	4.32 e+5	3.33e-4
new DVB-S2	2.95	4.752 e+5	0
	2.93	4.752 e+5	3.577e-5

TABLE V. E_s/N_0 PEROFORMANCE AT QUASI-ERROR-FREE, 10E-7 AWGN CHANNEL FOR QPSK 3/4

QPSK 3/4	E_s/N_0 [dB]	Number of simulated bits	LDPC BER
DVB-S2	3.96	4.86 e+5	0
	3.94	4.86 e+5	7.61e-4
new DVB-S2	3.94	5.346 e+5	0
	3.91	5.346 e+5	1.871e-6

Table 2 contains simulation results for the QPSK with a code rate of $r = 1/3$. The implemented model performs well with E_s/N_0 of $E_s/N_0 \geq -1.40$ dB. When E_s/N_0 becomes lower the system has no quasi-error-free performance any more. It is shown that our model has an allowed bit error rate even with this noise level. Table 2 shows that with $E_s/N_0 = -1.40$ dB the system has a quasi-error-free bit error rate.

Table 3 shows simulation results for the QPSK with the code rate $r = 1/2$.   The level of noise of $E_s/N_0 = 0.85$ dB has a bit error rate of satisfactory value. When the noise level becomes higher in system, the model shows a non-error free performance. Again, the new model has a quasi-error-free bit error rate, showing that the Reed-Solomon code could protect more bits from noise interference.

Table 4, showing the simulation results for the QPSK 2/3 leads to same conclusion, the new DVB-S2 model with an outer Reed-Solomon (204, 188) code performs quasi-error-free at noise level where the implemented model gained an unsatisfactory performance.

Table 5 contains the bit error rates for the QPSK 3/4. Again, the new model performed quasi-error-free in surrounding ($E_s/N_0 = 3.94$) where the implemented DVB-S2 model based on the BCH gained an unsatisfactory results.
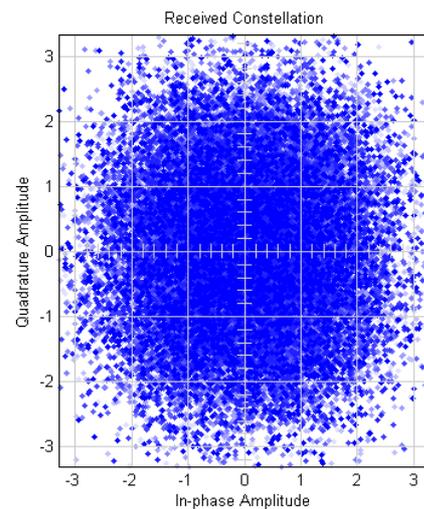
Figure 6.   The constellation graph for the existing DVB-S2 Matlab simulation model for the QPSK 1/4 mapping scheme at $Eb/N0 = -2.54$ dB.
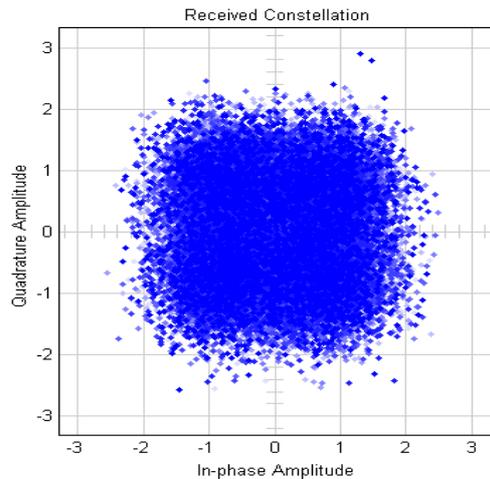
Figure 7. The constellation graph for the new DVB-S2 Matlab simulation model for the QPSK 1/4 mapping scheme at *Eb/N0 = -2.54 dB*.

Simulation results may be analyzed via the constellation graph every simulation resulted with. A representative example of such graphical system performance analyzes is given in Figure 6 and 7. Figure 6 shows the constellation graph for the QPSK 1/4 for the already existing and for the new DVB-S2 model. The results show that the new model performs better since the dots are not so dispersed as they are at the constellation graph for the existing DVB-S2 model.

V. CONCLUSION

This paper presented a new DVB-S2 model based on the Reed-Solomon outer coding since Reed-Solomon code showed to be resistant to burst errors. The type of chosen Reed-Solomon code is the Reed-Solomon (204, 188) code. The Reed-Solomon (204, 188) code maps the data words of 188 symbols into code words with length of 204 symbols. This code can correct up to 8 symbol errors.

Simulation model for the new DVB-S2 is designed with Matlab. Simulation results of the new DVB-S2 model are compared with results gained with an existing DVB-S2 model in Matlab.

The results show that our new DVB-S2 model with the Reed-Solomon (204, 188) code performed better than the existing DVB-S2 model with the BCH outer coding. At E_s/N_0 level, where the existing model gained a non-quasi-error-free performance, the new DVB-S2 model showed better results and a quasi-error-free bit error rate.

The negative aspect of the use of the Reed-Solomon code is a much higher complexity of the Reed-Solomon decoder as compared to the BCH decoder.

REFERENCES

[1] European Telecommunications Standards Institute (ETSI), "Digital video broadcasting (DVB) second generation framing structure for broadband satellite applications, EN 302 307 v1.1.1", www.dvb.org; retrieved: December, 2011.

[2] R. Zhou, A.Picart, R. Pyndiah, and A. Goalic, "Achieving channel capacity with low complexity RS-BTC using QPSK over AWGN channel", Proc. of the 11th Intermational Conference on Telecommunciation (ICT 2004), Aug. 2004, pp. 163-170.

[3] R. Stasiewcz, "DVB-S2 the new international standard", CBC Technology review, Jan. 2007

[4] A.L. Acevedo, "DVB-S vs DVB-S2", Course of certificate in communication system "Satellite", Sept. 2010, UCLA

[5] I. Bacic, K. Malaric, "Analysis for different codes for QPSK modulation in satellite communication", Proc. Modelling, identification and Control (MIC 09), IASTED, Feb. 2009, pp. 32-37, ISBN: 978-0-88986-782-6

[6] S. Muller, M. Scherger, M. Kabutz, M. Alles, F. Kienle, and N. Wehn, "A novel LDPC deecoder for DVB-S2 IP", Proc. Design, Automation and Test in Europe (DATE '05), April 2009, doi:10.1109/DATE.2005.39

[7] T.F. Peforar, F.A.L. Gomes, R.R. Lopes, R. Gallo, J.S.G. Panaro, M.C. Paiva, F.A. Oliveira, and F. Lumertz, "Design, simulation and hardware implementation of a digital television system: LDPC channel coding", Proc. IEEE Symp. Spread Spectrum Techniques and Applications (ISSSTA 06), Aug. 2006, pp. 203-207

[8] M. Yang, W. Ryanm, and Y. Li, "Design of eficiently encodable moderate-length high-rate irregular LDPC codes", IEEE Transaction on Communications, vol. 52, No. 4, pp. 564-571, April 2004.

[9] S.G. Harihora, M.G. Chandra, B.S. Adiga, D.N. Pramod, and P.Balamurilidhar, "A Proporsal for the LDPC decoder architecture for DVB-S2", Proc. Indian National Conference on Communication, pp. 284-288, January 2006.

[10] M. Eroz, F.W. Sun, and L.N. Lee, "DVB-S2 low density parity check codes with near Shannon limit performance", International Journal of Satellite Communications and Networking, vol. 22, June 2004, pp. 269-279,

[11] D. Roddy, Satellite communication, 4th ed., New York: The McGraw-Hill Companies, 2006, pp. 322-324.

[12] European telecommunications Standards Institute (ETSI), "Digital video broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services, EN 300 4217", http://www.etsi.org/; retrieved: December, 2011.

[13] T. Jokela, "Performance Analysis of substituting DVB-S2 LDPC code for DVB-T error control coding system," Proc IEEE Symp. Broadband Multimedia Systems and Broadcasting (BMSB 08), IEEE Press, March/April 2008, pp. 1-5.

Outage Probability Evaluation of Land Mobile Satellite Cooperative Diversity Communication System

Xu Wang, Mingchuan Yang, Qing Guo

Communication Research Center

Harbin Institute of Technology

Harbin, China

wangxu2012hit@yahoo.cn, mcyang@hit.edu.cn , qguo@hitcrc.hit.edu.cn

Abstract—**In this paper, a dual-hop cooperative diversity land mobile satellite system (LMSS) respectively adopted amplify-and-forward (AF), decode-and-forward (DF), and coded cooperation (CC) relaying protocols is proposed, and the outage probability behaviors of the three protocols have been analyzed. We use the signal to noise ratio (SNR) of each path to derive the closed-form expressions of outage probability for each protocol. The accurately approximated expressions in favor of computation simplicity have also been derived applying trapezoidal integration. The Monte Carlo simulations show good match with the proposed analytical results under different shadowing environments, showing that CC is the best strategy for low terminal transmitting SNR. For high transmitting SNR, AF protocol outperforms the high cooperation level CC and DF, while the low cooperation level CC turns out to be better than the other two protocols.**

Keywords-land mobile satellite system; cooperative diversity communication; amplify-and-forward; decode-and-forward; coded cooperation

I. INTRODUCTION

The success of Multiple Input Multiple Output (MIMO) technology in terrestrial systems enormously stimulates the intensive research on its application in satellite communications [1]. In the new generation of satellite systems with strictly restricted available spectrum and power, high transmission speed and spectral efficiency are demanded, therefore, giving rise to more urge to utilize the spatial resources provided by MIMO technology. However, due to the limit of size, power and hardware complexity, it is difficult to employ multiple antennas both on ground terminals and satellite, especially at the mobile terminal. The installation of only one antenna at each user terminal tremendously hampers the development of MIMO technology. Recently, cooperative diversity technique, which is based on user cooperation and relaying cooperation, has emerged as a promising technique and received considerable interests[2][3].The spirit of this technique lies in that single-antenna terminals in a multi-user scenario are allowed to share their antennas and other resources through distributed transmission and processing, so that a virtual multi-antenna transmitter is formed, hence achieving the spatial diversity benefits of the virtual MIMO system.

When one mobile terminal travels in a large area, severe channel conditions, like trees or buildings obstructing signals transmitting directly from terminal to satellite might happen, resulting in complete blockage of line-of-sight (LOS) path. In such situation, cooperative relaying to satellite from other terminals under better propagation environments may offer an effective way to combat channel fading, enable better reliability of communications, as well as reduce the burden of transmitting power for the terminal. This kind of cooperative relaying communication does not rely on the assistance of other terrestrial facilities, making it particularly suitable for the gap-filler-less scenario.

A LMSS is a communication system that provides communications to terrestrial mobile users using satellite. There have been papers focusing on the outage performance under LMSS environments with AF relaying protocol (e.g., [4][5]). But, to the best of our knowledge up to now, there has been few works on the outage probability analysis on DF and CC protocol of LMSS in the literature. However, these two protocols can be incorporated with different channel coding schemes, having the potential to decrease bit error rate, as well as enhance the system feasibility under different channel conditions. So, it is also of great importance to investigate the outage performance of DF and CC in LMSS.

In this paper, we are interested in proposing an uplink mobile terminal assisted cooperation diversity LMSS using AF strategy, DF strategy, and CC strategy, respectively. We derive the closed-form expressions of accurately approximated outage probabilities for the three protocols. The outage performances of these protocols are compared when signal from direct terminal to satellite path is undergoing different degree of shadowing. The analytical results are verified by Monte Carlo simulations.

The remainder of this paper is organized as follows. Section II presents the system models and the channel fading models. In section III we derive the closed-form expressions of approximated outage probabilities for the three relaying protocols. The simulation results are presented in section IV. Finally, conclusion remarks are made in section V.

II. SYSTEM MODELS

A. System Models

We provide the simple frequency nonselective fading system for AF protocol and DF protocol consisting of a source terminal, a cooperative relaying terminal, and the land mobile satellite (LMS) represented by a GEO satellite, as shown in Fig. 1. The source can transmit signal to the relay and the satellite at the same time with different frequency bands using
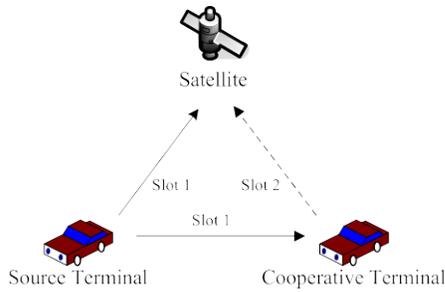
Figure 1. A three node LMS cooperative diversity system for AF and DF.
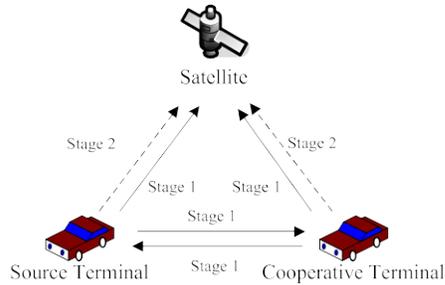
Figure 2. A three node LMS cooperative diversity system for CC.

two sets of antennas. But the relay node cannot transmit and receive signal using the same frequency at the same time, so the half duplex mode is adopted. The time scheduling consists of two time slots. In the first time slot, the source transmits its signal both to satellite and the cooperative terminal with transmitting power p_1. During the second time slot, the scaled version of the signal for AF or regenerated version of the signal for DF at the cooperative terminal is transmitted to the satellite with transmitting power p_2. At the satellite, signals from source and relay nodes are combined through maximum ratio combining (MRC) algorithm.

Based on Fig. 1, the cooperative diversity system for the CC protocol is provided in Fig. 2, where the cooperation process is divided into two stages, during which the source terminal and its cooperative terminal each transmits a total length of N bits containing K information bits. In the first stage, both terminals transmit the information bits to each other and to the satellite in N_1 bits code word with rate $R_1 = K / N_1$. Again, two pairs of antennas are used for terrestrial communication and satellite communication, respectively. If the source node or the relay node cannot correctly decode its partner's code word in the first stage, then it transmits $N_2 = (N - N_1)$ bits code word of its own parity bits to the satellite in the second stage, or its partner's parity bits are transmitted. An important parameter of the protocol is the cooperation level, defined by N_2 / N .

B. Channel Models

Channel fading coefficient from the source to the satellite h_{sd} is assumed to be shadowed Rician fading, and N_0 is the corresponding noise power. The probability density

function (PDF) of the instantaneous SNR $\gamma_{sd} = |h_{sd}|^2 p_1 / N_0$ per symbol is expressed as [6]

$$P_{\gamma_{sd}}(\gamma) = \exp(-\gamma / 2b_0\bar{\gamma}_{sd})[2b_0 m / (2b_0 m + \Omega)]^m$$
$$\times {}_1F_1(m,1,\Omega\gamma / [2b_0\bar{\gamma}_{sd} \times (2b_0 m + \Omega)]) / 2b_0\bar{\gamma}_{sd} \quad (1)$$

where b_0 is the average power of the multipath component, $\bar{\gamma}_{sd} = \sigma_{sd}^2 p_1 / N_0$ is the average SNR between the source and satellite with variance σ_{sd}^2, m is Nakagami parameter, Ω is the average power of the LOS component, and ${}_1F_1(g)$ is the confluent hypergeometric function.

The cumulative distribution function (CDF) of γ_{sd} is

$$P_{\gamma_{sd}}(\gamma) = (2b_0 m)^m / 2b_0\bar{\gamma}_{sd}(2b_0 m + \Omega)^m$$
$$\times \sum_{l=0}^{\infty} (2b_0\bar{\gamma}_{sd})^{(l+1)}\Gamma(l+1,\gamma / 2b_0\bar{\gamma}_{sd}) \quad (2)$$
$$\times (m)_l \Omega^l / (1)_l l![2b_0\bar{\gamma}_{sd} \times (2b_0 m + \Omega)]^l$$

where $\Gamma(g,g)$ is the incomplete gamma function according to (8.350.1) [7], $(g)_l$ is the Pochhammer symbol.

Channel fading coefficients of the source to cooperative relay h_{sr} and the relay to source h_{rs} are assumed to be Rayleigh fading, with the corresponding noise power N_0. The PDFs of the instantaneous SNR $\gamma_{sr} = |h_{sr}|^2 p_1 / N_0$ and $\gamma_{rs} = |h_{rs}|^2 p_2 / N_0$ per symbol are expressed as

$$p_{sr}(\gamma) = \exp(-\gamma / \bar{\gamma}_{sr}) / \bar{\gamma}_{sr} \quad (3)$$

$$p_{rs}(\gamma) = \exp(-\gamma / \bar{\gamma}_{rs}) / \bar{\gamma}_{rs} \quad (4)$$

The CDF of γ_{sr} and γ_{rs} is given as

$$P_{\gamma_{sr}}(\gamma) = 1 - \exp(-\gamma / \bar{\gamma}_{sr}) \quad (5)$$

$$P_{\gamma_{rs}}(\gamma) = 1 - \exp(-\gamma / \bar{\gamma}_{rs}) \quad (6)$$

where $\bar{\gamma}_{sr} = \sigma_{sr}^2 p_1 / N_0$ is the average SNR from the source to the cooperative relay with variance σ_{sr}^2, and $\bar{\gamma}_{rs} = \sigma_{rs}^2 p_2 / N_0$ is the average SNR from the relay to the source with variance σ_{rs}^2 .

We assume that there is always a relay terminal having a clear sight of the satellite, so the channel fading coefficient h_{rd} can be modeled as Rician fading, with the corresponding noise power N_0. And how to find the proper cooperative terminal is left for future work. The PDF of the instantaneous SNR $\gamma_{rd} = |h_{rd}|^2 p_2 / N_0$ per symbol is expressed as

$$p_{\gamma_{rd}}(\gamma) = [(1 + K_{rd}) / \bar{\gamma}_{rd}]\exp[-K_{rd} - (1 + K_{rd})\gamma / \bar{\gamma}_{rd}]$$
$$\times I_0(2\sqrt{K_{rd}(1 + K_{rd})\gamma / \bar{\gamma}_{rd}}) \quad (7)$$

where $\overline{\gamma}_{rd} = \sigma_{rd}^2 p_2 / N_0$ is the average SNR between the relay and the satellite with variance σ_{rd}^2, K_{rd} is the channel Rician K-factor, and $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind.

The CDF of γ_{rd} is given by

$$P_{\gamma_{rd}}(\gamma) = 1 - Q(\sqrt{2K_{rd}}, \sqrt{2(1+K_{rd})\gamma / \overline{\gamma}_{rd}}) \qquad (8)$$

where $Q(\cdot, \cdot)$ is the first-order Marcum Q function [8].

III. OUTAGE PROBABILITY ANALYSIS

When spectral efficiency is set to R, the whole communication system is in outage state when the maximum average mutual information $I_D < R$. From the perspective of information theory, I_D depends on the instantaneous SNR γ_d of the combined signal at the satellite. The outage probability of the source node is

$$P_{out} = \Pr\{\gamma_d < \gamma_{th}\} \qquad (5)$$

where γ_{th} is the threshold decided by R.

For AF relaying protocol, the instantaneous SNR of the combined signal at the satellite is given as [2]

$$\gamma_d = \gamma_{d\_AF} = \gamma_{sd} + \gamma_{sr}\gamma_{rd} / (\gamma_{sr} + \gamma_{rd} + 1) \qquad (6)$$

The outage probability of γ_{d_AF} is derived as

$$
\begin{aligned}
P_{out\_AF} &= \Pr\{\gamma_{d\_AF} < \gamma_{th}\} = \Pr\{\gamma_{d\_AF} < \gamma_{th\_AF}\} \\
&= \int_0^{\gamma_{th\_AF}} p_{\gamma_{AF}}(\gamma_1) \int_0^{\gamma_{th\_AF}-\gamma_1} p_{\gamma_{sd}}(\gamma_2) d\gamma_2 d\gamma_1 \quad (7) \\
&= \int_0^{\gamma_{th\_AF}} p_{\gamma_{AF}}(\gamma_1) P_{\gamma_{sd}}(\gamma_{th\_AF} - \gamma_1) d\gamma_1
\end{aligned}
$$

where $p_{\gamma_{AF}}(\cdot)$ denotes the PDF of the SNR $\gamma_{AF} = \gamma_{sr}\gamma_{rd} / (\gamma_{sr} + \gamma_{rd} + 1)$. Using the method in [9]

$$
\begin{aligned}
p_{\gamma_{AF}}(\gamma) &= (1+K_{rd})/(\overline{\gamma}_{sr}\overline{\gamma}_{rd}) \exp[-K_{rd} - (1+K_{rd})\gamma / \overline{\gamma}_{rd} - \gamma / \overline{\gamma}_{sr}] \\
&\times \{\sum_{t=0}^{\infty} [K_{rd}(K_{rd}+1)/\overline{\gamma}_{rd}]^t \\
&\times \sum_{k=0}^{t+1} 2\gamma^t \binom{t+1}{k}(\gamma+1)^k [(\beta/\eta)^{(1-k)/2} K_{(1-k)}(2\sqrt{\eta\beta}) \\
&+ \gamma(\beta/\eta)^{-k/2} K_{-k}(2\sqrt{\eta\beta})]/(t!)^2\}
\end{aligned}
\qquad (8)
$$

where $\beta = \gamma(\gamma+1)(K_{rd}+1)/\overline{\gamma}_{rd}$, $\eta = 1/\overline{\gamma}_{sr}$, and $K_v(\cdot)$ denotes the vth order modified Bessel function of the second kind.

For DF relaying protocol, the instantaneous SNR of the combined signal at the satellite is given as [2]

$$\gamma_d = \gamma_{d\_DF} = \min\{\gamma_{sr}, \gamma_{sd} + \gamma_{rd}\} \qquad (9)$$

where $\min(\cdot)$ returns the minimum value.

So the outage probability of γ_{d_DF} is derived as

$$
\begin{aligned}
P_{out\_DF} &= \Pr\{\gamma_{d\_DF} < \gamma_{th}\} = \Pr\{\gamma_{d\_DF} < \gamma_{th\_DF}\} \\
&= P_{\gamma_{sr}}(\gamma_{th\_DF}) + [1 - P_{\gamma_{sr}}(\gamma_{th\_DF})] \qquad (10) \\
&\times \int_0^{\gamma_{th\_DF}} p_{\gamma_{sd}}(\gamma_1) \int_0^{\gamma_{th\_DF}-\gamma_1} p_{\gamma_{rd}}(\gamma_2) d\gamma_2 d\gamma_1 \\
&= P_{\gamma_{sr}}(\gamma_{th\_DF}) + [1 - P_{\gamma_{sr}}(\gamma_{th\_DF})] \\
&\times \int_0^{\gamma_{th\_DF}} p_{\gamma_{sd}}(\gamma_1) P_{\gamma_{rs}}(\gamma_{th\_DF} - \gamma_1) d\gamma_1
\end{aligned}
$$

For CC relaying protocol, the instantaneous SNR of the combined signal at the satellite is much more complicated. Here we simply give the source node outage probability expression based on [10] as below

$$
\begin{aligned}
P_{out\_CC} &= [1 - P_{\gamma_{rs}}(2^{R/\alpha} - 1)][1 - P_{\gamma_{sr}}(2^{R/\alpha} - 1)] \\
&\times \int_0^{2^{R/\alpha}-1} p_{\gamma_{sd}}(\gamma_1) P_{\gamma_{rd}}(u_1) d\gamma_1 \\
&+ P_{\gamma_{rs}}(2^{R/\alpha} - 1) P_{\gamma_{sr}}(2^{R/\alpha} - 1) P_{\gamma_{sd}}(2^R - 1) \\
&+ [1 - P_{\gamma_{rs}}(2^{R/\alpha} - 1)] P_{\gamma_{sr}}(2^{R/\alpha} - 1) \\
&\times \int_0^{2^R-1} p_{\gamma_{sd}}(\gamma_1) P_{\gamma_{rd}}(u_2) d\gamma_1 \\
&+ P_{\gamma_{rs}}(2^{R/\alpha} - 1)[1 - P_{\gamma_{sr}}(2^{R/\alpha} - 1)] P_{\gamma_{sd}}(2^{R/\alpha} - 1)
\end{aligned}
\qquad (11)
$$

where $(1-\alpha)$ is the cooperation level of the cooperative diversity system, and also $u_1 = 2^{R/(1-\alpha)} / (1+\gamma_1)^{\alpha/(1-\alpha)} - 1$, while $u_2 = 2^{R/(1-\alpha)} / (1+\gamma_1)^{\alpha/(1-\alpha)} - \gamma_1 - 1$. The channel condition of the three paths for different α is the same as in AF and DF.

The integral operations in (7), (10), and (11) are difficult to compute. In this paper, through the trapezoidal integration with reasonably unit spacing, the accurately approximated integrals can be obtained.

Set $W(\gamma_1) = p_{\gamma_{AF}}(\gamma_1) P_{\gamma_{sd}}(\gamma_{th\_AF} - \gamma_1)$, and the closed-form outage probability expression for AF is written as

$$
\begin{aligned}
P_{out\_AF} &\approx \sum_{n=1}^{q_{AF}-1} W(n\gamma_{th\_AF} / q_{AF})\gamma_{th\_AF} / q_{AF} \\
&+ [W(0) + W(\gamma_{th\_AF})]\gamma_{th\_AF} / 2q_{AF}
\end{aligned}
\qquad (12)
$$

where q_{AF} is the number of the spacings for $W(\gamma_1)$.

Set $X(\gamma_1) = p_{\gamma_{sd}}(\gamma_1) P_{\gamma_{rd}}(\gamma_{th\_DF} - \gamma_1)$, and the closed-form expression of outage probability for DF protocol is written as

$$
\begin{aligned}
P_{out\_DF} &\approx P_{\gamma_{sr}}(\gamma_{th\_DF}) + [1 - P_{\gamma_{sr}}(\gamma_{th\_DF})] \\
&\times \{\sum_{n=1}^{q_{DF}-1} X(n\gamma_{th\_DF} / q)\gamma_{th\_DF} / q_{DF} \\
&+ [X(0) + X(\gamma_{th\_DF})]\gamma_{th\_DF} / 2q_{DF}\}
\end{aligned}
\qquad (13)
$$

where q_{DF} is the number of the spacings for $X(\gamma_1)$.

Set $Y(\gamma_1) = p_{\gamma_{sd}}(\gamma_1)P_{\gamma_{rd}}(2^{R/(1-\alpha)}/(1+\gamma_1)^{\alpha/(1-\alpha)}-1)$,

and $Z(\gamma_1) = p_{\gamma_{sd}}(\gamma_1)P_{\gamma_{rd}}(2^{R/(1-\alpha)}/(1+\gamma_1)^{\alpha/(1-\alpha)}-\gamma_1-1)$,

so the closed-form expression of outage probability for CC protocol is written as

$$P_{out\_CC} \approx [1-P_{\gamma_{rs}}(2^{R/\alpha}-1)][1-P_{\gamma_{sr}}(2^{R/\alpha}-1)]$$

$$\times\{\sum_{n=1}^{q_{CC1}-1} Y(n(2^{R/\alpha}-1)/q_{CC1})(2^{R/\alpha}-1)/q_{CC1}$$

$$+[Y(0)+Y([2^{R/\alpha}-1])][2^{R/\alpha}-1]/2q_{CC1}\}$$

$$+P_{\gamma_{rs}}(2^{R/\alpha}-1)[1-P_{\gamma_{sr}}(2^{R/\alpha}-1)]P_{\gamma_{sd}}(2^R-1)$$

$$+[1-P_{\gamma_{rs}}(2^{R/\alpha}-1)]P_{\gamma_{sr}}(2^{R/\alpha}-1) \qquad (14)$$

$$\times\{\sum_{n=1}^{q_{CC2}-1} Z(n(2^R-1)/q_{CC2})[2^R-1]/q_{CC2}$$

$$+[Z(0)+Z([2^R-1])](2^R-1)/2q_{CC2}\}$$

$$+P_{\gamma_{sr}}(2^{R/\alpha}-1)[1-P_{\gamma_{rs}}(2^{R/\alpha}-1)]P_{\gamma_{sd}}(2^{R/\alpha}-1)$$

where q_{CC1} is the number of the spacings for $Y(\gamma_1)$ and q_{CC2} is the number of the spacings for $Z(\gamma_1)$.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, the Monte Carlo simulation results and the theoretical analysis of the three relaying protocols above are presented with each path experiences nonidentical fading. The relaying protocols are compared with the no cooperative direct transmission as well. Set the transmitting SNR at the source and the cooperative terminal $p_1/N_0 = p_2/N_0$, and $\sigma_{sd}^2 = \sigma_{sr}^2 = \sigma_{rs}^2 = \sigma_{rd}^2 = 1$. In a GEO LMSS, K_{rd} value ranges from 7 to 15 dB [11], so it is assumed that $K_{rd}=11.14dB$. We set $R = 1b/s/Hz$, and the threshold $\gamma_{th} = \gamma_{th\_AF} = \gamma_{th\_DF} = 2^{2R}-1$. For CC protocol, set the cooperation level to 0.1, 0.3, 0.5 and 0.7, respectively. Set $q_{AF} = q_{DF} = q_{CC1} = q_{CC2} = 1000$ with the approximation accuracy of the outage probability. The parameters for different channel condition of the shadowed Rician model given in [6] are shown in Table Ⅰ.

Fig. 3, Fig. 4 and Fig. 5 compare AF, DF, CC and the no cooperation case in terms of the system outage probability

TABLE I. PARAMETERS FOR THE SHADOWED RICIAN UNDER DIFFERENT CHANNEL CONDITIONS

Channel Condition	b_0	m	Ω
Infrequent light shadowing	0.158	19.4	1.29
Average shadowing	0.126	10.1	0.835
Frequent heavy shadowing	0.063	0.739	0.000897
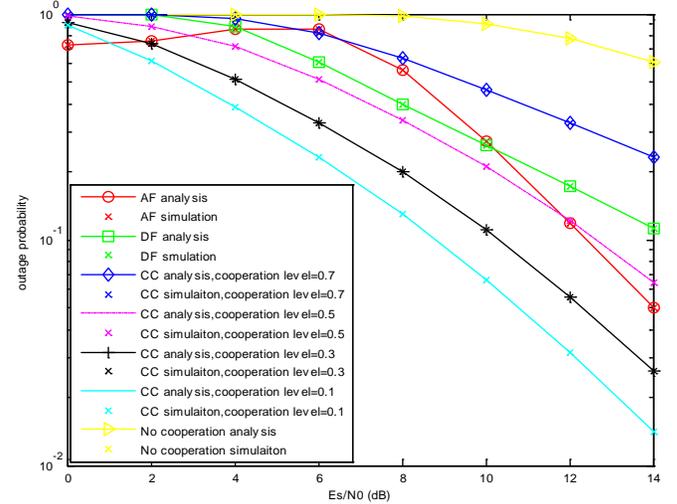
Figure 3. Outage probability for AF, DF and CC when shadowed Rician channel is in frequent heavy shadowing.

under different shadowed Rician channel conditions. The outage probabilities of the system are plotted versus the E_s/N_0 ($E_s/N_0 = p_1/N_0 = p_2/N_0$). It is evident from the figures that, the theoretical results excellently match with the simulation results, verifying the accuracy of the analysis.

Fig. 3 shows the outage probability comparisons between AF, DF, different cooperation level CC, and no cooperation transmission with the source to satellite path undergoing frequent heavy shadowing. In the figure, all three relaying protocols show diversity gain compared to no cooperation scenario. For CC protocol, as α gets lower, the slope of curves tends to get deeper, which means stronger channel coding algorithm for N_1 ensures better outage performance (also shown in Fig. 4 and Fig.5). It can be observed from the figure that when $E_s/N_0 > 6dB$, AF and DF outperform CC when α is high, $\alpha = 0.7$ for example, while the latter protocol is superior to the former two when its cooperation level is as low as 0.1. AF has higher outage probabilities than DF as E_s/N_0 is between $4dB$ and $10dB$. All protocols have high outage probabilities when E_s/N_0 is low, reaching more than 0.70 at $E_s/N_0 = 0dB$. The outage probability between $E_s/N_0 = 0dB$ and $E_s/N_0 = 6dB$ increases with transmitting SNR. This is due to the nonmonotone decreasing property of $p_{\gamma_{AF}}(\gamma)$ with respect to increasing SNR and the channel parameters of the heavy shadowing channel condition.

Fig. 4 shows the outage probability comparisons between the three protocols and no cooperation scenario when the source to satellite path is in average shadowing. Fig. 5 shows the comparisons in the infrequent light shadowing case. Combined with Fig. 3, the three figures illustrate that, the better shadowed Rician channel condition is, the lower the outage probabilities get, with the exception of DF. As a matter of fact, the outage probabilities for the DF protocol are almost the same at the same E_s/N_0 value under three different channel conditions. It can be also seen that, AF and CC outperform the no cooperation scenario except for CC
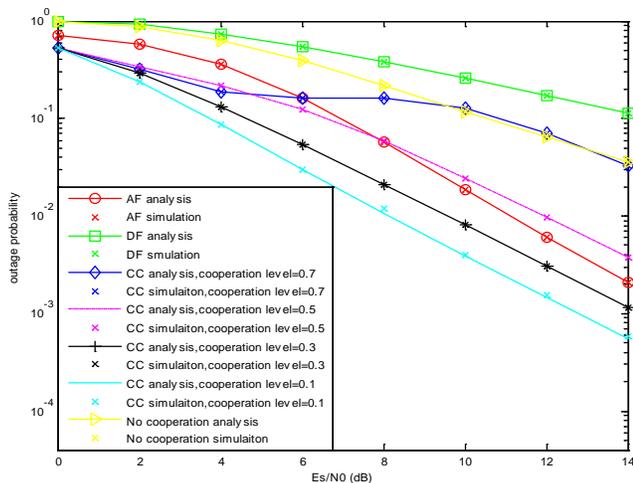
Figure 4.   Outage probability for AF, DF and CC when shadowed Rician channel is in average shadowing.

Figure 5.   Outage probability for AF, DF and CC when shadowed Rician channel is in infrequent light shadowing.

with $\alpha = 0.7$, and DF protocol has the worst outage probability behavior in Fig. 4 and Fig. 5, even worse than the no cooperation scenario.

When comparisons are made between the AF and CC at $E_s / N_0 > 6dB$, similar illustration as in Fig. 3 is shown. But at low E_s / N_0 values, unlike Fig. 3, CC protocol with different cooperation levels all have the obvious advantage over the other two protocols, with the probabilities as low as 0.5 in Fig. 4 and 0.3 in Fig. 5 at $E_s / N_0 = 0dB$, while the probabilities of the others are still higher than 0.7 in both figures.

For CC protocol in Fig. 4 and Fig. 5, it may be noticed that curves of all α values have almost the same trend as in Fig. 3 except for the curve of $\alpha = 0.7$. This result can be explained as follows. It is found that at higher E_s / N_0, the outage probability in (14) is mainly determined by the second term and the fourth term of the addition operation, the former of which is monotonic decreasing while the latter one is convex, thanks to the monotonic increasing outage probability in Rayleigh channel at high cooperation level $\alpha = 0.7$. Therefore, makes it possible for the overall outage possibility to be nondecreasing, or even increasing within certain range of E_s / N_0.

I.   CONCLUSIONS

In this paper, the outage performance of a dual-hop LMS cooperative diversity system is proposed. The closed-form expressions of approximated outage probability for three important and commonly used relaying protocols: amplify-and-forward, decode-and-forward, and coded cooperation, are derived. The simulation results verify the analytical results. The results show that for different shadowing conditions on source to terminal path, at high transmitting SNR, AF protocol is the preferable option to CC protocol with high cooperation level, but with low cooperation level, or at low transmitting SNR regardless of the cooperation level, CC protocol turns out to be the better option for the LMS cooperative diversity system. Both CC protocol and AF
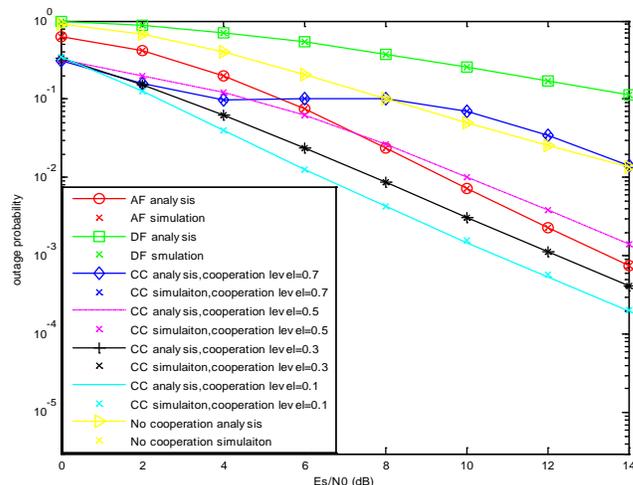
protocol show diversity gain over direct transmission. DF protocol is the last choice of the three protocols.

REFERENCES

[1]   P. Arapoglou, K. Liolis, M. Bertinelli, A. Panagopoulos, Panayotis Cottis and Riccardo De Gaudenzi, "MIMO over satellite: a review," IEEE on Communication Surveys and Tutorials, vol. 13, no. 1, pp. 27-51, 2011.

[2]   J. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior, " IEEE Trans. Inf. Theory, vol.50, no.12, pp.3062–3080, December 2004.

[3]   Gao Zhiwei, Wang Ping and Liu Fuqiang, "Multiuser diversity performance evaluation of 4G cooperative relay cellular network, " Proc. the 15th APCC, pp. 322–325, October 2009.

[4]   Ariflqbal and Kazi Mohiuddin Ahmed," Outage probability analysis of multi-hop cooperative satellite-terrestrial network," 2011 Conf. the 8th ECTI., pp. 256 – 259, May 2011.

[5]   W Limpakom, Y. D. Yao, and H. Man, "Outage probability analysis of wireless relay and cooperative networks in Rician fading channels with different K-factors," IEEE 69th Vehicular Technology Conference (VTC), pp. 1-5, 2009.

[6]   A. Abdi, W.C. Lau, M.S. Alouini, and M. Kaveh, "A simple model for land mobile satellite channels: first-and second-order statistics," IEEE Trans. Wireless Commun., vol.2, no. 3, pp. 519-528, May 2003.

[7]   I. S. Gradshteyn and I. M. Ryzhik,"Table of Integrals, Series, and Products,"7th ed. London, Academic Press, Inc, 2007.

[8]   M. K. Simon and M. S. Alouini, "Digital Communication over Fading Channels," 2 ed. New York: Wiley, 2005.

[9]   Lie-Liang Yang and Hsiao-Hwa Chen, "Error probability of digital communications using relay diversity over Nakagami-m fading channels," IEEE Trans. Wireless Commun., vol.7, no. 5, pp. 1806-1811, May 2008.

[10] Todd Edward Hunter,"Coded cooperation: a new framework for user cooperation in wireless networks," PhD thesis, University of Texas, USA, May 2004.

[11] J. Xu, J. H. Jong and C. Ravishankar, "Channel modeling for a land mobile satellite system," IEEE Conf. Globecom Telecommunication, pp. 4596-4600, 2007.

In Orbit Antenna Pattern Tests and Service Coverage Measurements of COMS

Jin Ho Jo, Moon Hee You, Seong Pal Lee, Jae Hoon Kim

Satellite & Wireless Convergence Research Department,
Electronics and Telecommunications Research Institute (ETRI),
Daejeon City, Rep. of Korea
E-mail: jhjo@etri.re.kr, moon@etri.re.kr, spallee@etri.re.kr, jhkim@etri.re.kr

Abstract—**Communication, Ocean and Meteorological Satellite (COMS) is the multi-purposed Korean geostationary satellite funded by Korean government ministry, and is to supply communication services, ocean and weather observation for 7 years. COMS was launched by Ariane-5 on 26th June 2010 and arrived successfully to operational geo-stationary orbit 128.2E. It features a multi-beam Ka-band Communications package, a set of equipments to study Ocean Ecosystem to aid the fishing industry and a Metrological weather observation system. The multi beam Ka-band antenna in orbit test campaign was conducted by Electronics and Telecommunications Research Institute team. A brief outline on functional and operational capability including software and hardware used for in orbit test validation of Ka-Band antenna system is addressed. After successful in orbit test, service coverage measurement for COMS Ka-band antenna was performed. Methodology and results for service coverage measurement are addressed also. The antenna in orbit test results show COMS Ka-band antennas performance was not degraded by satellite launch and service coverage are formed as we designed over the south Korea peninsula.**

Keywords-Communication, Ocean and Meteorological Satellite; Ka-band payload; In orbit test; Service coverage measurement.

I. INTRODUCTION

Communication, Ocean and Meteorological Satellite (COMS) is a multi-mission satellite and one of the key features of this satellite is the Electronics and Telecommunications Research Institute (ETRI) developed multi-beam Ka-band antenna system [1] that consists of two spot beams; South Korea and North Korea as shown in Figure 1. The antenna system consists of two reflectors installed on the east and west panel of the spacecraft. COMS is first of its kind that incorporates both earth observation and communications payloads from single geostationary satellite [2][3][4]. This paper describes field measurement of COMS antennas pattern and coverage of this unique satellite. Correlation between ground and field measurement shape for this multi beam Ka-band antenna pattern and coverage is presented.

In this paper, COMS Ka-band antenna configuration was reviewed at Section 2. In orbit test methodology for antenna pattern measurement was reviewed at Section 3 and in orbit test results was reviewed at Section 4. Section 5 and Section 6 describe filed measurement of COMS Ka-band payload service coverage and conclusion, respectively.
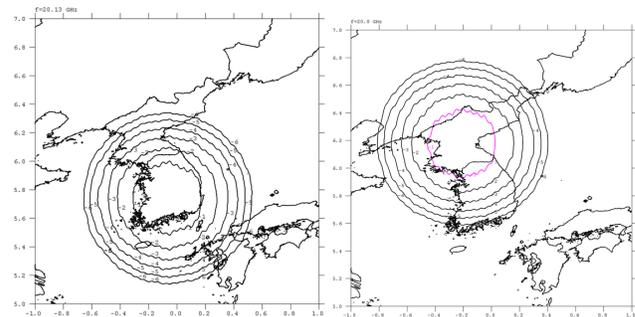
Figure 1. COMS service coverage for SK and NK beams

II. COMS ANTENNA AND GROUND TEST

The antenna system consists of two reflectors installed on the east and west panel of the spacecraft. Each reflector mounted on Antenna Deployment and Trim Mechanism (ADTM) comprised of offset reflector integrated with backing structure and sun shield, as illustrated in Figure 2. The ADTM provides on orbit capability to adjust beam pointing. The reflecting surface is graphite and support structures are honeycomb sandwich with graphite/epoxy face sheets and aluminum core. East reflector is illuminated by two feed horns generate South Korea and China Beams. West reflector is illuminated by single a feed horn to generate North Korea beam.
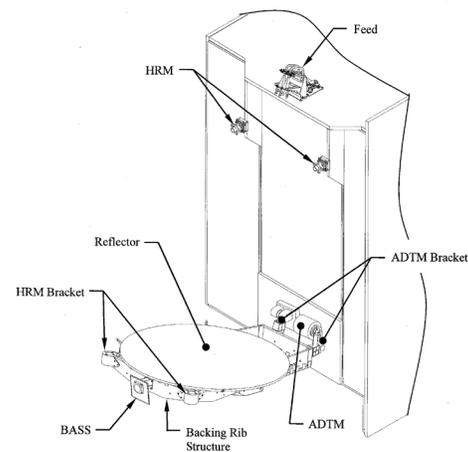
Figure 2. Illustration of Antenna Configuration (East Panel)

At unit level, pattern tests were performed at ETRI (Daejeon, Korea) Near Field Range before and after environmental tests for various parameters, such as gain, side-lobe, cross-polarization, co-polarization isolation, alignment information between reflector and feed. At the system level, these tests were independently performed by MDA (Montreal, Canada) in Compact Antenna Test Range (CATR). The two sets of test results are summarized in literature [5].

III. IN ORBIT TEST METHODOLOGY

After successful launch of a satellite and bus In Orbit Test (IOT), a series of measurements are performed to verify and identify if any mechanical or RF parameters of satellite antennas are degraded or changed. The satellite antenna IOT is essentially Power Flux Density (PFD) and Effective Isotropic Radiated Power (EIRP) measurements over selected points of the coverage pattern carried out by accurately measuring at an earth station. One of the most important tasks is thus earth station and test equipment calibration. ETRI did this calibration on the earth station located at Daejeon, South Korea, prior to IOT commencement using operating KOREASAT satellites. Antenna pattern tests involve a number of single (E- or H-) plane radiation pattern cut measurements [6][7]. The measurement can be done point by point (direct measurement, no processing required). In modern satellites, a antenna pattern cut can also be performed much faster by slewing the satellite at known speed and recording in a computer the earth station power level measurements time stamped that need complex data processing

A. Antenna IOT Objectives

The Antenna IOT, also known as antenna mapping, is not a replacement of range test because IOT tests are subject to many uncertainties and is not as detailed as CATR. IOT is a confirmation test with following objectives:

- Confirm that no RF or mechanical or damage due to harsh launch environment
- Ascertain correct reflector pointing and recommend if ADTM adjustment required
- Confirm pattern shape integrity and identify if there is any RF performance degradation
- Validation of TED model

B. Special Test Considerations

At Ka-Band IOT, weather impact is significant and antenna IOT was carried out only during dry periods. Antenna pattern cuts are relative measurements and error will be minimized if the variations of measurement uncertainties attributed by weather and equipment drift can be minimal during each of the pattern cuts. In consideration of above, for COMS Ka-band antenna pattern cut measurements, spacecraft slewing method [8] has been selected that implies measuring the SFD and EIRP at

Daejeon earth station while the satellite with antenna is moved at known speed that results a pattern cut. An antenna pattern cut was completed within 15~20 minutes instead of couple of hours with point by point method [8].

C. IOT Data Processing

Astrium (Toulouse, France), the satellite manufacturer using their restitution software processed the satellite positional time stamped data and ETRI, Korea processed the ground measured data also time stamped. ETRI/Telesat software was used to synchronize the two data sets that enabled to extract each pattern cut plots. The IOT measured pattern cut plot is then compared with ground based measurements to evaluate in orbit performance. Antenna IOT, also known as antenna mapping, is not a replacement of range test because IOT tests are subject to many uncertainties and is not as detailed as in CATR.

IV. IOT TEST RESULT SUMMARY

Astrium, using their restitution software processed the satellite positional time stamped data. ETRI extracted satellite position to measured signal strength data for each cuts. Because the beams are unshaped circular, it was considered to have only two pattern cuts (E- and H) for each beam instead of multiple cuts [9] for shaped beam antennas. This philosophy was adopted primarily to save antenna IOT time. ETRI/Telesat software was used to compare each pattern cut with ground based CATR measurements to validate beam shape integrity and identify precisely the desired bore-sight adjustment. In simple terms the software did try to match each pattern cut by shifting incrementally in both azimuth and elevation until best match is achieved for both E- and H-plane cuts.

A. Antenna Pattern Shape Integrity

Figures 3 and 4 show typical E- and H-plane plots of antenna pattern cuts measurements respect to the CATR measurement for South Korea and North Korea Beam. The plots show that the antenna patterns measured during IOT are very similar with CATR measurement.
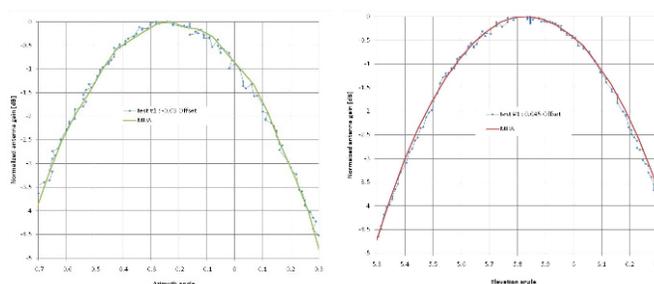
Figure 3.   South Korea Beam Plot (East Reflector)
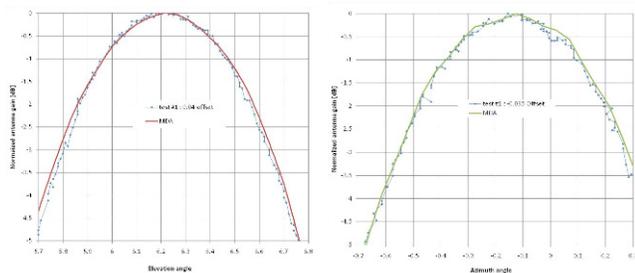
Figure 4.   North Korea Beam Plot (West Reflector)

B.   Antenna Beam Pointing Adjustment

The IOT data analysis predicted an average de-pointing of 0.045^0 for South Korea and 0.040^0 for North Korea Beam. This implies re-pointing of 0.025^0 for East reflector and 0.020^0 for West reflector considering that feeds are fixed and ration between mechanical and RF beam shift is 1+BDF (~1.9). But, antennas still meet the specified EIRP and G/T requirements with above de-pointing.   Such a small de-pointing is within measurement tolerance, and therefore, it is considered that no ADTM adjustment is necessary for either reflector

C.   Antenna Thermo Elastic Distortion (TED) Validation

An extensive thermal deformation analysis were carried out during design phase and maximum deformation is found to be 0.276mm for West reflector at begin of life for sun normal illumination, which corresponds to directivity loss of less than 0.2dB at edge of coverage. The thermal analysis also shows that maximum de-pointing due to TED is less than 0.02^0 and occurs around equinox.   The IOT was performed around summer solstice when the deformation is minimal and gain loss is below measurement tolerance and therefore no useful thermal distortion tests could be conducted.

D.   Cross-polarization and pointing Confirmation Tests

Antenna IOT had been performed at a single frequency and it is industrial standard to perform IOT based on co-polarization measurements.   This is because cross-polarization measurements are not accurate enough to get a meaningful result.   Both co- and cross-polarization performances were measured at three selected locations for each beam at low, mid and high frequency transponder. These results were compared to ground based predicted EIRP and SFD.   The measured values were within measurement uncertainties (0.5dB for co-polarization and 2dB for cross-polarization) and thus confirms both pointing and cross-polarization.

V.   SERVICE COVERAGE MEASUREMENT OVERVIEW

The purpose of service coverage measurement is to check how much EIRP contour map of COMS in orbit is similar with contour map gathered from the ground test in CATR facility. For the service coverage measurement, it needs to measure COMS Ka-band signal power over the South Korea

area. The service coverage measurements need two system supports, master ground station and mobile station. The master ground station located in ETRI transmits un-modulated RF signal and experimental 3DTV broadcasting signal to COMS. The master ground station receives measure returned un-modulated RF signal from the COMS and it also receives Ka-band beacon signal from the COMS. The measured un-modulated RF signal levels in the master ground station were used for reference values for mobile station measurements. The measured beacon signal was used for signal compensation attenuated by atmosphere.

The mobile station receives un-modulated RF signal generated by master ground station and repeated by COMS. Received beacon signal level was used for signal compensation attenuated by atmosphere to eliminate disturbance form the atmosphere condition changes. 3DTV broadcasting signal quality was measured by professional receiver to check Es/No and packet CRC errors. Figure 5 shows the simplified diagram for service coverage measurement [10].
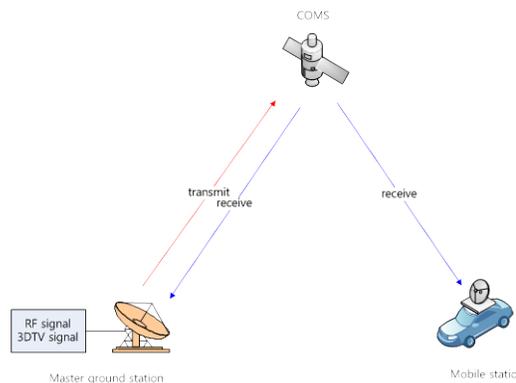
Figure 5.   COMS service coverage measurement diagram

Master ground station has 7.2 meter size of diameter satellite antenna and 175W output power high power amplifier. Mobile station has 1.8 meter size of diameter antenna with auto tracking equipment. Figure 6 shows the master ground station and mobile station, respectively, used for COMS service coverage measurement.

Figure 6.   Master ground station antenna and mobile station

The 17 measurement points over the South Korea peninsula are selected for mobile station measurement. Except reference point, all measurement points were selected on coast area because field measurements are to check

service coverage of COMS Ka-band payload. Selected measurement points are shown in Figure 7.

- Reference point: ETRI in Daejeon(14)
- West coast area: Ganghwa(1), Taean(2), Byeonsan (3), Jindo(4)
- South coast area: Goheung(5), Geoje(6), Busan(7)
- East coast area: Pohang(8), Uljin(9), Sokcho(10), Goseong(11)
- North area: Chuncheon(12), Pochecon(13)
- Jeju island: Chagwido(15), Mosulpo(16), Sungsan (17)
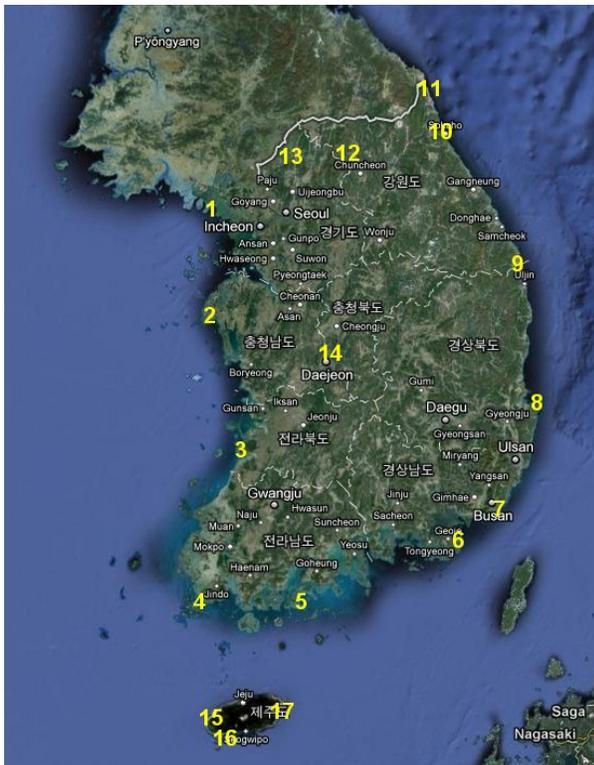
Figure 7.   Measurement points for service coverage

Figure 8.   Service coverage map of South Korea beam
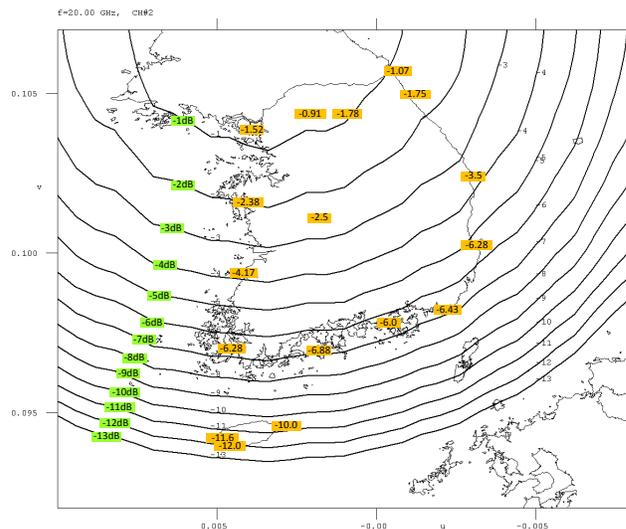
Figure 9.   Service coverage map of North Korea beam

VI.   SERVICE COVERAGE MEASUREMENT RESULTS

Measurement data from the mobile station at 17 measurement point were normalized by the measurement data at the reference point. The signal variation due to the atmospheric condition was compensated by the beacon signal level data. Figure 8 shows the service coverage contour map for South Korea beam and Figure 9 shows the service coverage contour map for North Korea beam respectively.

The above figures show that field measurement results over the South Korea at 17 points are similar with ground test results (solid line). The deviation between field measurement and ground test results are less than 1dB. This deviation value is acceptable when we consider measurement equipment stability of each ground stations and mobile station antenna pointing errors during field measurement. The reception of experimental 3DTV service was good for all measurement points. There were enough Es/No margins (3.5~8.5dB) and no packet CRC errors for 3DTV service over the South Korea. The field measurement results show that COMS Ka-band service coverage is formed well in orbit as we designed.

VII.   CONCLUSION AND FUTURE WORKS

In this paper, we reviewed in orbit antenna pattern test results and service coverage measurement results of COMS.

The in orbit antenna pattern cuts have shown excellent agreement confirming beam shape integrity. The EIRP and SFD measurement at selected locations was well within measurement uncertainties when compared to that predicted from ground based measurements. The field measurement of COMS service coverage for South Korea beam and North Korea beam show similar results with ground measurement contour maps in CATR facility. The similarity between in-orbit test results and ground test results are confirms that Ka-band antenna was well withstands launch environments and also performances are not degraded in space environments. At this moment, 3D-HDTV broadcasting and broadband VSAT communications are in trial service through COMS Ka-band payload.

REFERENCES

[1] J. H. Lee, J. H. Kim, S. P. Lee, and H. J. Lee, "Multi-Beam Satellite Antenna Design" IEEE Intl Symp on Antennas and Propagation, Session 84, July 2004

[2] S. P. Lee, "Development of Satellite Communication System for COMS", APSCC 2004, pp. 71-76

[3] S. P. Lee, J. W. Eun, K. S. Choi, C. S. Sin, J. W. Park, and Y. M. Lee, "Conceptual Design of the Satellite Communication System for Communications, Ocean and Metrological Satellite", KICS 2004 fall conference, pp. 148-152.

[4] Yoola Hwang, Byoung-Sun Lee, Hae-Yeon Kim, Hedong Kim, and Jaehoon Kim, "Orbit Determination Accuracy Improvement for Geostationary Satellite with Single Station Antenna Tracking Data", ETRI Journal, Vol. 30, No. 6, pp. 774-782, 2008.

[5] J. S. Choi, S. P. Lee, and I. B. Yom, "Performance Test of Ka Band Multi-Beam Antenna", Ka and Broadband Communications, Navigation and Earth Observation Conference, Cagliary, Italy, Sept 2009.

[6] Y. Imaizumi, Y. Suzuki and K. Ohata "In-Orbit Antenna Pattern Measurement of ETS-VIII", European Conference on Antennas and Propagation EUCAP 2007, Edinburgh, 11-16 November 2007.

[7] F. Rahman, A. Shoamanesh, N. Martens, P. Margittai, S. Reid, and T. Auyeung, "In Orbit Measurement of Satellite Communication Payload", IEEE Pacific Rim Conference, Victoria, Canada, May 1993

[8] W. D. Kelley, J. W. Opiekun, and S. L. Teller, "In orbit Satellite Antenna Measurement" IEEE Intl Symp on Antennas and Propagation Salt Lake City, USA, pp. 542-545, July 2000

[9] F. Rahman, "Boresight Pointing Calibration of Satellite through In-Orbit Antenna Tests", IEEE Transactions on Instrumentation and Measurement, vol.42, pp. 553–556, April 1993

[10] J. H. Jo, M. H. You, and S. P. Lee, "Field measurement of beam coverage for COMS Kaband communication payload" KICS summer conference, 8A-38, June 2011

Determining the Accuracy of the On Board Propagation Software for Optical Intersatellite Link

Umit C. Yilmaz
Faculty of Engineering
Karadeniz Technical University,
Trabzon,Turkey
Turksat Satellite Control Center
Golbasi, Ankara, Turkey
cyilmaz@turksat.com.tr

Ismail H. Cavdar
Faculty of Engineering
Karadeniz Technical University,
Trabzon,Turkey
cavdar@ktu.edu.tr

Cemal Sakaci
Turksat Satellite Control Center
Golbasi, Ankara, Turkey
csakaci@turksat.com.tr

Abstract- **In this paper, the accuracy of the satellite on board orbit calculation is compared with a ground based system calculation. The objective is to determine whether the on board satellite orbit propagator software is sufficient enough for Optical Intersatellite Link communication between a Low Earth Orbit and Geosynchronous Earth Orbit satellites and figuring out the possible reasons of any inaccuracy if any. The study is done by using TURKSAT 3A Geostationary satellites' data. Even the software on board the TURKSAT 3A does not intend to provide fine orbital position for intersatellite communication, this study is done in order to conclude the convergences and open points on the onboard orbit propagator. First, a quick summary is given for TURKSAT 3A satellite and its' station keeping philosophy. In the following sections, comparisons are done between on board and on ground orbit calculations with different time intervals of the year 2010. The possible root causes of the differences between these two calculations are discussed at the end.**

Keywords – Turksat; Optical Intersatellite Link; On Board Propagation Software.

I. INTRODUCTION

After the invention of lasers, lots of studies have been done for laser communications because of its theoretical advantages of higher data rates, lower power, smaller size, and lower mass. The studies for free-space laser communications have been in development in the United States since 1960s and in Japan and Europe since 1970s [1]. The advantage of small wavelength and low beam divergence make lasers attractive for optical communication [2], especially for intersatellite link communication between an observation satellite and a Geostationary Earth Orbit (GEO) satellite. It usually takes long time to send the image taken by a Low Earth Orbit (LEO) observation satellite to Earth. Therefore, one of the advantages of the intersatellite link is that it dramatically reduces the access time of a requested image of a land without a need of a polar station [3], [4], [5].

On the other hand, having a narrow laser beam came with the problem of tracking and finding the target spacecraft (SC). Since the access time between LEO and GEO satellite is limited due to different orbits, it is a time critical operation to find and lock the target SC as soon as both satellites see each other to begin the communication. That's why, finding the target SC on time give more chance to exchange more data between LEO and GEO satellites. And the GEO satellite can send the required information to ground station almost real time.

Successful laser communications between LEO and GEO satellites have been achieved after lots of studies. Two of the main steps of optical intersatellite links (OISL) are; the successful communication between French SPOT-4 (LEO) and ARTEMIS (GEO) on November 2001 and bidirectional laser communication ARTEMIS - Japanese OICETS (LEO) on December 2005. These tests showed that the GEO satellites can be very useful for relay purpose for OISL.

In order to compensate the movement of the satellites, known Ephemerides data and the signals from the electro optics tracking systems are used [6]. The concept of optical intersatellite links in communication satellite networks is discussed in [7] and [8]. The related studies are focused on: satellites optical communication network [9], analyzing the impact of random pointing and tracking errors in coherent and incoherent optical intersatellite communication [10], proposals for adapting the control system to reduce the vibration effect on the satellite[11]. Taking into account the classification of Pointing Errors in intersatellite link [12] ephemeris errors are one of the point-ahead errors which have to be minimized.

In this paper, a series of studies were done on a flying satellite (TURKSAT 3A) in order to calculate whether on board ephemeris propagation is sufficient enough for OISL. Firstly, a quick summary is given for TURKSAT 3A satellite and its' station keeping philosophy. Then, comparisons are done between on board and on ground orbit calculations with different time intervals of the year 2010 in the following sections. At the end, the possible root causes of the differences between these two calculations are discussed.

II. TURKSAT 3A AND ITS' STATION KEEPING PRINCIPLE

TURKSAT 3A, which was built on a Thales Alenia Spaces' SB4000 satellite platform, launched into space on 12 June 2008. It was placed into orbit at 42 degree East longitude where it is collocated with TURKSAT 2A and it has been working properly at that position since then.

Like most of the GEO satellite operators, Türksat AS. choose to make North and East maneuvers in 14 days cycle basis. Fig. 1 shows the Orbit Determinations (OD), North Maneuver (NS), On Orbit Propagator (OOP) initial orbital parameters update and East Maneuver (EW) plan in one cycle [13].
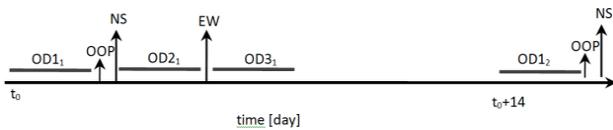
Figure 1. Station Keeping Principle of TURKSAT 3A

The ODs are performed by using an antenna which has tracking capability. During the year 2010, three different antennas, all of which are located in Ankara, Turkey, were used for this purpose. The orbit determination process uses the distance between the station to satellite and the ground station antenna's azimuth and elevation angles during the measurements. During one cycle, three orbit determination campaigns are performed:

1st, before the North maneuver to be used for both calculation of North maneuver and initial Keplerian parameters of OOP on satellite board.

2nd, after the North maneuver in order to calculate the efficiency of the applied maneuver and its cross coupling effect in tangential direction. This determination also used for East Maneuver calculation.

3rd, after the East maneuver to determine the efficiency of the maneuver and the last orbit after the maneuver sequence.

As it can be seen in the Figure 1, there are two maneuvers between two OOP update operations and three orbit determinations. That's why, the On Ground Orbit Calculation (OGOC) considered as more reliable depending on the antenna tracking performance. This calculation take into account all perturbation effects, SC mass, effective cross-section area opposed to sun and tracking antennas angles and distance when calculating the new orbit. Similarly, the OOP also take into account the perturbations, thruster activities on satellite and all other parameters special to Spacecraft.

Since the cumulative errors and on board unexpected activities may disturb the orbit, there should be a gap between OGOC and OOP. This gap may remain or change as long as new initial OOP parameters loaded. This means that Ephemeris data, which can be used for

rough pointing to target SC [14], may be different than expected.

III. CALCULATION OOP AND OGOC DIFFERENCES

A. Philosophy of the Calculations

The analysis is done by dumping the on board orbital telemetries at specific time interval and propagating the freshest orbit on ground for the same time. Since the on board orbit propagator does not give semimajor axis value, it has been calculated by using longitude drift parameter [16]. After that, the Keplerian parameters were transferred to position and velocity [15]. The Euclidean distance has been calculated from rms of the position errors.

The following figures are representing the differences between Keplerain parameters, position and velocity differences and Euclidean differences between these two calculations.

B. Comparison of OOP and OGOC for the year 2010

First, one year analysis was performed for the year 2010; everyday, two points were taken into account for calculation. The on board OOP telemetries were dumped each day at around 5:00:00 GMT and 17:00:00 GMT and transferred to position and velocity. On ground side the up-to-the-minute orbit was calculated by OGOC. Table 1 shows the rms values of the difference between OOP and OGOC, where μ is the mean and the σ is the standard deviation of the corresponding data. In table, 'a' is the semimajor-axis, 'i' is inclination, 'e' is the eccentricity of the orbit, 'RAAN' is the Right Ascension of Ascending Node, 'AoP' is the Argument of Perigee and MeanAnom represents Mean Anomaly of the orbit. x, y, z and V_x, V_y and V_z are position and velocity of the satellite with respect to Earth Center J2000 reference frame.

TABLE 1. OOP-OGOC PARAMETER DIFFERENCES FOR THE YEAR 2010

Δ OOP - OGOC	max	Min	μ	σ	unit
Δ a	1.714	0.00144	0.6415	0.399	km
Δ i	0.00862	1.59e-07	0.002	0.00192	degree
Δ e	3.294e-5	5.55e-09	4.77e-06	5.3e-6	-
Δ RAAN	11.546	9.18e-05	2.0263	2.2452	degree
Δ MeanAnom	4.1619	7.74e-06	0.6219	0.5921	degree
Δ AoP	11.669	1.207e-04	2.18	2.1702	degree
Δ x	16.283	5.47e-03	3.1232	2.9645	km
Δ y	27.38	1.13e-03	2.888	2.6346	km
Δ z	6.2129	6.75e-03	1.5243	1.22	km
Δ Vx	1.5617	5.80e-05	0.21327	0.195	m/s
Δ Vy	1.1834	2.258e-04	0.2172	0.2165	m/s
Δ Vz	0.6775	3.284e-05	0.125	0.141	m/s
Δ Euclidean	21.756	1.43e-01	5.166	3.307	km

As can be seen in this table, the maximum Euclidean distance between on ground and on board parameters is 21.7564 km.
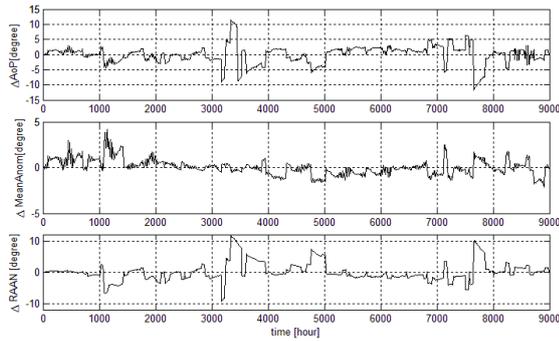
Figure 2. Differences in AoP, M and RAAN
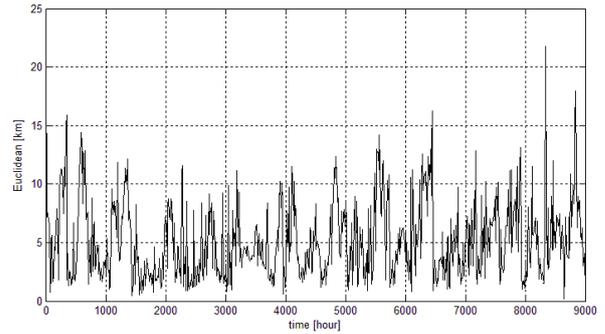
Figure 6. Differences in Euclidean distance for 2010
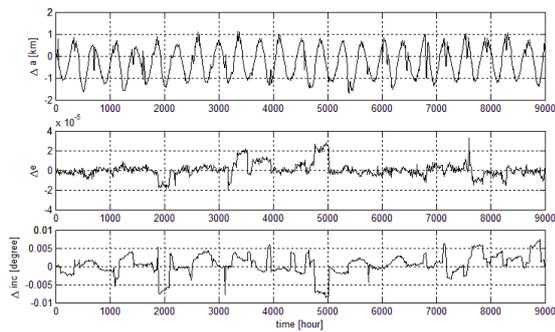
Figure 3. Differences in semimajor axis, eccentricity and inclination
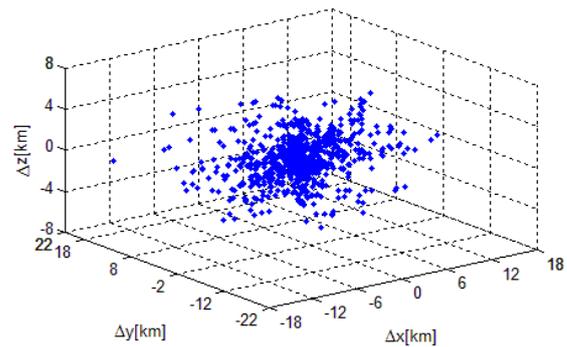
Figure 7. Differences in Euclidean in space for 2010

As shown in figures above and detailed in Table 1, for the year 2010 the mean Euclidean error between satellite calculation and on ground calculation is 5.116 km and the standard deviation is 3.307 km. Only 0.266% of the values are bigger than 20 km, 0.532% of the values are between 15 km to 20 km. 9.587% of the values are between 10 km to 15 km and finally 89.61% of the values are less than 10km.

C. *Two-Week Analysis (between 18.10.2010 – 02.11.2010)*

Further analysis was done between two OOP initialization operations, starting the calculation from 18.10.2010 to 02.11.2010 by comparing the orbital parameters every three hours. The following figures show the variations of position and velocity. It is observed from the orbital parameters, whenever the new OOP parameters are loaded on the board the differences are getting closer to zero. The following figures give the variations of position and velocity errors.

Figure 4. Variations in positions, x, y and z
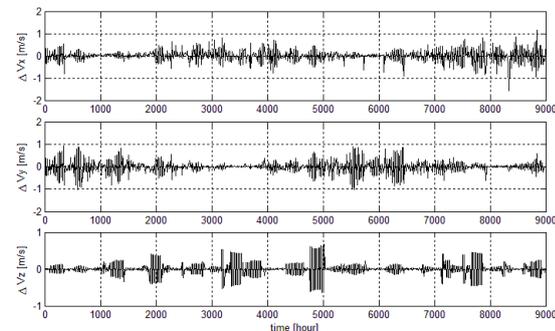
Figure 5. Variations in velocities Vx, Vy and Vz
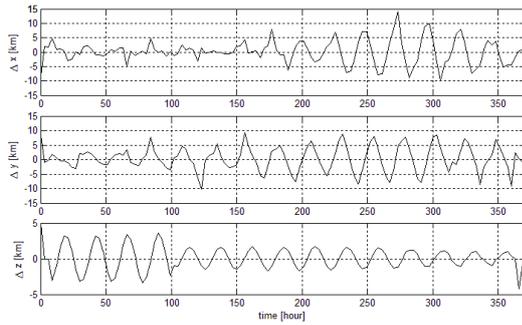
Figure 8. Variations in positions, x, y and z errors

Figure 11. Difference in Euclidean distance
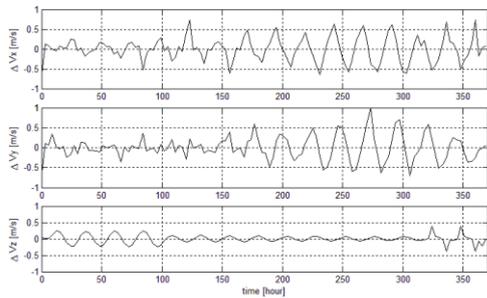
Figure 9. Variations in Vx, Vy and Vz errors

D. 4-Day Analysis (between 24.10.2010 20:00GMT – 29.10.2010 12:00GMT)

In order to analyze the OOP drift, the comparison study was done where there was not any thruster activity (North man, East man or auto wheel unloading) and new OD.

Starting from 24.10.2010 20:00 GMT to 29.10.2010 12:00 GMT, similar analysis was done with 1 hour interval. As can be seen from figures below the most drifting parameters during that time interval are Δx and $\Delta V y$.
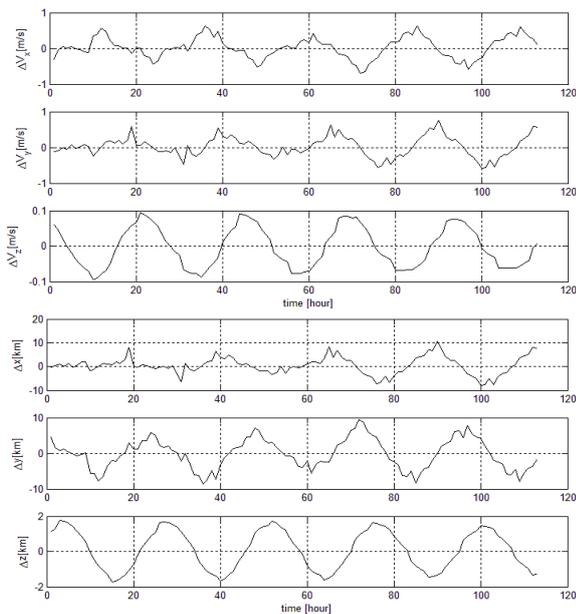
Figure 10. Differences in positions and velocities

IV. STUDY ON ANTENNA ANGLES USED FOR ORBIT DETERMINATION

As indicated in previous sections, most of the time the differences in Keplerian Elements are observed between the angle data; RAAN, Argument of Perigee and Mean Anomaly. The differences in semimajor axis, eccentricity and inclination do not have as much affect as other three parameters.

Further analysis was done in order to find the root cause of the differences in OOP and OGOC. Ranging measurements were analyzed for the year 2010. For that purpose, 75 ranging operations in 25 cycles were taken into account.

The software, which is used for orbit computation, calculates the standard deviation of the angle components and distance in normal distribution after calculating the biases. That's why the standard deviations may give some idea whether the qualities of the measurements are good or not.

After the studies on ranging measurements, it has been seen that, whenever the standard deviations of azimuth and elevation are respectively high, it affects all the Keplerian parameters in calculation. The distance measurement seems not as major disturbance as the angular ones.

V. CONCLUSION

In this paper, we have presented the comparison of orbital calculation using on board satellite computer and on ground orbit propagator. It can be seen that the accuracy of the orbital parameters are crucial and key factors in maintaining a pointing between a GEO and a LEO satellite due to limited time interval. One has to pay attention to load the most accurate initial values to satellites' On Orbit Propagator. The antennas which are used for the orbit determination must have a very accurate angular sensitivity in both azimuth and elevation. Most of the time less than 8 mdeg standard deviation has to be reached for angular resolutions. This issue is as much important as the tracking data that were used as initial parameters calculation for GEO satellite orbit propagator software. The maximization of the quality of the tracking antennas eventually gives better proximity to real position of the satellite in orbit and requires less tracking field of view for LEO satellite beacon and laser beam.

REFERENCES

[1] D. L. Begley, "Free-Space Laser Communications: The Agony and the Ecstasy", Laser and Electro-Optics Society 1999 12th Annual Meeting", Vol1 p 43-44, November 1999, USA

[2] J.E., Mulholland and S.A. Cadogan, "Intersatellite Laser Crosslinks" S.A, Aerospace and Electronic System, IEEE Transaction on, Vol. 32, Issue 3, pages 1011-1020,1996

[3] Z. Sodnik, B. Furch and H. Lutz, "Free-Space Laser Communication Activities in Europe: SILEX and beyond", IEEE Lasers and Electro-Optics Society, October 2006, pp. 78-79.

[4] N.Karafolas, Z. Sodnik, J. Maria, P. Armengol and I. Mckenzie,"Optical communications in space", ESA, 2009

[5] G.C. Baister and P.V. Gatenby, "Why optical communication links are needed for future satellite constellations",1996

[6] S. Arnon, "Power versus stabilization for laser satellite communication", Applied Optics Vol 38 No 15, May 1999

[7] Motorola Global Communications, Inc, "Application for Celestri multimedia LEO system" (Federal Communications Commission, Washington, D.C., 1997

[8] S.Arnon and N.S.Kopeika, "Laser satellite communication networks-vibration effects and possible solutions", Proc IEEE 85, 1646-1661 (1997)

[9] S. Arnon and N. S. Kopeika, "The performance limitations of free space optical communication satellite networks due to vibrations—analog case," Opt. Eng. 36, 175–182, 1997

[10] C. C. Chen and C. S. Gardner, "Impact of random pointing and tracking errors on the design of coherent and incoherent optical intersatellite communication links," IEEE Trans. Commun. 37, 252–260, 1989

[11] V. A. Skormin, M. A. Tascillo, and T. E. Busch, "Adaptive jitter rejection technique applicable to airborne laser communication systems," Opt. Eng. 34, 1263–1268, 1995

[12] H. Hemmati, "Near Earth Laser Communication", CRC Press, 2008

[13] U.C. Yılmaz, "Inclination Angle In Geostationary Satellites and Its' Effects on Communication", Karadeniz Technical University, MSc Thesis, Trabzon, Turkey, 2005

[14] S.Arnon, S.R. Rotman and N.S. Kopeika, "Bandwidth Maximization for Satellite Laser Communication", IEEE T-AES/35/2/04317, 1998

[15] C. Sakaci, "Two Simulation Models: Low Altitude Flows and the TURKSAT Satellite Orbit", Department of Aeronautical Engineering, Middle East Technical University, MS Thesis, Ankara, Turkey1996

[16] E. M. Soop, "Introduction to Geostationary Orbits", ESA, 1983

Complete Partitioning Policy with Different Handover Priority schemes for Multi-Class Traffic in LEO Mobile Satellite Systems

Amr S. Matar, Ibrahim I. Ibrahim, Gamal Abd Elfadeel and Hesham M. Z. Badr

Department of Communication, Electronics and Computer Engineering

Faculty of Helwan Engineering, University of Helwan

Cairo, Egypt

Email: eng_amr_s.matar@yahoo.com

Abstract—**In this paper, an analytical framework is developed to evaluate the performance of complete partitioning (CP) policy with two different handover priority schemes for multi-class traffic in Low Earth Orbit-Mobile Satellite Systems (LEO-MSS). The queuing of handover requests priority scheme is examined as the first priority scheme. Where, in the second priority scheme a combination of guard channel and handover request queuing priority schemes is developed.**

Keywords- LEO-MSS; multi-class; handover; queuing.

I. INTRODUCTION

The increase of public interest and mass marketing for mobile communications is leading to support third generation (3G) multimedia services requirements at anytime and anywhere. These make Mobile Satellite Systems (MSS) to be a good attractive choice for future global wireless communication networks. Compared with Geostationary Mobile Satellite Systems (GEO-MSS), LEO-MSS require lower transmit power, shorter propagation delay and higher traffic capacities. Therefore, LEO satellites are more suited for providing real-time interactive and multimedia services than other systems [1, 2].

The resource management strategies central issue for LEO-MSS system is to select the suitable policy for managing handover requests. From the user standpoint, the interruption of a conversation is more undesirable than blocking of a newly arriving call. Previous researches have considered various resource management strategies for LEO-MSS. One approach is to queue handover (QH) requests [3, 4]. Another approach for managing handovers (Guard channel) is to reserve resources before handover occurrence [5]. In the previous approaches, only single class of traffic was considered. For multi-class traffic, the performance analysis of a complete partitioning (CP) with or without fixed channel reservation was examined in [7].

Motivated by the above, in this paper, we present an analytical framework for evaluating the performance of LEO-MSS multi-class traffic using complete partitioning (CP) policy with two different handover priority schemes. In the first priority scheme, the handover requests queuing scheme is developed. Second, a combination of guard channel and handover request queuing approach is examined. The results are compared with the handover priority scheme developed in [7].

The remainder of this paper is organized as follows. In Section II, the basic assumptions are described. Section III presents a suitable mobility model. An analytical study for the CP policy with the two priority schemes is presented in Section IV. Section V shows the analytical results for the performance analysis. Finally, conclusions are drawn in Section VI.

II. BASIC ASSUMPTION

Similar to [9], this paper is based on the IRIDIUM system, which consists of 66 satellites orbiting over six near polar circular orbits at about 780 km of altitude. Due to the high value of the satellite ground-track speed, V_{trk} (about 26600 km/h in the LEO case), with respect to the user's motion, the relative satellite-user motion will be approximated by the vector V_{trk}. Moreover, mobile stations (MS's) cross the cellular network irradiated by a satellite according to a parallel straight lines.

The satellite footprint is divided into smaller cells or spotbeams in order to achieve efficient frequency reuse. Due to beam-forming, spot-beams are disposed on the earth according to a hexagonal regular layout (side R) with circular coverage of radius R'. The possible values for the ratio R'/R range is from 1 to 1.5 [9]. Clearly, the greater this ratio is, the larger the overlap area (between adjacent cells) as shown in Figure 1. Let us assume minimum possible extension for the overlap area such that $R'=R$. In the IRIDIUM case, the radius R is equal to 212.5 km. The centers of adjacent cells are separated by a distance equal to $\sqrt{3}R$.

To evaluate the performance of resource management strategies of multi-class traffic, the following qualities of service (QoS) parameters [7] are used:
1) P_{bk}, blocking probability of class-k new call attempts;
2) P_{fk}, handover failure probability of class-k calls;
3) P_{dk}, call dropping probability of class-k calls;
4) P_{usk}, unsuccessful call probability of class-k traffic.

Based on ITU-T recommendations for land mobile services [10], the values P_{bk} and P_{dk} should not exceed $5.10^{-4}, 10^{-2}$ respectively.

III. MOBILTY MODEL

In the following, let define source cell: the cell where the MS call starts and transit cell: any subsequent cell reached by the MS with the call in progress. From the call arrival in a cell, a random offset $z \in [-R, R]$ is associated to this call,

where z is the offset of the related MS according to the reference shown in Figure 1.

For class-k traffic, in order to characterize the user's (relative) mobility in multi-class traffic LEO-MSS's, we introduce the dimensionless parameter α_k as

$$\alpha_k = \frac{\sqrt{3}R}{V_{trk}T_{dk}} \qquad (1)$$

where

T_{dk} is the average duration time of class-k calls.

The proposed model for LEO mobility is based on the following assumptions [9]:

1) The new call origination is uniformly distributed over the mobile service area.
2) MS's cross the cellular network with a relative velocity, vector V_{trk} "orthogonal" to the side of the cells (as shown in Fig. 1).
3) When a handover occurs, the destination cell will be the neighboring cell in the direction of the relative satellite-user motion.
4) From the call arrival in a cell, the related MS travels a distance (depending on offset z) defined as:
   - Uniformly distributed between zero and $d(z)$, if the cell is the source cell of the call;
   - Deterministically equal to $d(z)$, if the cell is a transit cell of the call.

where

$$d(z) = \begin{cases} \sqrt{3}R, & if\ |z| \le \frac{R}{2} \\ 2\sqrt{3}(R - |z|), & if\ \frac{R}{2} \le |z| \le R \end{cases} \qquad (2)$$

Based on [6], the handover probabilities of class-k traffic from the source cell and transit one (P_{H1K} and P_{H2K}, respectively) are expressed as

$$P_{H1k} = \frac{2}{3}\left\{ P_{h1k} + \frac{1 - P_{h1k}}{\alpha_k} \right\} \qquad (3)$$

$$P_{H2k} = \frac{P_{h1k} + P_{h2k}}{2} \qquad (4)$$

where

$$P_{h1k} = \frac{1 - e^{-\alpha_k}}{\alpha_k} \quad , \quad P_{h2k} = e^{-\alpha_k} \qquad (5)$$

The channel holding time for a class-k call in cell x [8]:

$$t_{Hik} = \min[t_{dk}, t_{mci}], \quad i = 1,2. \qquad (6)$$

with expected value [7]:

$$E_k[t_{Hik}] = T_{dk}(1 - P_{Hik}), \quad i = 1,2. \qquad (7)$$

where $i = 1$ refer to the call in its source cell and $i = 2$ refer to the call in the transit cell.
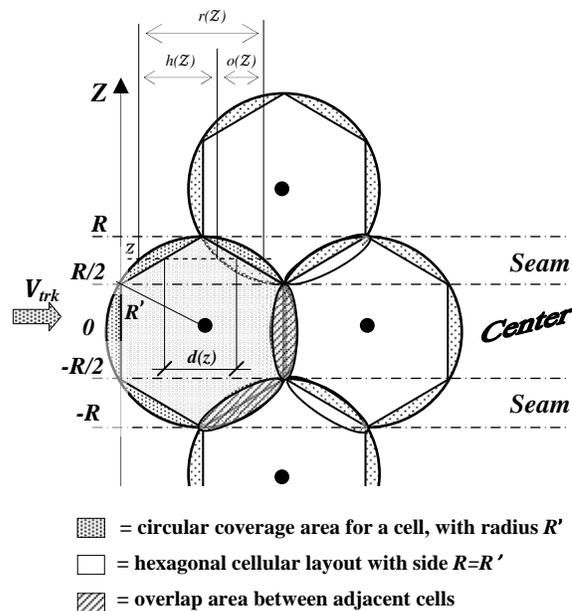
Figure 1. The shape of the cells and the distance crossed in the cell in the overlap area for a given height z.

IV. COMPLETE PARTITIONING POLICY PERFORMANCE ANALYSIS

In this section, analytical approaches for evaluating the CP performance with two different handover priority schemes for multi-class traffic are presented. In performing our analysis, we have assumed the following [6]:

- C channels are assigned per cell.
- The maximum number of the traffic classes in the system is K.
- New call arrivals and handover attempts of class-k traffic are two independent Poisson processes, with mean rates λ_{nk} and λ_{hk} respectively. And with λ_{hk} related to λ_{nk} by

$$\frac{\lambda_{hk}}{\lambda_{nk}} = \frac{2}{3}(1 - P_{bk})\left\{ \frac{P_{h1k}}{1 - (1 - P_{fk})P_{h2k}} + \frac{1 - P_{h1k} + (1 - P_{fk})(P_{h1k} - P_{h2k})}{\alpha_k - \alpha_k(1 - P_{fk})^2 P_{h2k}} \right\} \qquad (8)$$

- Whether class-k handover requests are queued or not, the channel holding time in a cell (for both new call arrivals and handovers) is approximated by a random variable with an exponential distribution and mean $1/\mu_k$ given by

$$\frac{1}{\mu_k} = \frac{\lambda_{nk}(1 - P_{bk})}{\lambda_{nk}(1 - P_{bk}) + \lambda_{hk}(1 - P_{fk})}E_k[t_{H1k}] + \frac{\lambda_{hk}(1 - P_{fk})}{\lambda_{nk}(1 - P_{bk}) + \lambda_{hk}(1 - P_{fk})}E_k[t_{H2k}] \qquad (9)$$

- The maximum waiting time is approximated by a random variable exponentially distributed, with expected value equal to $1/\mu_w = E[t_{w\ max}]$, where $E[t_{w\ max}]$ is obtained as

$$E[t_{w\ max}] = \frac{E[O(z)]}{V_{trk}} = \alpha_k T_{dk}\beta \qquad (10)$$

where $O(z)$ is the distance covered by the MS in the overlap area, which due to both the regular cellular

layout and the mobility assumptions, it remains the same for any handover request. While β is given by:

$$\beta = \frac{4}{9}\left(\frac{\sqrt{3}}{3}\pi - \frac{3}{2}\right) \approx 0.1394 \qquad (11)$$

Under the complete partitioning (CP) policy, all C channels available in a cell are partitioned into independent K subsets, with C_k ($1 \le k \le K$) channels allocated to class-k traffic and $C_1 + C_2 + \cdots + C_k \le C$. To efficiently assign the channels among traffic classes, an optimal channel partitioning scheme that can maximize channel utilization can be found in [11].

A. Complete Partitioning (CP) with Handover Queuing Priority Scheme

In this subsection, an analytical approach to queuing of handover requests scheme is developed. We denote C_k as the number of channels allocated to class-k traffic. In general, when there are free channels in the class-k subset, class-k new and handover calls are equally likely to get service. However, when all the C_k channels are occupied, class-k new calls are blocked whereas handover call requests are queued in their queue (Q) of Length L for a maximum time $t_{w\,max}$, waiting for a free channel. If the queue is full, the class-k handover calls are dropped.

Let $\Lambda(j)$ denotes the number of free channels in class-k channel subset in the cell j. According to this queuing scheme, the inter-beam handover requests are as follows:
1) If $\Lambda(j) \ne \emptyset$, the class-k new and handover calls get service immediately in cell j.
2) If $\Lambda(j) = \emptyset$, the class-k new calls are blocked and the class-k handover requests are queued waiting for an available channel in cell j. In the meantime, the handover call is served by its originating cell. A handover request leaves the queue for one of the following reasons:

   a) *The handover procedure is successful:* The handover request is served, before the call is ended and its maximum queuing time has expired.
   b) *The handover procedure has been useless:* The call ends before the corresponding handover request is served and its maximum queuing time has expired.
   c) *The handover procedure fails and the call is dropped:* The handover has not been performed within $t_{w\,max}$ and the call is not ended before its maximum queuing time has expired.

According to the queuing scheme described, the queuing scheme can be modeled as an $M/M/C_k/S$ queue. Its state is defined as the sum of the number of class-k calls in service and the number of queued class-k handover requests. The state transition diagram is shown in Figure 2. The steady state probability of the state j, P_j can be derived as:

$$P_j = \begin{cases} \dfrac{\lambda_k^j}{j!\,\mu_k^j}\,P_0, & 0 \le j \le C_k \\[3mm] \dfrac{\lambda_k^{C_k}\,\lambda_{hk}^{j-C_k}}{C_k!\,\mu_k^{C_k}\,\prod_{i=0}^{j-C_k}[C_k\mu_k + i(\mu_k+\mu_w)]}\,P_0, & C_k+1 \le j \le C_k+L \end{cases} \qquad (12)$$
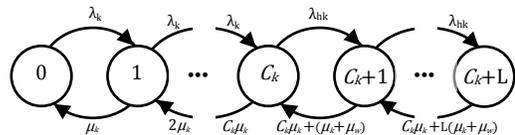
Figure 2. State Transition Diagram of CP Policy with Handover Queuing Priority Scheme.

where the rate $\lambda_k = \lambda_{nk} + \lambda_{hk}$ is the total class-k arrival rate and the idle system probability P_0 is

$$P_0 = \left\{ \sum_{j=0}^{C_k}\left[\frac{\lambda_k^j}{j!\,\mu_k^j}\right] + \sum_{j=C_k+1}^{C_k+L}\left[\frac{\lambda_k^{C_k}\,\lambda_{hk}^{j-C_k}}{C_k!\,\mu_k^{C_k}\,\prod_{i=0}^{j-C_k}[C_k\mu_k + i(\mu_k+\mu_w)]}\right] \right\}^{-1} \qquad (13)$$

Class-k new call arrivals are blocked when all the available C_k channels are in use. Therefore, the steady state blocking probability for the class-k new call (P_{bk}) can be expressed as

$$P_{bk} = \sum_{j=C_k}^{C_k+L} P_j \qquad (14)$$

Class-k handover failure occurs if a handover call arrival finds all class-k subset channels are occupied and its respective request queue is full *or* the handover call request is queued in its respective queue; however, it is dropped before getting service because its waiting time in the queue is expired before the handover call gets served or finished its service. The steady-state class-k handover failure probability is given as

$$P_{fk} = P_{C_k+L} + \sum_{i=0}^{L-1} P_{C_k+i}P_{fk/i} \qquad (15)$$

where the first term is describe the event that the class-k handover request queue is full. While the second term describes the event that the class-k handover call request is queued, but it is dropped before getting service because its waiting time is expired before a channel is released. The term $P_{fk/i}$ gives the probability of handover failure for a class-k handover call request in the queue given the handover call request joined the queue as the (i+1) call. This is found as [8]:

$$P_{fk/i} = \frac{(i+1)\mu_w}{C_k\mu_k + i(\mu_k+\mu_w)} \qquad (16)$$

The probability of an admitted class-k handover call being forced into termination during the ith handover can be expressed as

$$P_{dki} = P_{Fk}\left[P_{h1k}(1-P_{Fk})^{i-1}P_{h2k}^{i-1}\right] \qquad (17)$$

By summing over all possible values of i, P_{dk} can be obtained as follows

$$P_{dk} = \sum_{i=1}^{\infty} P_{dki} = \sum_{i=1}^{\infty} P_{Fk}\left[P_{h1k}(1-P_{Fk})^{i-1}P_{h2k}^{i-1}\right]$$

$$= \frac{P_{Fk}P_{h1k}}{1 - P_{h2k}(1-P_{Fk})} \qquad (18)$$

The unsuccessful call probability of class-k traffic P_{usk} is also used as an important parameter for evaluating overall system performance and can be derived as

$$P_{usk} = P_{Bk} + P_{dk}(1 - P_{Bk}) \qquad (19)$$
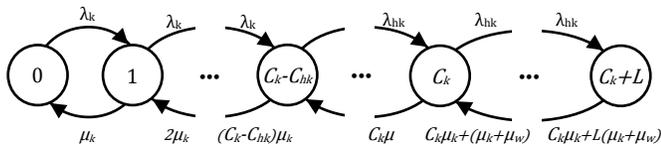
Figure 3. State Transition Diagram of CP Policy with Combination of Guard Channel and Handover Request Queuing Priority Scheme.

B. Complete Partitioning (CP) with Guard Channel and Handover Queuing Combination Priority Scheme

This subsection presents an analytical model for the combination of guard channel and handover request queuing scheme. In this model, when there are free channels in the class-k subset, class-k new and handover calls are equally likely to get service. However, when the number of occupied channels is equal to threshold (C_k - C_{hk}), class-k new calls are blocked whereas class-k handover calls are gets service. When all the C_k channels are occupied, class-k handover call requests are queued in their queue (Q) of Length L for a maximum time $t_{w\,max}$, waiting for a free channel according to the same scenario discussed in the previous scheme.

As it is shown in Figure 3, the queuing scheme can be modeled as an M /M /C_k /S queue. Its state is defined as the sum of the number of class-k calls in service and the number of queued class-k handover requests.

Let us analyze the state probabilities for the state transition diagram in Fig. 3, the steady state probability of the state j, P_j can be obtained as:

$$P_j = \begin{cases} \dfrac{\lambda_k^j}{j!\,\mu_k^j}\,P_0, & 0 < j \leq C_k - C_{hk} \\[2ex] \dfrac{\lambda_k^{C_k-C_{hk}}\,\lambda_{hk}^{j-(C_k-C_{hk})}}{j!\,\mu_k^j}\,P_0, & C_k - C_{hk} < j \leq C_k \\[2ex] \dfrac{\lambda_k^{C_k}\,\lambda_{hk}^{j-C_k}}{C_k!\,\mu_k^{C_k}\,\prod_{i=0}^{j-C_k}[C_k\mu_k+i(\mu_k+\mu_w)]}P_0, & C_k < j \leq C_k + L \end{cases}$$

(20)

where the idle system probability P_0 is

$$P_0 = \left\{ \sum_{j=0}^{C_k-C_{hk}}\left[\frac{\lambda_k^j}{j!\,\mu_k^j}\right] + \sum_{j=C_k-C_{hk}}^{C_k}\left[\frac{\lambda_k^{C_k-C_{hk}}\,\lambda_{hk}^{j-(C_k-C_{hk})}}{j!\,\mu_k^j}\right] \right.$$
$$\left. + \sum_{j=C_k+1}^{C_k+L}\left[\frac{\lambda_k^{C_k}\,\lambda_{hk}^{j-C_k}}{C_k!\,\mu_k^{C_k}\,\prod_{i=0}^{j-C_k}[C_k\mu_k+i(\mu_k+\mu_w)]}\right]\right\}^{-1}$$

(21)

Class-k new call arrivals are blocked when (C_k-C_{hk}) channels are in use. Therefore, the steady state blocking probability for the class-k new call (P_{bk}) can be expressed as:

$$P_{bk} = \sum_{j=C_k-C_{hk}}^{C_k+L} P_j$$

(22)

Similar to the previous scheme, the class-k handover failure probability can be computed as (15).

Using (18) and (19), P_{dk} and P_{usk} can then be computed, respectively.
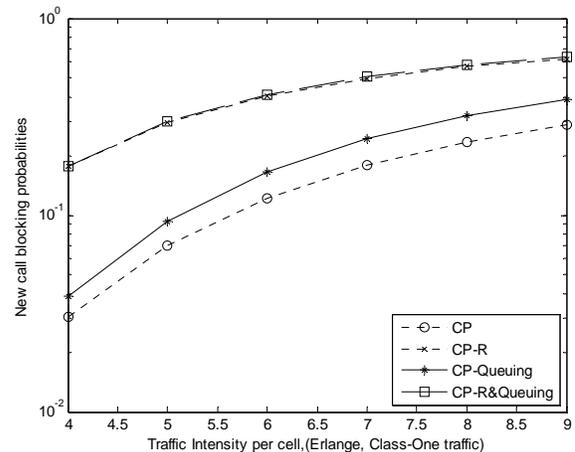
(a)
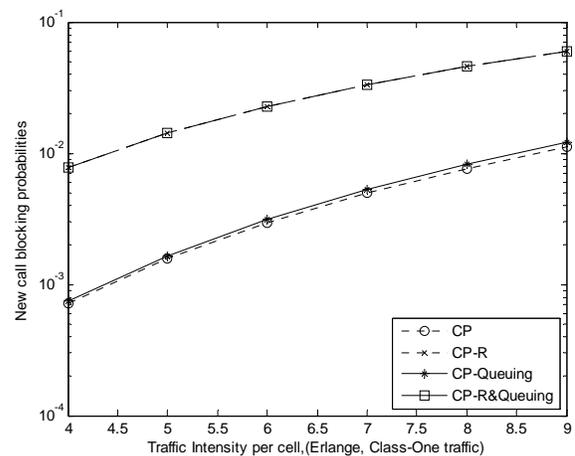
(b)

Figure 4. Analytical results for new call blocking probabilities as function of class-one traffic intensity of CP policy with different handover priority schemes.
(a) Class-One traffic.   (b) Class-Two traffic.

V. ANALYTICAL RESULT

In this section, we analyze the analytical results of the CP policy for multi-class traffic with the handover request queuing priority (named as CP-Queuing) scheme and the CP policy with the combination of guard channel and handover request queuing priority (named as CP-R&Queuing) scheme, which have been presented in section IV.

In the following, we consider from K-class of traffic in the cell two different class of traffic with the following parameter values: the total number of channel assigned for class-k traffic (C_1 = 8, C_2 = 4), reserved 25% of the total channel of each class for handover (C_{h1} = 2, C_{h2} = 1), the average duration time of class-k calls (T_{d1} = 180, T_{d2} = 540), the handover request queue length for class-k traffic are (L_1 = 4, L_2 = 2) and the traffic intensity of class-two traffic is 0.1 of the traffic intensity of class-one traffic.

(a)
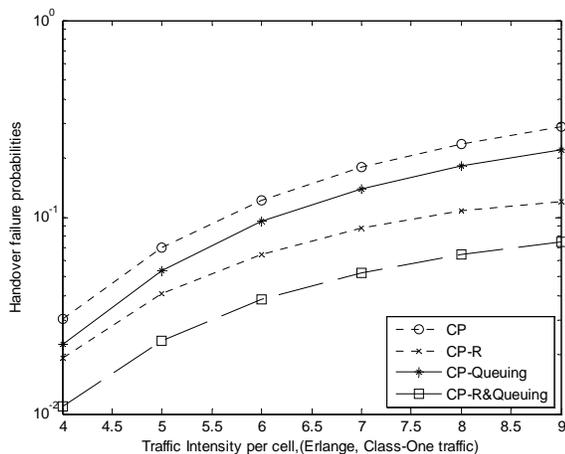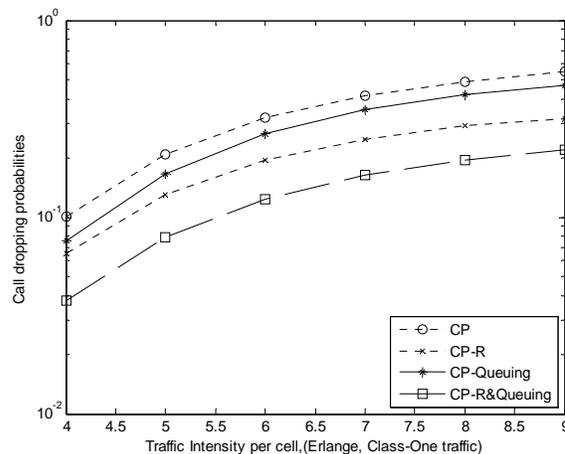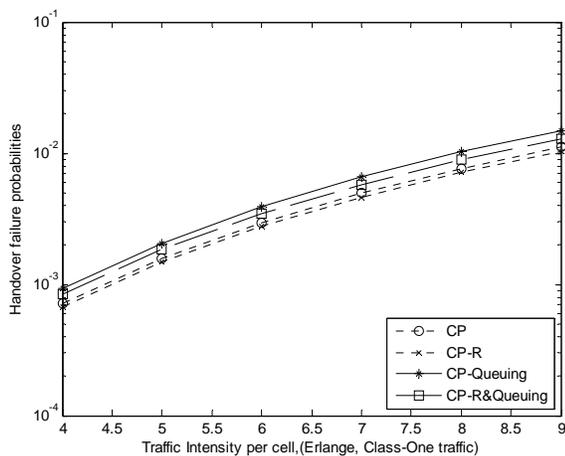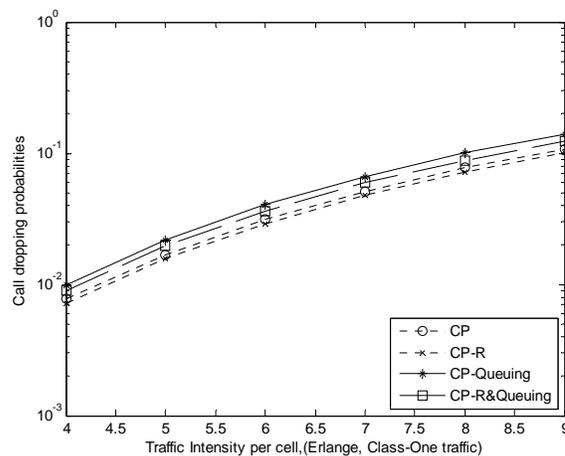
(b)

Figure 5. Analytical results for handover failure probabilities as function of class-one traffic intensity of CP policy with different handover priority schemes.
(a) Class-One traffic.   (b) Class-Two traffic.

(a)

(b)

Fig. 6. Analytical results for call dropping probabilities as function of class-one traffic intensity of CP policy with different handover priority schemes.
(a) Class-One traffic.   (b) Class-Two traffic.

Figures 4-7 shows analytical results of CP policy under different handover priority schemes in terms of P_{bk}, P_{fk}, P_{dk} and P_{usk} respectively. In these graphs, the behavior of CP with no priority (named as CP) and CP with fixed channel reservation (guard channel) priority (named as CP-R) scheme examined in [7] have been also considered.

In Figures 4(a) and 4(b), the analytical results for class-k new call blocking probability show that the handover queuing (CP-Queuing) priority scheme achieves a better performance than the handover queuing with guard channel combination (CP-R&Queuing) priority scheme for class-one and class-two traffic respectively. However, for class-two traffic the CP-Queuing scheme is very close in performance to non-prioritized CP scheme, it is slightly higher for class-one traffic. The difference between the two schemes is duo to that the handover call requests which wait in a queue will occupy the channel as soon as the condition of channel varies from busy to free. Therefore, the new calls have smaller opportunity of occupying channel than the handover calls.

From the Figures 5(a) and 6(a), we can see that the combination (CP-R&Queuing) priority scheme for class-one traffic in terms of handover failure probability (P_{fk}, see Figure 5(a)) and also of call dropping probability (P_{dk}, see Figure 6(a)) is the best among all other different priority schemes, and the performance difference between it and other schemes increase as the class-one traffic intensity increased. However, for class-two traffic the non-prioritized CP scheme and CP-R scheme attain better results than our schemes as shown in Figures 5(b), 6(b).

From the performance of unsuccessful call probability (P_{usk}, i.e. representing the fraction of new class-k calls that are not completed because of either being blocked initially or being dropped due to the failure of subsequent handover requests) of class-one and class-two traffic shown in Figures 7(a), 7(b) respectively, the handover queuing (CP-Queuing) priority scheme and non-prioritized CP scheme have a good response over other priority schemes. In Figure 7(a), the CP-Queuing scheme is less than non-prioritized CP scheme in
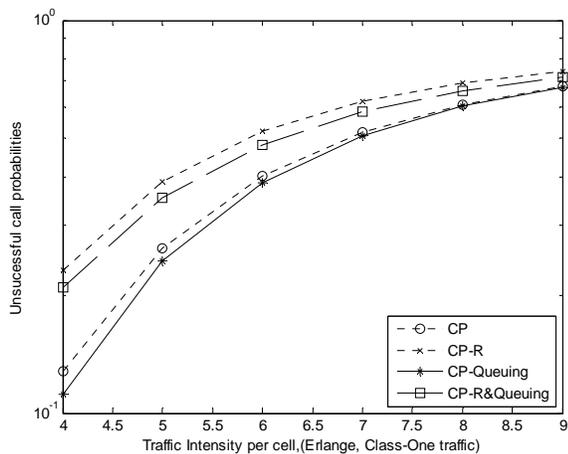
(a)
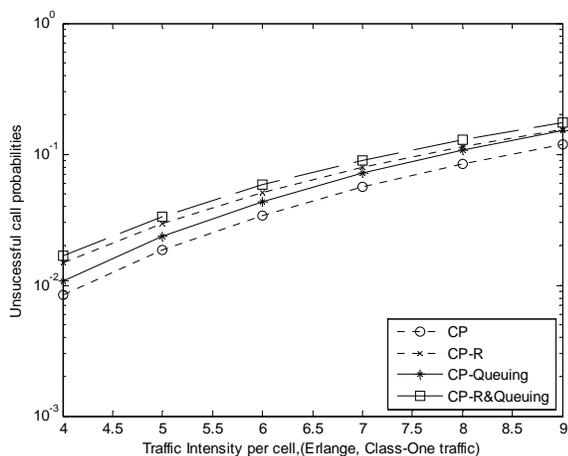
(b)

Figure 7. Analytical results for unsuccessful call probabilities as function of class-one traffic intensity of CP policy with different handover priority schemes.
(a) Class-One traffic.   (b) Class-Two traffic.

(a)

(b)

Figure 8. The effect of the handover request queue length (L_k) on the class-one and class-two traffic:

a)New Call Blocking Probability b)Handover Failure Probability

terms of P_{usk} at low traffic intensity. And as the class-one traffic intensity increase the performance begin to close to each other till be the same. However, for class-two traffic the non-prioritized CP scheme is the best performance.

For CP-Queuing priority scheme and as we can see in Figure 8, the increasing of handover request queuing length (L_k) has a approximately the same effect on the performance of new call blocking probability (Figure 8(a)) and handover failure probability (Figure 8(b)) for both class of traffic.

In the CP-R&Queuing priority scheme, the new call blocking probability increases significantly as the number of channel reservation increase as shown in Figure 9(a). The performance difference for class-two traffic is significantly higher than the class-one traffic; this is due to the small number of channel assigned for this class. This increase in P_{bk} results with a decrease in the handover failure probability as can be seen in Figure 9(b). However, for class-two traffic as the reserved channel increase it adopt with a little decrease in P_{fk}.

In conclusion, at low traffic intensity, the CP policy with queuing of handover requests scheme for class-one traffic effectively reduce the P_{bk} and P_{usk} at the expense of a little increase in the P_{fk} and P_{dk} than did the combination of handover request queuing with guard channel priority scheme. However, as traffic intensity increase, the performance difference in terms of P_{bk} and P_{usk} decrease to be very close and for P_{fk} and P_{dk} is increased. For class-two traffic, the non-prioritized CP priority scheme attains good result in all performance parameters.

VI. CONCLUSION

In this paper, we have developed an analytical work to evaluate the performance of CP resource management policy for multi-class traffic in LEO-MSS. Two different handover priority schemes have been introduced: the handover request queuing priority scheme and the combination of handover request queuing with guard channel priority scheme.
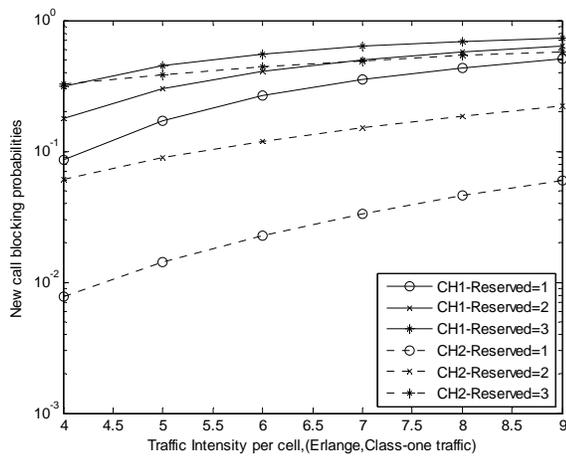
(a)
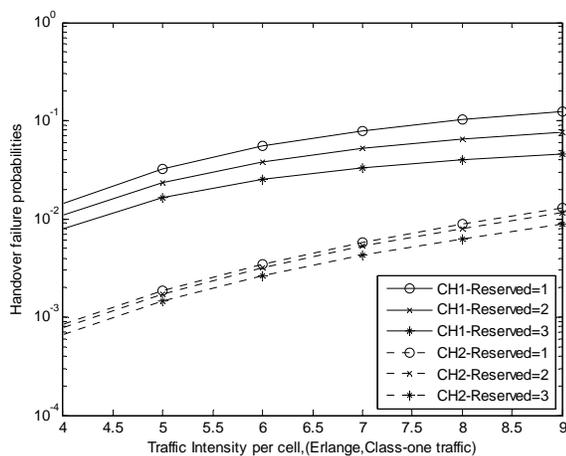
(b)

Figure 9. The effect of the number of reserved channel for handover
on the class-one and class-two traffic :
a) New Call Blocking Probability b) Handover Failure Probability.

Analytical results have shown that: for class-one traffic, the CP policy with queuing of handover requests scheme provides a good result at low traffic intensity. However at high traffic intensity, it is recommended to use the combination of handover request queuing with guard channel priority scheme. For class-two traffic, we have seen that the CP policy with no priority scheme attains the best performance results.

REFERENCES

[1] A. Jamalipour and T. Tung, "The Role of Satellites in Global IT: Trends and Implications," IEEE Pers. Commun., vol.8, pp. 5–11, June 2001.

[2] W. Kiamoucheand, M. Benslama, "A Novel Method For evaluating Parameters Of Ongoing Calls In Low Earth Orbit Mobile Satellite System," International Journal of Information Technology, vol. 5, pp. 43-47, 2009.

[3] E. D. Re, R. Fantacci, and G. Giambene, "Different queuing policies for handover requests in low earth orbit mobile satellite systems," IEEE Trans. Veh. Technol., vol. 48, pp. 448-458, Mar. 1999.

[4] Z. Wang and P.T. Mathiopoulos, "On the Performance Analysis of Dynamic Channel Allocation with FIFO Handover Queuing in LEOMSS," IEEE Trans. Commun., vol. 53, no. 9, pp. 1443-1446, Sep. 2005.

[5] E. Papapetrou, S. Karapantazis, G. Dimitriadis, and F.-N. Pavlidou, "Satellite Handover Techniques for LEO Networks," Int. J. Sat. Commun., vol. 22, no. 2, pp. 231–245, March/April 2004.

[6] A. S. Matar, G. A. Elfadeel, I. I. Ibrahim and H. M. Z. Badr, "Handover Priority Schemes for Multi-Class Traffic in LEO Mobile Satellite Systems," IJSCI, vol. 9, no. 1, pp. 46–56, Jan. 2012.

[7] Z. Wang, P.T. Mathiopoulos and R. Schober, "Channeling Partitioning Policies for Multi-Class Traffic in LEO-MSS," IEEE Trans. Aerospace and electronic systems, vol. 45, no. 4, pp. 1320-1332, Oct. 2009.

[8] D. Hong and S.S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoff procedures-Version 2a," College of Eng. Appl. Sci., State Univ. New York at Stony Brook, Tech. Rep. 773, June 1, 1999.

[9] E. D. Re, R. Fantacci, and G. Giambene, "Handover queuing strategies with dynamic and fixed channel allocation techniques in low earth orbit mobile satellite systems," IEEE Trans. Commun., vol. 47, no. 1, pp. 89–102, Jan. 1999.

[10] ITU-T Recommendation E.771, "Network grade of service parameters and target values for circuit-switched land mobile services".

[11] Z. Wang and P.T. Mathiopoulos, "Optimal Channel Partitioning and Channel Utilization for Multiclass Traffic in a LEO-MSS," IEEE Trans. Aerospace and electronic systems, vol. 46, no. 4, Oct. 2010.

A Raptor-based Adaptive Erasure Code
over Ka-band Space Channel

Zhou Jie, Jiao Jian, Yang Zhi-hua, Gu Shu-shi, Zhang Qin-yu

Communication Engineering Research Center of HITSZ
Harbin Institute of Technology Shenzhen Graduate School
Shen Zhen, China
zqy@hit.edu

Abstract—**In order to efficiently resist the influence of weather dynamics over Ka-band link, this paper designs a practical Raptor based erasure code combined with a forecast model, which has adaptive code rate on different weather. The forecast model could capture the weather correlation through the Gilbert-Elliot model, and lower the forecast errors by employing CFDP (CCSDS File Delivery Protocol) asynchronous NAKs (Negative Acknowledgement) to feed back the channel state information (CSI). We provide analytical results on forecast error and simulation shows that the algorithm can significantly enhance the decoding performance, especially in the bad weather state, and maintain the link availability and high throughput over Ka-band weather dependent link.**

Keywords-Ka-band; Raptor code; forecast model; CFDP asynchronous NAK mode

I. INTRODUCTION

The success of Mars Reconnaissance Orbiter (MRO) demonstrated the feasibility that future exploration missions employ the Ka-band [1-2]. Though 32GHz Ka-band signal increases the downlink throughput, file delivery is highly vulnerable to fluctuating weather states of ground station area and experiences severe errors which lead to frequent link outages [3-5]. For high reliable data return (i.e. with 99% weather availability), there is unexpectedly a 4.7 dB penalty at Ka-band compared to a 1.17 dB penalty at X-band [6].

To improve Ka-band link efficiency, many research focus on the rate adaption and long erasure correcting codes (LEC) at present [7-11]. The rate adaptive transmission scheme mitigates the adverse effect of fluctuation weather by dynamically adjusting the link margins, and it is also indicated the adaptive rate transmission scheme has significant advantages in the throughput and link availability compared to the constant rate transmission in [7], but there's no specific coding scheme and the Markov model used is less accurate. On the other hand, the existing LEC schemes are RS, LDPC (IRA code, GeIRA code) and LDGM codes. RS code has limited optional parameters and large complexity, and the LDPC code needs a long code-length. These limitations are beyond the capacity of the probers in processing and storage. Contrast to LDPC code, the rateless fountain code, as one of the LDGM codes, is more suitable to resist the burst error. Therefore, the paper discusses an adaptive concatenated erasure coding technique with small code-length and linear encoding and decoding complexity, which could mitigate the error-floor by concatenating a pre-coding scheme.

From the adaptive erasure code perspective, it is more meaningful to predict the weather state for the rate adaption. As a general method, Markov model could be exploited to forecast the weather state. However, the high forecast error in this model cannot meet the requirement of the reliable scientific data return. Hence, we propose a forecast model to degrade the forecast error, by employing CFDP asynchronous NAKs to feed back the channel state information (CSI). Based on the forecast model, we design a specific Raptor based erasure code with rate adaption to maintain link availability and high throughput. Raptor code, as a type of fountain code, could recover the whole file as the destination receives N packets a little more than k raw packets. Consequently the adaptive Raptor code could accommodate itself to the Ka-band weather dependent space channel.

The rest of this paper is organized as follows: Section II describes a Ka-band channel forecast model and gives the analytical results on the forecast error. The details of Raptor based concatenated erasure coding scheme are presented in Section III. Section IV discusses parameter design and throughput analysis of adaptive erasure coding scheme. In Section V, simulation and discussion are presented. Finally, Section VI gives the conclusion.

II. KA-BAND CHANNEL FORECAST MODEL

A. Gilbert-Elliot channel

Ka-band is highly vulnerable to weather impairments. The downlink data integrity is affected by packet errors badly, which are incurred by the system noise from the rain and cloudy. Therefore, it is necessary to establish a dynamics channel model to capture the weather correlation.

A rain attenuation of Ka-band model is built to simulate link burst BERs associated with the changing weather states in [3-5]. In this model, a noise temperature threshold is defined as T_{th}. If the sampled noise temperature is less than T_{th}, define the weather as good state, denoting it as *G*, and most of the transmitted packets will be received successfully; the space downlink is available with low BERs ($10^{-8} \sim 10^{-5}$). Otherwise, it is defined as bad state, representing it as *B*, significant error

will occur for the high noise temperature of the receiver antenna. Relatively high BERs ($10^{-4} \sim 10^{-3}$) are applied to capture severe packet loss rates (50%~99%), and make the link unusable. The dynamics of the Ka-band link are modeled as the Gilbert-Elliot model based on the AWGN channel, as shown in Fig. 1,
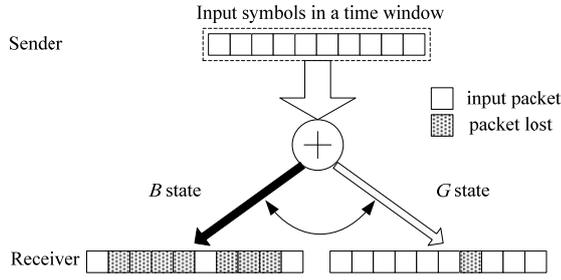
Figure 1.   Two error state AWGN channel Gilbert-Elliot model

Suppose we use BPSK modulation and we have

$$BER = Q\left(\sqrt{2E_b / N_0}\right) \qquad (1)$$

E_b / N_0 is the SNR, where $N_0 = K \cdot T_{th}$. Then, the relationship between noise temperature T_{th} and BER is:

$$T_{th} = \frac{2E_b}{K \cdot \left(Q^{-1}\left(BER\right)\right)^2} \qquad (2)$$

In GE model, the next state is only determined by the current state, so we define the transfer probabilities from G to B as $P(B|G) = \lambda_G$, and from B to G state as $P(G|B) = \lambda_B$, respectively. The state transfer matrix P can be written as:

$$P = \begin{bmatrix} P(G|G) & P(B|G) \\ P(G|B) & P(B|B) \end{bmatrix} = \begin{bmatrix} 1-\lambda_G & \lambda_G \\ \lambda_B & 1-\lambda_B \end{bmatrix} \qquad (3)$$

Using standard methods, the eigenvalues of P are found to be $\lambda_1 = 1$ and $\lambda_2 = 1 - \lambda_G - \lambda_B$, with corresponding eigenvectors $S_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ and $S_2 = \begin{bmatrix} \lambda_G & -\lambda_B \end{bmatrix}^T$, where T denotes transpose. Consequently, we can write

$$P = S \Lambda S^{-1} \qquad (4)$$

where $S = \begin{bmatrix} 1 & \lambda_G \\ 1 & -\lambda_B \end{bmatrix}$, $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$.

So the stationary probability of each state is:

$$\begin{aligned} P_G &= \lambda_B / \left(\lambda_G + \lambda_B\right) \\ P_B &= \lambda_G / \left(\lambda_G + \lambda_B\right) \end{aligned} \qquad (5)$$

We define the matrix in (3) as the one-step transfer matrix, which can predict the next time weather state with the current state. And we define the matrix in (6) as the m-step transfer

matrix, which can predict the weather state at the mth instants of time with the current state.

$$P^m = S \Lambda^m S^{-1} = \begin{bmatrix} P_G \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m & P_B \cdot \left(1-\lambda_2{}^m\right) \\ P_G \cdot \left(1-\lambda_2{}^m\right) & P_B \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m \end{bmatrix} \qquad (6)$$

Based on the Markov chain, if the channel at time 0 is in the G state, then the weather state forecast error at time m is $2 \times \left(P_G \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m\right) \cdot P_B \cdot \left(1-\lambda_2{}^m\right)$, while be in B state, then the forecast error is $2 \times P_G \cdot \left(1-\lambda_2{}^m\right) \cdot \left(P_B \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m\right)$. So the weather state forecast error $F_e(m)$ at time m associated with (5) is defined as:

$$\begin{aligned} F_e(m) = &\, P_G \cdot \left(2 \times \left(P_G \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m\right) \cdot P_B \cdot \left(1-\lambda_2{}^m\right)\right) \\ &+ P_B \cdot \left(2 \times P_G \cdot \left(1-\lambda_2{}^m\right) \cdot \left(P_B \cdot \left(1-\lambda_2{}^m\right) + \lambda_2{}^m\right)\right) \end{aligned} \qquad (7)$$

From (7), we can easily see the forecast error will become higher when the value of m increases. So we need exploit a high quality forecast model to reduce the error for the data integrity.

B. The forecast model based on the CFDP asynchronous NAK mode

CFDP provides a store-and-forward file delivery capability operating across an end-to-end space link [12-13]. In the acknowledged transmission mode, communication reliability is guaranteed through negative acknowledgment (NAK) issued by the CFDP receiving entity. Four NAK algorithms can be chosen to adopt the different scenarios. Considering the asynchronous NAK mode, the receiver can set a timer to control the file checking time and then feedback the results to ensure reliability [14]. In this paper, we build a weather state forecast model combined the asynchronous NAK mode, as shown in Fig. 2.
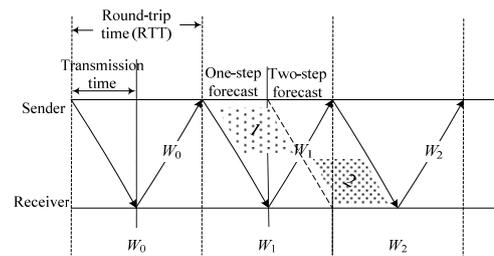
Figure 2.   The weather state forecast model

In this forecast model, assume the duration of a weather state W is equal to round-trip time (RTT) and the duration of each transmission is the full resolution of the raw data, which is about 1.44 minute time scale, chosen to be consistent with [3]. We define the duration of each transmission as a time window. Upon receipt of the first packet, the receiver will feed back the CSI on current weather state every other RTT until the file ends. It is in the state W_1 that the sender will obtain the state W_0 feed backed by the receiver after experiencing a

transmission time. In this period W_1, it needs to predict two steps weather parameters according to the previous state W_0 through the forecast model, as shown in following steps:

a) one-step forecast area: the first part of data will reach the receiver being in this period W_1 after a downlink transmission time (as "1" in Fig. 2), we can get the rate parameter for coding these data, using one-step transfer matrix.

b) two-step forecast area: the second part will be sent during the second half of the state and then arrive the receiver at the next state W_2 (as "2" in Fig. 2), which can be predicted by two-step transfer matrix for the parameter.

As the one-step and two-step areas are the same length in the state W_1, the forecast error F_e is expressed by

$$F_e = 0.5 \times F_e(1) + 0.5 \times F_e(2) \tag{8}$$
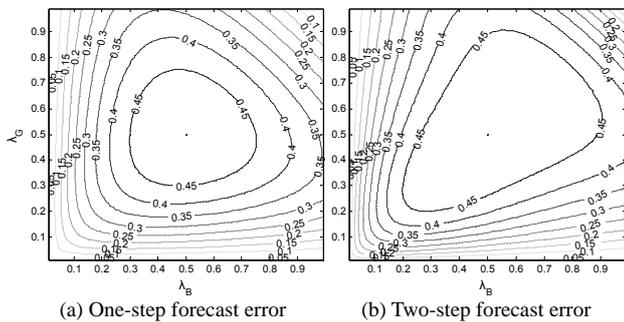
(a) One-step forecast error   (b) Two-step forecast error
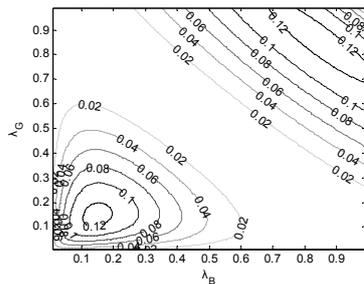
(c) The difference between the one-step and two-step forecast errors

Figure 3.   Forecast error on different transition probabilities

The Fig. 3 shows the influence of different λ_B and λ_G on the one-step and two-step forecast errors. It can be seen that the error is close to maximum of 0.5 with both of the transfer probabilities tending to 0.5 gradually. If any probability is less than 0.05 or both exceed 0.95, the forecast error is within 10%. This result illustrates when one weather state is stable for a long time or two states switch steadily, the model has high forecast accuracy.

Compared with (a) and (b) in Fig. 3, the two-step forecast error tends to maximum rapidly, which means two-step forecast error is larger than the one-step forecast error, using the same transfer matrix. Moreover the forecast error is larger when the transfer probability deviate more from 0.5. On the other hand, the error falls back to 0, when both of the transfer probabilities are close to edge (0 or 1), as shown in Fig. 3(c). Compared to the Markov model in Fig. 1, the forecast model

combined with the CFDP asynchronous NAK mode in Fig. 2 has a lower forecast error.

III. DESIGN AND PERFORMANCE OF RAPTOR-BASED CONCATENATED ERASURE CODING SCHEME

A. The design of concatenated erasure coding scheme

Based on the proposed forecast model, we design a concatenated erasure coding scheme with small code length and linear encoding and decoding complexity, as shown in Fig. 4.
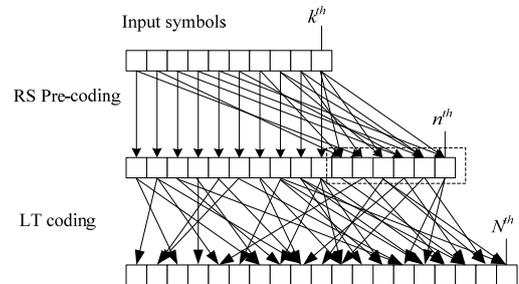
Figure 4.   Raptor based concatenated erasure code scheme

The Raptor based concatenated erasure code parameter is defined as (k, n, N, Ω), where k is the original information packets, n is the intermediate packets, N is fountain code erasure packets and Ω is degree distribution. The encoding process as follows:

a) genetate n intermediate packets from the k raw packets using RS pre-code.

b) generate N erausre packets from n intermediate packets using LT code with corresponding degree distribution.

And the decoding process as follows:

a) recovers n' $(\leq n)$ intermediate packets from the received erasure packets using BP algorithm.

b) then recover corresponding raw packets from the n' intermidiate packets using RS decoding algorithm.

The BER determines the efficient values of k and n. The sender computes the BER related with a weather state to acquire code rate.

B. The encoding and decoding performance of concatenated erasure code

As mentioned above, RS code, as the pre-code of Raptor code, could recover the k raw packets from the n' intermediate packets LT code recovers. In this section, we use weaken robust soliton distribution (WRSD) as the degree distribution with linear complexity, which can recover the majority of raw packets [10]. The generator function is defined as:

$$\Omega(x) = \left[ux + \sum_{i=2}^{D} x^i / (i \cdot (i-1)) + x^{D+1}/D \right] / (u+1) \tag{9}$$

If Raptor code adopts the WRSD, it has been proved that any set $(1+\varepsilon/2) \cdot n + 1$ of received packets are sufficient to recover at least $(1-\delta) \cdot n, (\delta = (\varepsilon/4)/(1+\varepsilon))$ intermediate packets via BP decoding, where the complexity is $O(\ln(1/\varepsilon))$.
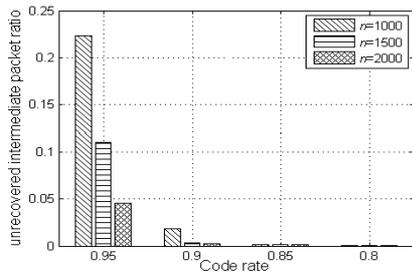
Figure 5. Recovery capability of fountain code with WRSD

Fig. 5 shows the relationship between code rate and unrecovered intermediate packet ratio as the raw packets are 1000, 1500 and 2000, respectively. It can be seen that the ratios are lower than 0.05 when code rate is less than 0.9. So we can recover all raw packets by means of adding the RS code with rate less than 0.95.

IV. PARAMETER DESIGN AND THROUGHPUT ANALYSIS OF ADAPTIVE ERASURE CODING SCHEME

A. Parameter design of adaptive erausre code

As mentioned in Section II, the BER varies with weather states transition, and we accommodate the BER variations by designing an adaptive parameter coding scheme. For the BER value, the forecast model in Fig. 2 can obtain it by forecasting the weather state when data arrive at the receiver. Consequently, we can achieve rate adaption as follows.

Given a time window T, we assume the downlink data rate R and the code length L, then $N = T \cdot R / L$ erasure packets will be transmitted during a time window. The relationship between packet loss rate and BER becomes:

$$P_{packet} = 1 - (1 - BER)^L \qquad (10)$$

The scheme handles the changing packet loss rates by adjusting code rates of the rateless fountain code adaptively.

GPC and BPC are defined as the G and B states parameter coding scheme respectively. We can receive N_G erasure packets successfully and recover k_G raw packets in G state with the BER P_{egood} and the same as N_B, k_B and P_{ebad} in B state. So the code rates of G and B state are given by:

$$Coderate_{good} = k_G / N \qquad (11)$$

$$Coderate_{bad} = k_B / N \qquad (12)$$

The above analysis in Fig. 5 explains that if the error correction capacity of RS code is more than δn, $(1 - P_{packet}) \cdot N$ erasure packets that the receiver get in certain BER, are sufficient to recover the entire file. When T and other communication conditions are fixed, the value of N is constant and there exist $P_{egood} \le P_{ebad}$ and $Coderate_{good} \ge Coderate_{bad}$. Hence, this paper proposes an adaptive parameter coding scheme APC, which accommodates the code rates of Raptor code adaptively based on the weather state transition. Its

statistical average code rate during a time window is obtained as follows:

$$Coderate_{adaptive} = \frac{P_G \cdot k_G + P_B \cdot k_B}{N} \qquad (13)$$

B. Analysis of link throughput

The GPC can recover the entire file in G state, but recover little in B state. A time window throughput of the GPC is represented by:

$$Throughput_{good} = P_G \cdot k_G \qquad (14)$$

The BPC could recover all information in both states and then maintain the link availability. Its throughput in one time window is given by:

$$Throughput_{bad} = k_B \qquad (15)$$

Compared with (14) and (15), the link occur outage when the GPC is in B state. On the other hand, though the BPC keeps the link continuous, it results in a small throughput. In this section, the proposed APC could maintain the link availability and high throughput by updating the parameters according to the CSI. Its statistical average throughput during a time window can be expressed by:

$$Throughput_{adaptive} = (P_G - F_e / 2) \cdot k_G + P_B \cdot k_B \qquad (16)$$

V. SIMULATION AND DISSCUSSION

We choose 20K as the noise temperature threshold in [3] and the one-step transfer matrix is defined as:

$$P = \begin{bmatrix} 0.9773 & 0.0227 \\ 0.1667 & 0.8333 \end{bmatrix} \qquad (17)$$

TABLE I. SIMULATION PARAMETERS

Parameter	Value
Code length L	1Kbyte
A time window T	1.44min
Downlink data rate R	20KBps
Weather state probabilities	P_G=0.88, P_B=0.12
Erasure packets of each transmission N	1728
The weather state forecast error F_e	0.0869
BER	G state: 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5} B state: 10^{-4}, 1.5×10^{-4}
GPC, BPC and APC code rates	G state: 0.828,0.824, 0.820,0.752 B state: 0.3357, 0.1325
LT code rates	G state: 0.69,0.689,0.684,0.614 B state: 0.2519, 0.1621

Table I shows the simulation parameters of GPC, BPC, APC and adaptive LT code for comparison. The values refer to the typical Earth-Mars communication scenario. Code rates are obtained by the corresponding simulation of coding schemes with the Monte Carlo under the decoding failure rate 10^{-4}. We execute 10000 simulations at each BER and average the result.

Figure 6. Throughput of GPC, BPC and APC with bad BER 10^{-4} and 1.5×10^{-4}. The throughput of the LT code is also included for comparison.

Fig. 6 shows the throughput in one time window of the GPC, BPC, APC and also adaptive LT code. It is indicated that with increased bad BER, the APC and LT coding schemes reduce their throughputs slightly, which illustrates that the bad BER is a factor of adaptive code throughput decline except the forecast error. Moreover, the proposed APC is superior to the adaptive LT code of same type with respect to throughput apparently. However, the throughput of GPC does not vary with the bad BERs since the link using GPC is equivalent to outage in the bad state. Similarly, the throughput of BPC is stable with the good BERs for $P_{egood} \leq P_{ebad}$. Both the APC and GPC have significantly better throughput than the BPC has. The APC is close to the GPC on throughput, but keeps the link continuous compared to the GPC.

VI. CONCLUSION AND FUTURE WORK

In this paper, for combating the adverse effects of frustration weather on data delivery, we design a Raptor based adaptive erasure coding scheme over a weather state forecast model, which has excellent encoding and decoding performance and low complexity. In the proposed scheme, CFDP asynchronous NAKs are employed to degrade the weather state forecast error. We design an adaptive parameter coding scheme APC on code rate, which could recover the total data especially in the bad weather state, enhance dramatically the decoding performance and keep the link available compared to the fixed parameter GPC. It also achieves 200% throughput increase compared to the reliable BPC and maintains high throughput. Therefore the adaptive erasure code is suitable for high-level reliable file delivery in exploration missions.

Systematic fountain code will be designed in the future work, in order to improve flexibility of encoding and decoding schemes.

REFERENCES

[1] S. Shambayati, F. Bavarian, and D. Morabito. Link design and planning for Mars Reconnaissance Orbiter (MRO) Ka-band (32 GHz) telecom demonstration [C]. IEEE Aerospace Conference, May 2005, pp. 1559-1569.

[2] S. Shambayati. Deep space Ka-band link: design, continuity and completeness [C]. IEEE Aerospace Conference, March 2008, pp. 1-14.

[3] I. Sung and J. Gao. CFDP performance over weather-dependent Ka-band channel [C]. SpaceOps 2006 Conference, Rome, Italy, June 2006, pp. 1-21.

[4] B. Maruddan, A. Kurniawan, S. Sugihartono, and A. Munir. Performance evaluation of Ka-Band satellite communication system in rain fading channel at tropical area [C]. International Conference on Electrical Engineering and Informatics, July 2011, pp. 1-5.

[5] M. Felice, C. Candilio, H. Xiong, Q. Gao, and R. Ricardo. Long-term rain attenuation prediction for Ka-Band link over Venezuela [C]. 2010 5th International Conference on Pervasive Computing and Applications (ICPCA), 2010, pp. 46-49.

[6] G. Noreen, T. Komarek, and R. Diehl. Mars Telecommunications Orbiter Ka-band system design and operations [C]. 9th Ka-band Utilization Conference, Lacco Ameno, Italy, November 2003.

[7] J.Sun, J. Gao, and S. Shambayati. Ka-band link optimization with rate adaptation [C]. IEEE Aerospace Conference, March 2006, pp. 1-7.

[8] T. Cola, H. Ernst, and M. Marchese. Application of long erasure codes and ARQ schemes for achieving high data transfer performance over long delay networks [J]. Satellite Communications and Navigation, May 2008, pp. 643-656

[9] M. Luby. LT codes [C]. 43rd Annual IEEE Symposium on Foundations of Computer Science. Vancouver, BC, Canada, November 16-19, 2002.

[10] A. Shokrollahi. Raptor codes [J]. IEEE Transactions on Information Theory, vol. 52, no. 6, June 2006, pp. 2551-2567.

[11] Q. Li, L. Yin, and J. Lu. Application of LDPC codes for Deep Space Communication under Solar Scintillation Condition. ICOSSSE'11 Proceedings of the 10th WSEAS international conference on System science and simulation in engineering, 2011.

[12] J. Jiao, Q. Guo, and Q. Zhang. Packets Interleaving CCSDS File Delivery Protocol in Deep Space Communication [J]. IEEE Aerospace Electronic System Magazine, vol.26, no.2, 2011, pp. 5-11.

[13] F. Flentge. Study on CFDP and DTN Architectures for ESA Space Missions [C]. The Third International Conference on Advances in Satellite and Space Communications, 2011, pp. 75-80.

[14] J. Jiao, Q. Zhang, and H. Li. An Optimal ARQ Timer Design of file delivery time in CFDP NAK Model [C]. Wicom09, Beijing, 2009, pp. 3946-3950.

A LT-Based Distributed Codes Over Erasure Channel

Liting Wei, Jian Jiao, Zhihua Yang, Shushi Gu, Qinyu Zhang, Youhua Fan

Harbin Institute of Technology Shenzhen Graduate School

Shenzhen, China

weiliting2005001@163.com, j.jiao1983@gmail.com, yangzhihua@hitsz.edu.cn, guss_2301@163.com, zqy@hit.edu.cn, yhfan@hit.edu.cn

Abstract—**For the future space network scenarios, where multi-rovers can return science data through one obiter simultaneously, we design a new distributed LT (NDLT) code between two sources, one relay and a single sink, which could enhance the efficiency and reliability of packets transferring from source to sink. This paper proposes a method to decompose weaken robust soliton distribution (WRSD) exploited by relay node into two deconvolved weaken robust soliton distributions (DWSDs), which are used in the sources and have lower average degree than traditional RSD-based deconvolved soliton distributions (DSDs). Low operation complexity at sources and only including a simple XOR operation at relay node could be provided by the proposed degree distribution. This paper provides analytical results about the coding complexity of proposed and traditional distributed LT (DLT) code. The simulation results show that the proposed NDLT has a higher decoding probability than that of traditional DLT, under the conditions of recovering a certain proportion of original information, such as some space mission scenarios with specific data completeness requirement.**

Keywords- Distributed coding; Relay; LT code; WRSD

I. INTRODUCTION

An erasure channel in [1] is a communication channel model wherein errors are described as erasures events. In the packetized data transmission, such as scientific data return in space communication, packets may be deleted if the destination fails to recover the packets. In order to recover the packets lost in file delivery, a kind of forward erasure coding is proposed, i.e., fountain codes [1]. Luby transform (LT) code is the first realization of fountain code [2], where the input symbols are been encoded according to RSD (robust soliton distribution). Raptor code in [3], as a kind of improved LT code, consists of pre-coding and LT-coding. Raptor codes have better decoding efficiency than LT codes by using pre-coding. The intermediate symbols are the symbols generated by pre-code from the input symbols. The output symbols are the symbols by LT-code from the intermediate symbols.

In space explorations, several scenarios are common, where probers on the explored planet are much more than the obiter or relay satellites. To provide the simultaneous bulk data return capacity for multi-sources through limited relay satellites and enhance the efficiency of the relay satellites, in this paper, we consider such a scenario, where two sources transmit information to a sink through a relay, as shown in Fig. 1.

In [4, 5], the method to decompose LT codes into DLT codes based on the deconvolution of RSD is proposed. The constructed DLT has large redundancy, coding complexity and large decode failure probability. Density evolution is used in [6] to find optimal codes over the network in which multiple sources transmit information to a sink through a relay. The work in [7] introduces Soliton-like rateless coding, which exploits the benefits of fountain coding and network coding over a Y-network. This method has larger operation complexity than that in [5, 6]. In space communication scenarios with multiple sources, single relay and single sink node, a new distributed LT code is designed for enhancing the efficiency and reliability of packets transferring from source to sink. In this network of Fig. 1, as a first step, the symbols at each of the two sources are encoded using the DWSD as the degree distribution. Then, the relay selectively XORs the bit streams it receives from each source and transmits the resulted NDLT code which approximatively follows WRSD to sink node. In this paper, we propose a method to deconvolve the WRSD for yielding DWSD, by which less coding packets is consumed to cover original packets with the maximum degree of $D+1$. In this paper, the belief propagation (BP) decoding algorithm is adopted in the sink node. The proposed method could obtain the DWSD-based NDLT code by constructing a special function, which has lower coding complexity and lower decode failure probability compared with DLT under the condition of only recovering a certain proportion of original packets. We provide analytical results about coding complexity of proposed and traditional distributed LT code.

This paper is organized as follows. Section II gives a review on WRSD and proposes a specific method to deconvolve the WRSD into DWSD. Section III provides analytical results about the coding complexity of distributed LT code. The simulation results and discussion about overhead and decode failure probability are shown in Section IV. Section V presents a conclusion of this paper.

II. DECOMPOSING WRSD INTO DWSD

A. WRSD

WRSD is applicable to the LT-coding in Raptor codes [3]. In the decoding process, it could recover a certain proportion of original packets and the rest packets are recovered by pre-coding. WRSD has the lower average degree than RSD if they have the same number of original symbols. WRSD has lower decode failure probability than

RSD under the condition of recovering a certain proportion of original information.

Definition 1: For constants $\varepsilon > 0$, $D = \left[4(1+\varepsilon)/\varepsilon\right]$ and $S = (\varepsilon/2) + (\varepsilon/2)^2$, the weaken robust soliton distribution (WRSD) $R(\cdot)$ is given by

$$R(i) = \begin{cases} \dfrac{S}{S+1}, & \text{for } i = 1 \\ \dfrac{1}{(i-1)i(S+1)}, & \text{for } 2 \le i \le D \\ \dfrac{1}{D(S+1)}, & \text{for } i = D+1 \\ 0, & \text{for } D+2 \le i \le k \end{cases} \quad (1)$$

WRSD has above advantages because of its degree distribution which has largest degree-$D+1$ and two obvious spiky values, as shown in Fig. 2. It is apparent to see that the first spiky value is in degree-2, which could enhance the decode success probability in BP decoding process. The other spiky value is in largest degree-$D+1$, which could consume low coding packets to cover original packets. Thus, this could reduce the decode failure probability.

B. Deconvolution of the WRSD
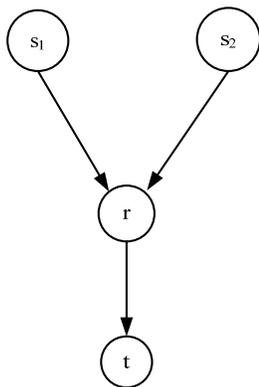
Figure 1. A two-source single-sink relay network

We consider that, as shown in Fig. 1, two sources s_1 and s_2 transfer packets to the same sink t through the relay r. Each source has $k/2$ input symbols to be transferred to the same sink. In this scenario, the source-to-relay link is lossless and all erasures occur on the relay -to -sink link. We define X_1 as a code symbol generated at s_1 with a degree d_1 and X_2 is a code symbol generated at s_2 with a degree d_2. Both d_1 and d_2 have the same degree distribution. Symbols from the two sources encoded in the relay are the same as [5]. We expect that the degree of $X_1 \oplus X_2$ is a random variable with the degree distribution of WRSD $R(\cdot)$, which WRSD has the lower average degree decode failure probability than RSD under the condition of recovering a certain proportion of original information. We define this problem as follows.

$$(p * p)(\cdot) = R(\cdot), \quad (2)$$

where d_1 and d_2 both obey $p(\cdot)$ and $d_1 + d_2$ obey $R(\cdot)$. We could obtain $p(\cdot)$ by deconvolving WRSD directly.

Direct deconvolution of the WRSD $R(\cdot)$ in (2), however, does not necessarily yield a valid probability distribution similar with [5]. To avoid direct deconvolution, we attempt to split the WRSD $R(\cdot)$ into two distributions $R1(\cdot)$ and $R2(\cdot)$. $R2(\cdot)$ captures the problematic part of the WRSD in [5](i.e., the degree-one symbols and the spike at $i = D+1$) and $R1(\cdot)$ is a smooth distribution that is easier to deconvolve .

Then we define $R1(\cdot)$ as follows:

$$R1(i) = \begin{cases} 0, & \text{for } i = 1 \\ \dfrac{1}{(i-1)i(S+1)b_1}, & \text{for } 2 \le i \le D+1 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

with the normalization factor $b_1 = \sum_{i=2}^{D+1} \dfrac{1}{(i-1)i(S+1)}$.

Similarly, $R2(\cdot)$ is given by

$$R2(i) = \begin{cases} \dfrac{S}{(S+1)b_2}, & \text{for } i = 1 \\ \dfrac{1}{(D+1)(S+1)b_2}, & \text{for } i = D+1 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

with the normalization factor $b_2 = \dfrac{S}{(S+1)} + \dfrac{1}{(D+1)(S+1)}$.

Thus, $b_1 + b_2 = 1$, and the WRSD can be rewritten as

$$R(i) = b_1 \cdot R1(i) + b_2 \cdot R2(i) = b_1 \cdot R1(i) + (1-b_1) \cdot R2(i) \quad (5)$$

So the WRSD is a mixture of the distributions $R1(\cdot)$ and $R2(\cdot)$ with mixing parameter b_1.

The approach taken in this paper is to deconvolve the distribution $R1(\cdot)$ and use the result in the construction of the new DLT codes.

$$(f * f)(i) = R1(i), \text{ for } 2 \le i \le \frac{D+1}{2} + 1 \quad (6)$$

Direct deconvolution of $R1(\cdot)$ in (6) yields $f(\cdot)$, whose independent variable value is from 1 to $\dfrac{D+1}{2}$. Give a degree distribution $p(i) = \lambda \cdot f(i) + (1-\lambda) \cdot R2(i)$, for $1 \le i \le \dfrac{D+1}{2}$, $f(\cdot)$ from direct deconvolution of (6) could make $p(i)$

obtain the largest degree- $\dfrac{D+1}{2}$, which could not make the degree distribution in relay to sink approximate WRSD using the low operation complexity protocol in the relay. Without improving obviously the operations complexity at the relay, we hope the probability of choosing the largest degree $D+1$ in the relay is larger. We define $\hat{f}(\cdot)$ recursively by

$$\hat{f}(i) = \begin{cases} \sqrt{R1(2)}, & \text{for } i=1 \\ \dfrac{R1(i+1) - \sum_{j=2}^{i-1} \hat{f}(j)\hat{f}(i+1-j)}{2\hat{f}(1)}, & \text{for } 2 \le i \le D+1 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

We can get the relationship between $f(\cdot)$ and $\hat{f}(\cdot)$.

$$\hat{f}(i) = \begin{cases} f(i), & \text{for } 1 \le i \le \dfrac{D+1}{2} \\ \dfrac{R1(i+1) - \sum_{j=2}^{i-1} \hat{f}(j)f(i+1-j)}{2f(1)}, & \text{for } \dfrac{D+1}{2} \le i \le D+1 \end{cases} \quad (8)$$

Let $\hat{R1}(i) = \left(\hat{f} * \hat{f}\right)(i)$.

We now define a new distribution constructed by $\hat{f}(\cdot)$ and $R2(\cdot)$.

Definition 2: The deconvolved weak robust soliton distribution (DWSD) $\hat{p}(\cdot)$ is given by

$$\hat{p}(i) = \lambda \cdot \hat{f}(i) + (1-\lambda) \cdot R2(i), \text{ for } 1 \le i \le D+1, \quad (9)$$

where $\lambda = \sqrt{b_1}$.

According to definition of probability distribution, we must proof $\hat{p} = \sum_{i=1}^{D+1} \hat{p}(i) \approx 1$. In order to verify $\hat{p} = \sum_{i=1}^{D+1} \hat{p}(i) \approx 1$, we need to proof $\hat{f} = \sum_{i=1}^{D+1} \hat{f}(i) \approx 1$. And in order to verify the result of new deconvolution $\hat{f}(\cdot)$ is reasonable, we hope to get $\left(\hat{f} * \hat{f}\right)(i) = \hat{R1}(i) \approx R1(i), \text{ for } 1 \le i \le k$.

*Proposition*1:
$\hat{f} = \sum_{i=1}^{D+1} \hat{f}(i) \approx 1$ and $\hat{R1}(i) \approx R1(i), \text{ for } 1 \le i \le k$.
Proof:

Similar with [5], we have $\lim_{D \to \infty} \sum_{i=1}^{\frac{D+1}{2}} f(i) = 1$.

Let $f = \sum_{i=1}^{\frac{D+1}{2}} f(i)$ and $\hat{f} = \sum_{i=1}^{D+1} \hat{f}(i)$.
We have

$$\Delta \hat{f} = \hat{f} - f = \sum_{i=\frac{D+1}{2}+1}^{D+1} \hat{f}(i)$$
$$\le \sum_{i=\frac{D+1}{2}+1}^{D+1} R1(i) = \frac{1}{D} \quad (10)$$

with D increasing，$\Delta \hat{f}$ tends to a arbitrarily small value ε_1 .
Consequently,

$$\hat{f} = \sum_{i=1}^{D+1} \hat{f}(i) \approx 1 \quad (11)$$

According to [5], we have

$$\sum_{i=1}^{2D+2} \hat{R1}(i) = \sum_{i=1}^{2D+2} \left(\hat{f} * \hat{f}(i)\right) = \left(\sum_{i=1}^{D+1} \hat{f}(i)\right)^2 \approx 1. \quad (12)$$

We can easily get $\hat{R1}(i) = R1(i), \text{ for } 2 \le i \le \dfrac{D+1}{2}+1$, and add $R1(i)$ from 2 to $\dfrac{D+1}{2}$ as follows.

$$\sum_{i=2}^{\frac{D+1}{2}} R1(i) = \frac{D+1}{D} \cdot \frac{D+1}{D+3} \quad (13)$$

with D increasing，$\sum_{i=2}^{\frac{D+1}{2}} R1(i)$ tends to 1.

Obviously, $\hat{R1}(i) = R1(i) = 0, \text{ for } 2D+3 \le i \le k$.

And for $\dfrac{D+1}{2}+2 \le i \le D+2$, we give a maximum limit which has the most number of items and each takes the biggest probability value as follows.

$$\Delta \hat{R1}(i) = \hat{R1}(i) - R1(i)$$
$$\le 2f(1)f(\hat{D}+1) + \cdots + 2f(1)f\left(\frac{\hat{D+1}}{2}+1\right)$$
$$\le 2 \cdot \frac{D+1}{2} f\left(\frac{\hat{D+1}}{2}+1\right) \le (D+1) \cdot R1\left(\frac{D+1}{2}+2\right) \quad (14)$$
$$= \frac{4(D+1)^2}{D(D+3)(D+5)}$$

with D increasing，$\Delta \hat{R1}(i)$ tends to a arbitrarily small value ε_1 .

Consequently, $\hat{R1}(i) \approx R1(i), \text{ for } \dfrac{D+1}{2}+2 \le i \le D+2$
Similarly, for $D+2 \le i \le 2D+2$

$$\Delta \hat{R1}(i) = \hat{R1}(i)$$

$$\leq 2f(1)f(\hat{D}+1) + \cdots + 2f(1)f\left(\frac{\hat{D}+1}{2}+1\right)$$

$$\leq 2 \cdot \frac{D+1}{2}f\left(\frac{\hat{D}+1}{2}+1\right) \leq (D+1) \cdot R1\left(\frac{D+1}{2}+2\right) \quad (15)$$

$$= \frac{4(D+1)^2}{D(D+3)(D+5)}$$

When D increases, $\hat{R1}(i)$ tends to a arbitrarily small value ε_1.

Consequently, $\hat{R1}(i) \approx R1(i)$, for $D+2 \leq i \leq 2D+2$

In summary,

$$\hat{R1}(i) \approx R1(i), \quad \text{for } 1 \leq i \leq k \quad (16)$$
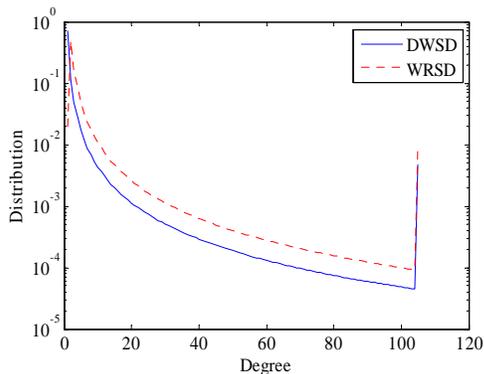
Figure 2. The WRSD $R(\cdot)$ and the DWSD $p(\cdot)$ both with ε=0.04.

We simulate $\hat{p}(\cdot)$ in Fig. 2 and find that DWSD is the same as WRSD. It is noted that, the largest probability value of DWSD is about of 71% in degree-one, while the largest probability value of WRSD is in degree-two. However, DWSD also has a spiky value in degree-D+1, and is the same to WRSD. Thus, it has similar properties with WRSD.

C. DWSD Applied to a Two-Source Single-Sink Relay Network

Similar with [5], the two sources encode symbols following the DWSD, which can be used to encode information in the network of Fig. 1. We define a sequence of code symbols produced in this process as a NDLT-2 code, and the sequence of symbols transmitted by the relay as a NMLT-2 code. It is necessary to construct a randomized decision protocol in which the relay transmits a symbol whose degree approximately follows the WRSD to sink. This approach is applied in fountain network coding in [7-10].
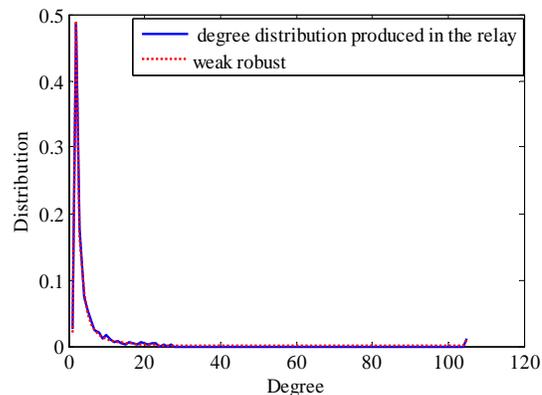
Figure 3. Comparison of WRSD and the degree distribution produced in the relay with ε=0.04

Considering the Fig. 3, the probability distribution produced by the decision protocol from [5] in the relay is approximate to WRSD. The distribution approximates WRSD in the Fig. 3. They have the same soliton waveforms coinciding with the theoretical analysis. WRSD has the lower average degree than RSD if they have the same number of original symbols. WRSD has lower decode failure probability than RSD in the condition of recovering a certain proportion of original information.

III. THE CODING COMPLEXITY OF DISTRIBUTED LT CODE

Definition 3: For constants $c > 0$ and $\delta \in [0,1]$, the robust soliton distribution (RSD) $\mu(\cdot)$ is given by

$$\mu(\cdot) = \frac{\rho(i) + \tau(i)}{\beta}, \quad \text{for } 1 \leq i \leq k \quad (17)$$

where $\beta = \sum_{i=1}^{k}(\rho(i) + \tau(i))$.

And, $\rho(i)$ and $\tau(i)$ are given by

$$\rho(i) = \begin{cases} 1/k, & \text{for } i=1 \\ 1/(i+i^2), & \text{for } i = 2 \cdots K \end{cases} \quad (18)$$

$$\tau(i) = \begin{cases} S/ik, & \text{for } 1 \leq i \leq \frac{k}{S}-1 \\ S\ln(S/\delta)/k, & \text{for } i = \frac{k}{S} \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

The RSD $\mu(\cdot)$ could be split into two distributions $\mu'(i)$ and $\mu''(i)$:

$$\mu'(\cdot) = \begin{cases} 0, & \text{for } i = 1 \\ \dfrac{\rho(i) + \tau(i)}{\beta'}, & \text{for } i = \dfrac{k}{S} \\ \dfrac{\rho(i)}{\beta'}, & \text{otherwise} \end{cases} \quad (20)$$

where $\beta' = \sum_{i=2}^{k} \rho(i) + \sum_{i=2}^{k/S-1} \tau(i)$;

$$\mu''(\cdot) = \begin{cases} \dfrac{\rho(i) + \tau(1)}{\beta''}, & \text{for } i = 1 \\ \dfrac{\tau(k/S)}{\beta''}, & \text{for } i = \dfrac{k}{S} \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where $\beta'' = \rho(1) + \tau(1) + \tau(\dfrac{k}{S})$.

A. Analysis of the Coding Complexity

In some specified scenarios with requirement of low complexity, such as deep space communication, overhead and decode failure probability and encoding complexity are important factors to measure the encoding algorithm. The algorithm of coding complexity is the same as [3].

We derive the expression of coding complexity as follows: assuming that the number of encoding symbols is $n = (1 + \varepsilon) \cdot k$; and defining

$$H(i) = \sum_{i=2}^{i-1} f(j) f(i+1-j), \quad \hat{G}(i) = \sum_{j=2}^{i} i \cdot \hat{H}(i) \quad (22)$$

1) Encoding complexity of traditional distributed LT

Similar with methods in [3], the average degree of DSD in the source is

$$\sum_{i=1}^{k/2} i \, \hat{p}(i). \quad (23)$$

The average degree of one and two degree distribution in the relay is

$$2 \cdot \frac{\beta'}{\beta} + 1 \cdot \left(1 - \frac{\beta'}{\beta}\right). \quad (24)$$

Thus, encoding complexity of distributed LT is

$$2n \cdot \sum_{i=1}^{k/2} i \, \hat{p}(i) + n \cdot \left(2 \cdot \frac{\beta'}{\beta} + 1 \cdot \left(1 - \frac{\beta'}{\beta}\right)\right)$$

$$= 2n \cdot \left(\sqrt{\frac{\beta'}{\beta}} f(1) + \left(1 - \sqrt{\frac{\beta'}{\beta}}\right) \mu^*(1) + \frac{k}{S}\left(\sqrt{\frac{\beta'}{\beta}} \frac{\left(\mu'\left(\frac{k}{S}+1\right) - H\left(\frac{k}{S}\right)\right)}{2f(1)}\right)\right. \quad (25)$$

$$\left. + \left(1 - \sqrt{\frac{\beta'}{\beta}}\right) \frac{\tau\left(\frac{k}{S}\right)}{\beta^*} + \sum_{\substack{i=2 \\ i \neq k/S}}^{k/2} i \sqrt{\frac{\beta'}{\beta}} \frac{\mu'(i+1) - H(i)}{2f(1)}\right) + n\left(1 + \frac{\beta'}{\beta}\right)$$

2) Encoding complexity of new distributed LT

The average degree of DWSD in the source is

$$\sum_{i=1}^{D+1} i p(i). \quad (26)$$

The average degree of one and two degree distribution in the relay is

$$2 b_1 + 1 \cdot (1 - b_1). \quad (27)$$

Thus, encoding complexity of new distributed LT is

$$2n \cdot \sum_{i=1}^{D+1} i p(i) + n\left(2 b_1 + 1 \cdot (1 - b_1)\right)$$

$$= n \cdot \left(\left(2\sqrt{b_1} \cdot R1(2) + \frac{2\left(1 - \sqrt{b_1}\right)}{b_2} - \frac{\sqrt{b_1} \cdot H(\hat{D} - 1)}{f(1)}\right)\right. \quad (28)$$

$$\left. + \left(2 + 2 b_1 + \frac{\ln(D+1) + C}{f(1)(S+1)\sqrt{b_1}} + \frac{\sqrt{b_1} \cdot \hat{G}(D)}{f(1)}\right)\right)$$

B. Results about the Coding Complexity

We assume that the number of encoding symbols is k=500, 800, 1000; and DSD with constants $c = 0.05$, $\delta = 0.5$ and DWSD with $\varepsilon = 0.04$.

TABLE I. DWSD AND DSD

k	k=500		k=800		k=1000	
d	DWSD	DSD	DWSD	DSD	DWSD	DSD
1	0.7101	0.6826	0.7101	0.684	0.7101	0.6848
2	0.1167	0.1142	0.1167	0.1144	0.1167	0.1145
3	0.0486	0.483	0.0486	0.0483	0.0486	0.0483
4	0.0269	0.0271	0.0269	0.027	0.0269	0.0027
5	0.0172	0.0175	0.0172	0.0175	0.0172	0.0174
64	0.0002	0.0193	0.0002		0.0002	
76				0.0182		
84						0.0177
105	0.0047		0.0047		0.0047	
$a1$	3.1102	4.4985	3.1102	4.8943	3.1102	5.0934

Table I shows the probability value from degree one to five and max degree of DWSD and DSD. The average degree $a1$ is given in table I. The average degree of DWSD is far less than DSD. Coding complexity at source of NDLT codes is $3.1102 \times n$, and the complexity of DLT codes is $4.4985 \times n$.

TABLE II. DEGREE DISTRIBUTION IN THE RELAY

k	k=500		k=800		k=1000	
d	NDLT	DLT	NDLT	DLT	NDLT	DLT
1	0.0293	0.0536	0.0293	0.9511	0.0293	0.9533
2	0.9707	0.9464	0.9707	0.0489	0.9707	0.0467
$a2$	1.9707	1.9464	1.9707	1.9511	1.9707	1.9533

Table II shows the probability value from degree one to two in the relay of new and traditional DLT codes. The average degree $a2$ is given in the end. But the average degree of distribution in the relay from DWSD is slightly larger than DSD.

Combining complexity of coding at source with relay, the total encoding complexity of new distributed LT is

$$3.1102 \times 2n + 1.9707 \times n = 4.2497 \times n . \tag{29}$$

Meanwhile encoding complexity of traditional distributed LT is

$$4.4985 \times 2n + \cdot 1.9464 \times n = 7.0506 \times n . \tag{30}$$

It can be simply seen that, encoding complexity of new distributed LT is less than traditional distributed LT with the same overhead. In the practical application scenarios, such as the deep space communication, this communication requests low complexity which can save the consumption of energy. Thus, it can enhance efficiency of information transferring from detectors on the objective planet to the earth station.

IV. SIMULATION AND DISCUSSION

In the simulation, two sources encode symbols by DWSD or DSD, and the sink node decodes by belief propagation decoding algorithm. If it dose not recover all the original symbols or a certain proportion of original symbols, we will call this decode failure. Overhead is the ration between the number of extra symbols required for decoding to succeed and k. The formula of calculating overhead is $(K-k)/k$. We assume that the number of original symbols is $k=500,\ 800,\cdot 1000$, DLT with constants $c = 0.05$, $\delta = 0.5$ and NDLT with $\varepsilon = 0.04$.

(a) k=500

(b) k=800

(c) k=1000
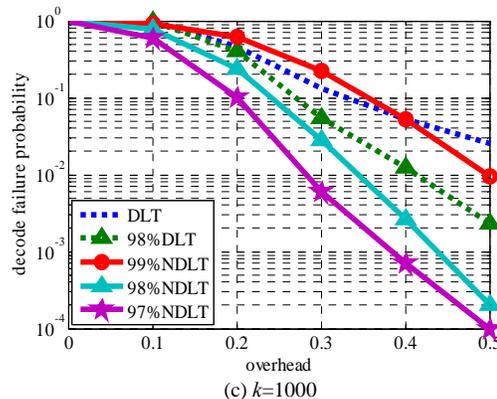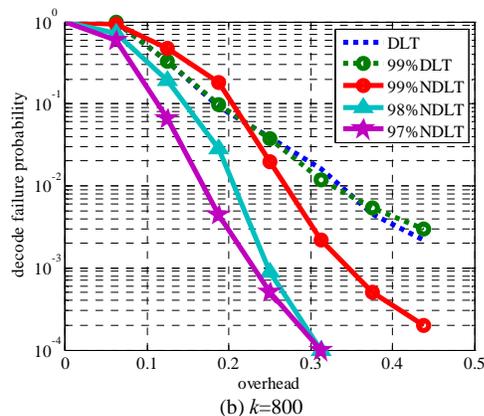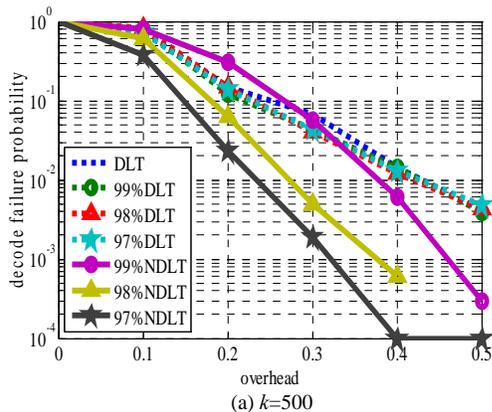
Figure 4. Overhead and decoding failure probability of distributed LT and new distributed LT of recovering 99%, 98%, 97% of original information.

Decode failure probability is shown in Fig. 4. with total coding overhead. It is indicated that distributed LT can recover all the original symbols, but its decode failure probability is larger than the NDLT codes under the condition of recovering a certain proportion of the original symbols. The NDLT codes only recover a certain proportion of the original symbols, but the rest proportion of the information could be recoverd by traditonal code in [3]. The results show if the new distributed LT code recover the quatity of original information below 98%, decode failure probability is far less than traditional distributed LT. It ensures the efficiency and reliability of the information transmission in some degree. But when the NDLT recovers more than 99% of total original data packets, decode failure probability is more than traditional distributed LT. Thus, we should choose proper traditional code, which can be more effective to recover all the original symbols. Moreover, it also show that the NDLT code has lower decode failure probability than the traditional DLT code on the condition of recovering the same proportion of original information. Thus, the NDLT code could be applicable to several specified scenatios for recovery on a certain propotion of original packets.

V. CONCLUSION

In this paper, we proposed an efficient method of deconvolving degree distribution of decentralized LT code and designed a specified distributed LT codes about the network model with communication between two sources and single sink through the relay. The proposed NDLT has lower coding complexity than traditional DLT. The simulation results indicated that coding complexity of NDLT is approximate one half of DLT. The NDLT has lower decode failure probability than traditional DLT under the condition of recovering a certain proportion of original packets. The decode failure probability is far less than DLT when the NDLT recover below 98% original information. The proposed NDLT is applicable to several scenarios to recover a certain proportion of original packets, such as file delivery in space communication.

REFERENCES

[1] D. J. C. Mackay, "Fountain codes," IEEE Proceedings on Communications, vol. 152, no. 6, May. 2005, pp. 1062-1068.

[2] M. Luby, "LT codes," 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), Vancouver, BC, Canada, Nov. 2002, pp. 271-282.

[3] A. Shokrollahi, "Raptor Codes," IEEE Trans. Info. Theory, vol. 52, No.6, June 2006, pp. 2551-2567.

[4] S. Puducheri, J. Kliewer, and T. E. Fuja, "The Design and Performance of Distributed LT Codes", in IEEE Trans. Info. Theory, vol. 53, no. 10, Oct. 2007, pp. 3740-3754.

[5] S. Puducheri, J. Kliewer, and T. E. Fuja, "Distributed LT codes", in Proc. IEEE Int. Symp. Information Theory, Seattle, WA, Jul. 2006, pp. 987-991.

[6] D. Sejdinovic, R. Piechocki, and A. Doufexi, "AND-OR tree analysis of distributed LT codes," in Proc. ITW, 2009, pp. 261-265.

[7] A. LIau, S. Yousefi, and I. –M. Kim, " Binary Soliton-like rateless coding for the Y-nerwork," IEEE Trans. On Commun, Jan. 2011, pp. 3217-3222.

[8] A. Liau, S. Yousefi, and Kim Il-Min, "Soliton-like network coding for a single relay", Information Theory (CWIT), 2011 12th Cannadian workshop on, May. 2011, pp. 86-89.

[9] S.R. W. Yeung and N. Cai, "Linear network coding," IEEE Trans. Inf. Theory, vol. 49, no. 2, Feb. 2003, pp. 371-381.

[10] A. Limmanee and W. Henkel, "A cooperative scheme for shaping degree distribution of LT-coded symbols in network coding multicast," in International ITG Conference on Source and Channel Coding, 2010, pp. 1-6.

A CFDP-based Multi-state Model on Ka-band Weather-dependent Channel

Li Huiyuan, Yang Zhihua, Jiao Jian, Zhang Qinyu

Department of Electronics & Information Engineering, Harbin Institute of Technology Shenzhen Graduate School,
Shenzhen 518055, China
lihuiyuan737@163.com, yangzhihua@hitsz.edu.cn, hellojesus@sina.com, zqy@hitsz.edu.cn

Abstract——In this paper, a channel model on transport layer over Ka-band links is proposed, which apparently reflects the characteristics of rain attenuation and equivalent noise temperature from the earth's atmosphere. The proposed model is built on a four-state Markov chain, which considered the different features of various packet types during CCSDS File Delivery Protocol (CFDP) transactions. We provide the exact criterion on how to divide different weather conditions at each state, and give the analytical results on transition probability matrix. In particular, we take Guangzhou in China as an example to find the probability distribution function of rain attenuation, and thus the transition matrix between different states, based on ITU-R rain attenuation model. The simulation results show that, compared with two state channel model, the proposed model could efficiently describe the error characteristics of different packet types during the file delivery, which leads to obvious degradations on CFDP performance.

Keywords-rain attenuation; equivalent noise temperature; ka-band; CFDP; Markov process

I. Introduction

Ka-band with largely available frequency bandwidth will greatly increase the capabilities and capacity of the future space networks. Although Ka-band links can bring obvious gains for communication system, it is highly vulnerable to fluctuating weather condition, such as the rainfall. Rainfall not only attenuates the signal power level of the signal at the receiver, but also makes the system noise temperature increased. These two aspects further worsen the signal-to-noise ratio (SNR) and make BERs of the system over long distance increased. On the other hand, high BERs will degrade the reliable delivery on the scientific data, due to frequent retransmissions on the missing data packets.

For file-based scientific data return, CCSDS puts forward CCSDS File Delivery Protocol (CFDP), instead of its terrestrial counterpart (File Transfer Protocol, FTP), which is designed to provide file-based data storage management, store-and-forward relay, and reliable data transfer over space links characterized by large propagation delay and intermittent availability[1]. The CFDP employs Automatic Repeat reQuest (ARQ) retransmission mechanism to file reliable delivery, which is significantly affected by the earth's weather over Ka-band links. In [2], the latency of CFDP over Ka-band is analyzed theoretically, based on a two-state Markov chain channel. This two-state model, denoted by Gilbert-Elliot channel model [3], divides simply the different weather conditions over Ka-band link into two states of "good" and "bad" respectively. Especially, it

exploits a large number of atmospheric noise temperature data sampled by DSNs of NASA, from which a threshold is determined to separate two weather states. In [4], a new statistical approach for deep-space communications links design was formulated and optimal link design strategies over Ka-band has been presented by using the atmospheric noise temperature statistics and the ground antenna characteristics to optimize the average data return on the link. A wide-area environment two-state channel model including "ideal state" and "non-ideal state" was proposed in [5], model parameters under three kinds of typical channel environments were fitted by using the statistics method of experimental data. Maruddani et al. [6] evaluated the performance of Ka-band wireless communication system in tropical area by comparing several channel coding schemes.

In the CFDP, PDUs are divided into different types due to their specified roles in the file transfer, with different length. Weather impairments on different types of PDUs over Ka-band links could produce unequal performance degradation, since the PDUs with different length could be lost or received independently at the same BER. Therefore, it is necessary to take the diversified influence of different types of PDUs into account when we model the channel characteristics over Ka-band weather dependent links. In this paper, we attempt to solve this problem by constructing multi-state channel model, in which each state reflects on the individual effects of the different PDUs. This paper has two following novel contributions: a) we built a four-state Markov channel model on transport layer over Ka-band links constructed with rain attenuation and system noise temperature together, which is not considered in the original two-state Gilbert-Elliot channel model. Thus combine the physical characteristics over Ka-band links on transport layer; b) In the model given, we determined the different states of channel by considering specified functions of PDU types of the CFDP, which could exactly characterize the influence on file delivery performance over Ka-band channel.

In Section II, we describe the CFDP file transfer process, and Gilbert-Elliot channel model has been introduced. In Section III, we divide the weather state into several types based on different types of PDUs in the CFDP, and give the derivation of the theoretical formulas, then some simulations is shown to find out the weather-related parameters in this model. The simulation of the delay performance of CFDP under this model with Monte Carlo algorithm, indicating the validity of the model by comparing the channel with others, have been shown in Section IV, followed by conclusions in Section V.

II. CFDP AND GILBERT-ELLIOT CHANNEL MODEL

A. Process of CFDP deferred NAK mode transmission

CFDP has four error control modes to handle the link disruptions and outages frequently encountered in space. In this study, we used the CFDP deferred NAK mode since it is the most generally used form of CFDP for space link [1].

For the CFDP file transfer, each file is divided into multiple protocol data units (PDUs). The sender will send Metadata PDU (MPDU) first, and then send PDUs to the receiver. The end of file PDU (EOF-PDU) is sent at the end of the initial transmission attempt. The receiving entity saves all information about missing data until the EOF is received. It then issues a NAK listing all the PDUs that were not successfully received. Upon the reception of this NAK message, the sender side retransmits the PDUs and the receiver will again responds with NAK messages until all PDUs are received correctly; then, the receiver will send finished message (FIN) to the sender and close the file transmission [7], as shown in Fig. 1.
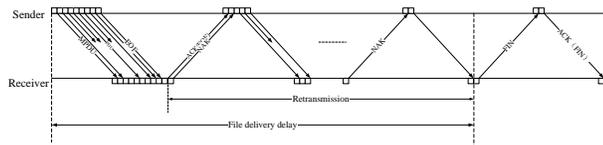
Figure 1. Process of CFDP deferred NAK mode transmission

B. Gilbert-Elliot Channel

We define the weather-dependent transport layer channel of CFDP as an erasure channel, which is generally described by Gilbert-Elliot model. There are two weather states in Gilbert-Elliot channel model, "good" and "bad" weather conditions. If it is good weather state, most of the transmitted packets will be received successfully. If not, however, most of the transmitted packets will experience errors due to bad weather condition. Therefore, two different BERs are applying to two different weather conditions. The weather state undergoes stochastic transitions, as shown in Fig. 2, and the transition from one state to another is defined by the transition matrix P.
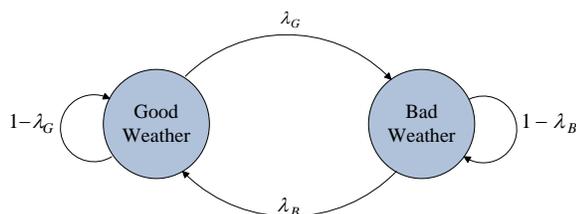
Figure 2. Gilbert-Elliot Channel model

$$P = \begin{bmatrix} P(G|G) & P(B|G) \\ P(G|B) & P(B|B) \end{bmatrix} = \begin{bmatrix} 1-\lambda_G & \lambda_G \\ \lambda_B & 1-\lambda_B \end{bmatrix} \quad (1)$$

We attempt to give the fact that, the performance of the CFDP over two-state weather channel model is worse than that of single-state AWGN channel model, through the comparison between them by Monte-Carlo simulations. The simulation parameters are shown in Table I.

TABLE I. SIMULATION PARAMETERS OF THE COMPARISON

File size	10MB	
The number of simulation	10000	
BERs in two-state weather model	Good state	1e-5, 1e-6, 1e-7, 1e-8
	Bad state	1e-3
BERs in single state model	1e-5, 1e-6, 1e-7, 1e-8	

Fig. 3 shows that the average transmission number of the CFDP over two-state weather model is slightly more than that of the single-state model, as a result of taking the bad weather into account. It is indicated that the two-state model could reflect the characteristic of the real weather dependent channel to a small extent. Also, the result of simulation with different combinations of good and bad weather BER pairs: (1e-5, 1e-3), (1e-6, 1e-3), (1e-7, 1e-3), and (1e-8, 1e-3) in two-state model, shows that different pairs produce different average retransmission numbers. It means that, the division criterion of the weather state have a nontrivial effect on the performance of the CFDP.
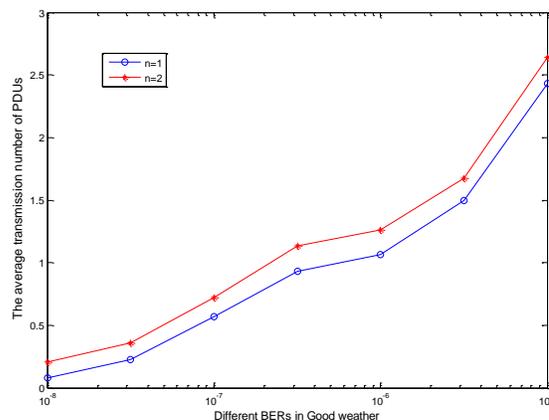
Figure 3. The average numbers of transmissions required under different channel models

Therefore, well-designed criterion of the division is considerably indispensable for the efficiency of the channel model, which could simulate the real channel properly.

III. FOUR-STATE WEATHER CHANNEL MODEL

A. The method of modeling

We determined the different states of channel by considering specified functions of PDU types of the CFDP. In the CFDP, different types of PDUs have different lengths and functions [8]. We group the PDUs into three types according to the different lengths. The first type includes

Prompt PDU, ACK PDU, and EOF PDU with the length of approximate 30Bytes; the second type contains Metadata PDU, File data PDU, and Finished PDU with the length of approximate 1024Bytes; another type of PDU is NAK PDU with the variable length, which changes as the number of the lost PDUs during the process of file transport, as shown in Table II.

TABLE II. THE TYPES OF PDUS IN THE CFDP

Types of PDUs	L_{pdu}(Byte)
EOF/ACK/PROMPT PDU	30
META/File-data/FIN PDU	1024
NAK	variable

We attempt to include respective possibility of errors occurred in different types of PDUs into the channel model, since the lengths of different types of PDUs (L_{pdu}) have obvious disparity. In this section, we design a multi-state channel model to distinguish the error rate of different types of PDUs, which could combine better with CFDP characteristics than the traditional two-state model. Our proposed weather dependent channel has four state divided by different BERs. The method of modeling is presented in the following way:

We know the relationship between packet error rates P and bit error rates:

$$P = 1 - (1 - \gamma)^{L_{pdu}} \qquad (2)$$

Thus,

$$\gamma = 1 - (1 - P)^{1/L_{pdu}} \qquad (3)$$

The error probability of the first PDU type is very small for its length is very short, and the second PDU type is the longest one, so it will lost when the environment of transmission is bad. We assume that the receive message is right when packet error rates is under 0.01, and the links break off when the packet error rates is above 0.5. So, we take 0.01 and 0.5 as two threshold values. We calculate the upper boundry BERs γ of different PDUs shown in Table III.

TABLE III. THE RELATIONSHIP BETWEEN BERS AND RAIN ATTENUATION

Types of PDU	Threshold values of packet error rates P	Maximum BERs
EOF/ACK/PROMPT PDU	0.01	4.19e-05
META/File-data/FIN PDU	0.01	1.23e-06
	0.5	8.46e-5

The proposed model contains a good state and three bad ones. Given a good weather state we predetermined, we divided three rainfall states by two BERs threshold in the case of packet error rates is 0.01, We use the BER corresponding to packet error rates is 0.5 as the upper boundary of the BERs. By doing this, we obtain a four weather states of channel model.

In this model, we assume that the transition occurs only between neighboring states since the changing of weather

conditions could conforms better to the characteristic of the first order Markov process, as shown in Fig. 4. Let T be the state transmission matrix as shown below, the channel could be represented from a state transmission matrix:
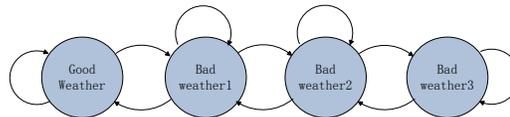
Figure 4. The Four-state weather channel model

$$T = \begin{pmatrix} P_{1,1} & P_{1,2} & 0 & 0 \\ P_{2,1} & P_{2,2} & P_{2,3} & 0 \\ 0 & P_{3,2} & P_{3,3} & P_{3,4} \\ 0 & 0 & P_{4,3} & P_{4,4} \end{pmatrix} \qquad (4)$$

B. The calculation on the proposed model parameters

Given BPSK modulation employed in this paper, we have:

$$\gamma = Q\left(\sqrt{\frac{2\varepsilon_b}{N_o}}\right) \qquad (5)$$

where ε_b is the signal energy of each bit and,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt, \qquad x \geq 0 \qquad (6)$$

So, we get:

$$\frac{2\varepsilon_b}{N_0} = (Q^{-1}(\gamma))^2 \qquad (7)$$

We used the general antenna gain calculated by follows:

$$G = \eta \frac{4\pi A}{\lambda^2} \qquad (8)$$

where λ is the wavelength of the signal, A is the receiving antenna area, and η is the antenna efficiency. For directional antenna, the power density of the destination is:

$$W = \frac{P_s G_s}{4\pi r^2} \ (W/m^2) \qquad (9)$$

where r is the distance of transmission, if the effective receiving antenna area is $A \times \eta$, the signal power received at the receiver can be expressed as:

$$P_r' = WA\eta = \frac{P_s G_s A \eta}{4\pi r^2} = \frac{P_s G_s G_r}{L_f} \qquad (10)$$

where $L_f = (4\pi r/\lambda)^2$ is the loss of free space.

Here, we consider the rain attenuation to the signal power A_p, and Pr at the receiver is:

$$[P_r]dB = [P_s + G_s + G_t - L_f - A_p]dB \qquad (11)$$

The energy of signal at the receiver is:

$$\varepsilon_b = \frac{P_r}{b} \qquad (12)$$

where b is bit rates (bit/sec) and we have $N_0=KT$, then according to (7) we have:

$$T = 2 \cdot P / Kb \cdot (Q^{-1}(\gamma))^2 \qquad (13)$$

Thus, we get the system noise temperature of earth station T, it is below as in [9]:

$$T_{sys} = \sigma_f T_{Ant} + T_{receiver}(K) \qquad (14)$$

where σ_f is the coefficient of coupling，T_{ant} is the equivalent noise temperature in receiving antenna and $T_{receiver}$ is the noise temperature of the receiver.

Having T, we can get T_{ant}:

$$T_{ant} = [T_{sys} - T_{receiver}] / \sigma_f \qquad (15)$$

By [10], we have:

$$T_{Ant} = T_m(1 - 10^{-0.1Ap}) + T_c * 10^{-0.1Ap}(K) \qquad (16)$$

where, T_c is cosmic background temperature, usually chosen as 2.7 K. T_m is atmospheric mean radiating temperature (K).

As a result, we obtain

$$A_p = 10 \lg \frac{T_c - T_m}{T_{ant} - T_m} \qquad (17)$$

Above formulas give a specific relationship between BERs and rain attenuations corresponding to different types of PDUs. We use these exact values of rain attenuation as the thresholds to divide weather conditions.

We take Guangzhou in China (typical subtropical monsoon climate) as an example in this section, in order to find the PDF of the rain attenuate by using a set of rain data (A_{Ri}, P_i), $i=1,\dots,N$, which are sampled from the ITU_R rain attenuate model[11,12]. In particular, A_{Ri} is the value of the rain attenuation being exceeded in P_i time in every statistics years, so we have:

$$P_i = prop\{A_i \ge A_{Ri}\}; \quad i = 1,\dots,N \qquad (18)$$

Fig. 5 shows the simulation results of the ITU_R rain attenuation model, and we can find the relationship between A_{Ri} and P_i.
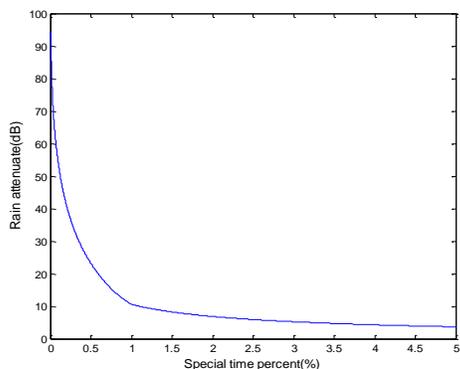
Figure 5. The relationship between rain attenuate and special time percent

Since the rain attenuation follows the logarithmic normal distribution, we can approximate the distribution of rainfall,

with its mean and square as that in [13]. The derivation is as below:

$$P = F(\ln A_R) = \int_{\ln A_R}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-m)^2}{2\sigma^2}) dx \qquad (19)$$

Similarly:

$$P = Q(\frac{\ln A_R - m}{\sigma}) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln A_R - m}{\sigma}}^{\infty} \exp(-\frac{x^2}{2}) dx \qquad (20)$$

Then, we get:

$$\frac{\ln A_R - m}{\sigma} = Q^{-1}(P) \qquad (21)$$

Since mean and square are unknown parameters, we make it as:

$$Q^{-1}(P) = C \cdot \ln A_R + B \qquad (22)$$

Where the C and B can be obtained from the linear regression of a set of data known below:

$$(\ln A_{Ri}, Q^{-1}(P_i)), \quad i = 1,\dots,N \qquad (23)$$

Thus, we can get the value of m and σ parameters as:

$$m = -\frac{B}{C}, \quad \sigma = \frac{1}{C} \qquad (24)$$

We obtained the value of m and σ parameters by a large number of simulations on linear regression, using the data from the ITU-R model, shown in Fig. 6. As a result, we get the value of mean and square separately: $m= - 1.9020$, $\sigma = 1.8950$.
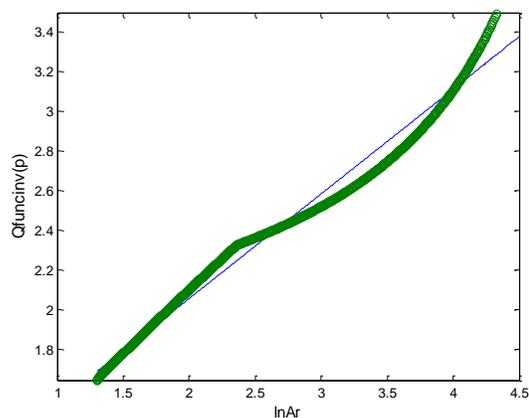
Figure 6. The linear regression of the data from the ITU-R model

Based on obtained m and σ, we represent the long-term PDF of rain attenuation as follows:

$$P(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} \exp(-\frac{(\ln x - m)^2}{2\sigma^2}) & x > 0 \\ 0 & x \le 0 \end{cases} \qquad (25)$$

The probability distribution function of rain attenuation in Guangzhou is shown in Fig. 7.
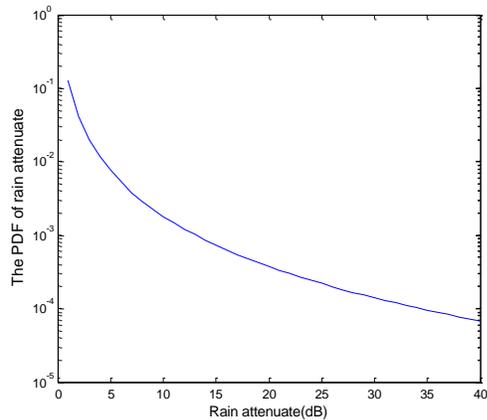
(a) The PDU of rain attenuation
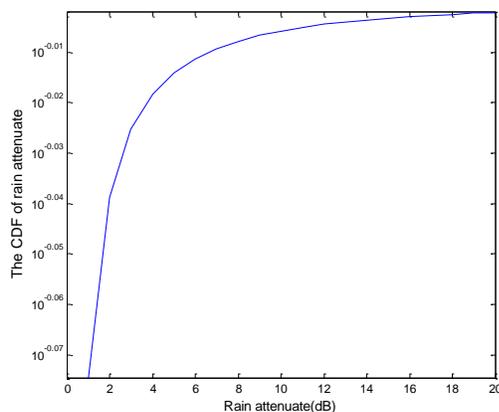
(b) The CDF of rain attenuation

Figure 7.   The probability distribution of rain attenuation in Guangzhou.

To obtain the state transmission matrix (4), we firstly need to calculate N_k, which is the number of times that attenuation crossed a given threshold l_k in single direction per second. The fade slope f_k is a measure of the time rate of change of attenuation at a given threshold l_k in units of (dB/sec) as defined in [14]. Then, we can calculate N_k approximatively by the multiplication of f_k and the value of PDF at a given threshold l_k in [15]:

$$N_k = f_k \cdot \frac{1}{l_k \sigma \sqrt{2\pi}} \exp(-\frac{(\ln l_k - m)^2}{2\sigma^2}), \quad k>0 \, \& \, k<K$$

$$N_0 = f_0, k = 0 \quad (26)$$

We also need π_k, which is state probability of the k_{th} state. We approximate the state probability of good weather state to 0.7 in Guangzhou for example, and the probability of the k_{th} state π_k, by long-term probability distribution, is:

$$\pi_k = 0.3 \cdot \int_{l_{k-1}}^{l_k} p_u du$$

$$= 0.3 \cdot \int_{l_{k-1}}^{l_k} \frac{1}{u\sigma\sqrt{2\pi}} \exp(-\frac{(\ln u - m)^2}{2\sigma^2}) du \quad (27)$$

State transition probability $t_{k,k+1}$ is approximated to:

$$t_{k,k+1} \approx N_k / \pi_k \quad (28)$$

Similarly,

$$t_{k,k-1} \approx N_{k-1} / \pi_k \quad (29)$$

Then $t_{k,k}$ is obtained by follows:

$$t_{k,k} = 1 - t_{k,k+1} - t_{k,k-1} \quad (30)$$

According to above formulas, the settings of parameters are shown in Table IV, we take these as example, and the transition matrix we achieved is in Table V.

TABLE IV.   THE SETTINGS OF PARAMETERS FOR TRANSITION MATRIX

Latitude	32.24'N	Longitude	118.46'E
Frequency(GHz)	20	Ps(w)	5
Bit rates(bit/sec)	1M	$T_{receiver}$(K)	160
Antenna efficiency	0.65	T_m(K)	270

TABLE V.   THE TRANSITION MATRIX

P	Good	Bad1	Bad2	Bad3
Good	0.9886	0.0114	0	0
Bad1	0.0269	0.973	0.0001	0
Bad2	0	0.198	0.492	0.31
Bad3	0	0	0.0604	0.9396

We use the state transmission matrix of the four-state Markov chain to imitate dynamic process of weather states.

IV.   SIMULATION AND DISCUSSION

In this section, we evaluate the performance of CFDP under the four-state weather channel model by Monte Carlo simulations.

First, CFDP is evaluated over various BERs with different sizes of files as 1MB, 5MB and 10MB separately in single-state channel. Fig. 8 shows that with the BERs increases, the average retransmission number of PDUs increase significantly. Under the same BER, the average retransmission number of PDUs increases with the size of files. In Fig. 9, the loss of NAK message occurs in the case that BER is more than 1e-5, which indicates the performance of the CFDP is seriously affected by weather conditions, especially bad weather. Therefore, weather effect must be considered in analyzing of the CFDP performance.
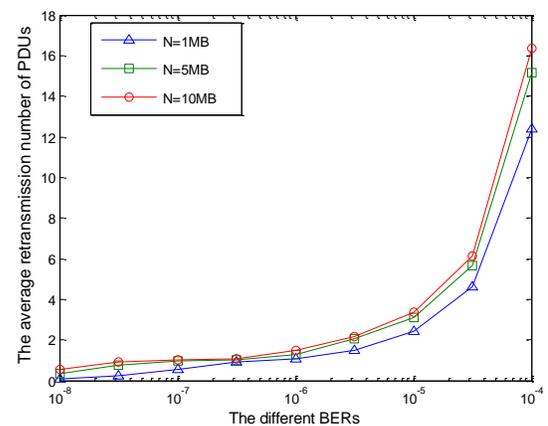
Figure 8.   The average retransmission number of PDUs
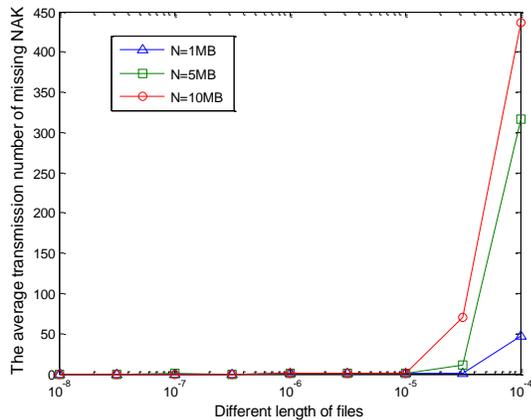
Figure 9.   The average number of missing NAK

Fig. 10 shows the comparison of the CFDP performance with different sizes of files over channel models with different weather states *n*, including *n*=1, *n*=2 and *n*=4. The settings of simulations are shown in Table VI.

TABLE VI.   THE SIMULATION PARAMETERS

File size(MB)		1, 2, 4, 6, 8, 10, 12
The times of simulation		10000
PDU size (Byte)	*long-PDU*	1024
	Short-PDU	30
	NAK	variable
The BERs of channel models	*Single-state*	1e-7
	Two-state model	(1e-7,1e-4)
	Four-state model	(1e-7,4.19e-5,1.23e-6, 1e-4)

We can find that the results of the CFDP performance under the four-state weather model are between that under the two-state model and single-state model. The results have evident decrease on the base of two-state model due to the further revision of weather conditions. It means that the channel we modeled characterizes better the dynamic process of weather conditions.
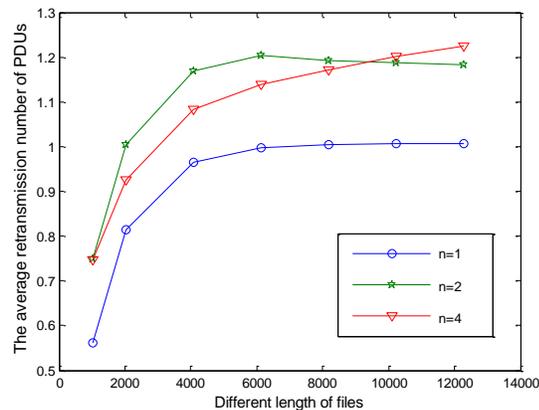
Figure 10. The average retransmission number of PDUs under different channel models

From the channel model, we can get the detailed NAK message which is vital to the whole link. Fig. 11 (a) shows the average transmission number of NAK. And the average number of missing NAK is shown in Fig. 11 (b). Similarly, The result of the four-state weather model has an evident decrease compared with the two-state model. The model could approach to the real channel characteristics more efficiently as a result of the careful revision of weather conditions change.
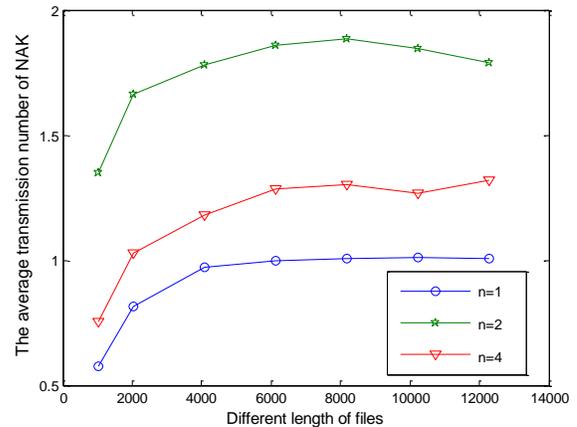
(a)The average transmission number of NAK
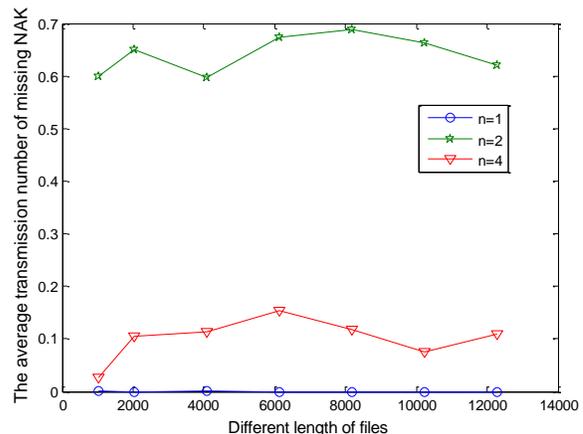
(b) The average transmission number of missing NAK

Figure 11. The performance of transmitted NAK under different channel models.

As a result, the four-state weather channel model gives a accurate description of the detailed dynamic process of weather states, and reflects more efficiently on the impact of weather conditions change on channel performance.

V.   CONCLUSION AND FUTURE WORK

In this study, we established the four-state weather channel model on transport layer of CFDP considering the propagation characteristics, including rain attenuation and equivalent noise temperature over Ka-band links. The proposed model could evaluate the performance of CFDP over Ka-band channel model efficiently, by reflecting on the impact of weather conditions into modeling channel. In

particular, we design a elaborated-designed division criterion to the weather conditions and divided the weather states by the BERs from the transport layer. The results indicated that the proposed model is more reasonable as it gives realistic results. Moreover, the four-state weather model give a accurate description of the detailed dynamic process of weather states, thus reflects more efficiently on the impact of weather conditions on channel.

Future work should take an emphasis on validation the results with actual measurements, meanwhile, using a high-order Markov chain model in order to accurately simulate weather processes.

VI. ACKNOWLEDGMENT

REFERENCES

[1] J. L. Gao and J. S. SeGui. "Performance Evaluation of the CCSDS File Delivery Protocol – Latency and Storage Requirement", IEEEAC paper #1084, Version 7, pp. 1300-1312, December, 2005.

[2] I. U. Sung and J. L. Gao. "CFDP Performance over Weather-Dependent Ka-band Channel", in: Proceeding of SpaceOps 2006, June, 2006.

[3] E. N. Gilbert. "Capacity of a burst-noise channel," Bell Syst. Tech.J., vol. 39, pp. 1253-1265, September, 1960.

[4] S. Shambayati. "Ka-Band Telemetry Operations Concept A Statistical Approach," Proc. IEEE, Vol 95, No. 11, pp. 2171-2179, November, 2007.

[5] M. Yang, F. Meng, S. Shi, and Q. Guo. "Markov Chain Based Two-State Satellite Mobile Channel Model". Vehicular Technology Conference (VTC Spring). pp. 1-5, July. 2011.

[6] B. Maruddani, A. Kurniawan, Sugihartono, and A. Munir. "Performance Evaluation of Ka-Band Satellite Communication System in Rain Fading Channel at Tropical Area", in Electrical Engineering and Informatics (ICEEI), 2011 International Conference on. pp. 1-5, September, 2011.

[7] CCSDS. CCSDS file delivery protocol (CFDP) part 1: introduction and overview. CCSDS 720.1-G-3. 2007,4.

[8] CCSDS. CCSDS file delivery protocol (CFDP) part 2: implementers guide. CCSDS 720.2-G-3. 2007,4.

[9] T Pratt, C Bostian, and J. Allnutt. Satellite Communications. Second Edition. pp. 101-105, July, 2005.

[10] F. S. Marzano and M. Montopoli. "Millimeter-wave antenna noise temperature due to rain clouds: theoretical model and statistical Prediction". European Conference on Antennas and Propagation. Proceedings, pp. 2906-2910, June, 2009.

[11] Rec. ITU-R P. 618-10, Propagation Data and Prediction Methods Required for the Design of Earth-space Telecommunication Systems[R]. 2009.

[12] Rec. ITU-R P. 838-3, Specific attenuation model for rain for use in prediction methods[R]. 2005.

[13] M. Filip and E. Vilar. "Optimum Utilization of the Channel Capacity of a Satellite Link in the Presence of Amplitude Scintillations and Rain Attenuation", IEEE Trans. Commun., VOL.38, NO.11, pp. 1958-1965, Nov. 1990.

[14] W.L. Stutzman, A. Safaai-Jazi, T. Pratt, B. Nelson, J. Laster, and H. Ajaz. "Results from the Virginia Tech Propagation Experiment Using the Olympus 12, 20, and 30 GHz Satellite Beacons", IEEE Trans. Antenn. Propagat. vol. 4354, pp. 54-62, Jan. 1995.

[15] J.-P. Ebert and A. Willig. "A Gilbert-Elliot Bit Error Model and the Efficient Use in Packet Level Simulation" TKN Technical Report TKN-99-002, Technical Univ. Berlin, Telecomm. Networks Group, Mar. 1999.

An Improved Double Retransmission Scheme to Deferred NAK CFDP

Haitao Wu[1,2], Zhihua Yang[1], Jian Jiao[1], Qinyu Zhang[1]

[1] Department of Electronics & Information Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China

[2] Faculty of Electronic Information & Mechanical Electrical Engineering, Zhaoqing University, Zhaoqing, China

wht@hitsz.edu.cn, yangzhihua@hitsz.edu.cn, hellojesus@sina.com, zqy@hitsz.edu.cn

Abstract—In this paper, we put forward an improved scheme based on deferred NAK (Negative Acknowledgement) mode in CCSDS (Consultative Committee for Space Data Systems) file delivery protocol, to make a trade-off between the file delivery time and the throughput. We employ K retransmissions on the missing packets to decrease the packet error probability and lower file delivery time. We define a ratio $\Delta(K-1)$ to evaluate the performance trade-off, which denotes the decrement of file delivery time by one extra retransmitted PDU (Protocol Data Unit) relative to standard deferred NAK for different K. We determine the proper retransmission times $K = 2$ by calculating the ratio of different K via a bulk of simulations. The simulation results indicate the performance of this scheme outperforms the standard deferred NAK mode in terms of file delivery time, especially in the scenarios with small connectivity time, long propagation delay and high Packet Error Rate (PER>0.1) greatly.

Keywords-CCSDS file delivery protocol; deferred NAK mode; double retransmission; file delivery time; power management

I. INTRODUCTION

Over the recent years, Consultative Committee for Space Data Systems (CCSDS) has standardized a set of communication protocols based on state-of-the-art techniques. Indeed, since the needs of different missions, a single type of standard would not satisfy all scenarios and possible selections will be offered to users [1].

Currently there are two kinds of reliability mechanisms for various mission scenarios, respectively based on retransmission and redundancy schemes. The ultimate frontier for ensuring reliable data transfer is represented by the joint application of ARQ and erasure codes, so as to provide a two-level protection in deep space communication[1]. Coding solutions applying the concept of digital fountain, such as LT, Tornado, and Raptor codes, yet deserve great attention and were initially considered within the CCSDS standardization process. Despite the virtues they may offer in terms of rateless coding, they still require the availability of a feedback channel for signaling the completion of the decoding procedure to the sender, which otherwise would continuously transmit new redundancy symbols, wasting power and link bandwidth[2].

As one type of reliability mechanisms based on retransmission, the CCSDS File Delivery Protocol (CFDP) offers end-to-end transport services for the file transfer to/from onboard mass memories. It is designed to get reliable transfer of files by following an FTP-like paradigm. Its implementation spans the application and transport layers. It consists of two operative procedures: core and extended. CFDP adopts an ARQ (Automatic Repeat reQuest) mechanism, while uses NAK (Negative Acknowledgement) to replace ACK (Acknowledgement). Considering the features of links and transmission requirements, CFDP offers four selectable modes: deferred, immediate, prompted and asynchronous [3].

The remainder of this paper is organized as follows. Section II presents the related work. Section III describes and analyzes K-retransmission-based deferred NAK scheme. Section IV defines a reasonable ratio to validate and evaluate our scheme, and determines the proper K value. Section V makes a comparison between standard deferred NAK and our scheme based on $K = 2$. Finally, conclusions are drawn.

II. RELATED WORK

Some work have been done to analyze the performance of CFDP. W. Baek and D. C. Lee evaluate the expected file delivery time in deferred and immediate NAK modes [4][5]. All parameter values are set under the constraint that the throughput is never compromised. T. de Cola proposes Repeated Transmission (RT) mechanism to extend the CFDP features, transmitting the same CFDP PDU for K times consecutively [6]. The relationships between packet loss rate, packet size, throughput and the number of transmissions are merely discussed under the cislunar scenario，whereas the file delivery time is not mentioned. G. Papastergiou and I. Psaras advance the Deep-Space Transport Protocol (DS-TP), which transmits each packet twice importing some delay between the original transmission and the retransmission. Although DS-TP can obtain two times faster than conventional protocols in terms of file delivery time, it induces high overhead [7].

From the above work, we find that the same PDU allowed to be delivered either two or K times consecutively during the whole period of file transfer. As a result, extraordinary power will be consumed in the context of multiple repeated transmissions. How to measure the compromise between the file delivery time, extraordinary power and throughput of CFDP, is an open issue. The purpose of this paper is to bring forward an improved scheme to standard deferred NAK CFDP, in which we use a K-retransmission-based strategy to make a trade-off between the file delivery time and the throughput. We measure the

proposed scheme under several typical scenarios, such as GEO, Earth to Moon and Earth to Mars.

III. *K*-RETRANSMISSION-BASED SCHEME OF CFDP

In CFDP, file-transfer is called a "transaction". Each file is segmented into PDUs of variable length before transmission. For the first transmission attempt, a meta-data PDU is sent at the beginning, followed by the file data PDUs sent out in sequence with an end-of-file (EOF) PDU, which marks the end of the first transmission attempt and signals the receiver to respond. If missing or error-corrupted PDUs were detected, a NAK message is issued back to the sender, with a list of PDUs that need to be retransmitted. Upon reception of the NAK, the sender will retransmit the requested PDUs until all missing PDUs have been correctly received. If all PDUs are successfully received, the receiver will send a Finished (FIN) message to the sender, signaling the completion of the transaction. As a widely used mode, in this paper we regard deferred NAK CFDP as an improved object. We consider the transport layer channel of CFDP to be an erasure channel and make some assumptions as follows, details refer to [4][5]:

- Link is full-duplex and power budget is fixed.
- Processing delay of PDU is neglected.
- The meta-data PDU and all file data PDUs have an identical length.
- Each EOF is transmitted successfully at a time.

Notations we use in this paper are specified in Table I.

TABLE I.   NOTATIONS.

Symbol	Definition
K	Retransmission times for each PDU in retransmission stage
N	Total number of PDUs in a transaction
P_{ef}	Probability of PDU error or lost
P_{er}	Probability of error in delivering NAK
T_{prop}	One-way propagation time
T_{PDU}	Transmission time of meta-data or file data PDU
RT	Transmission time of the PDUs in retransmission stage
P	Probability
H	A random variable, transmission number of a PDU in retransmission stage

A. Scheme Description

As depicted in Fig. 1, we chronologically divide the total file delivery time into two stages: first sending stage and retransmission stage. The first sending stage starts at the sender's transmission of meta-data initiating the transaction and ends after EOF reception at the receiver entity. The retransmission stage begins at the transmission immediately following ACK (EOF) transmission and finishes at the instant when all missing PDUs have been correctly received. Note that our definition of file delivery time does not include the time for FIN-ACK(FIN) procedure. Let us have a quick overview of *K*-retransmission-based deferred NAK scheme. The algorithm is described as follows:
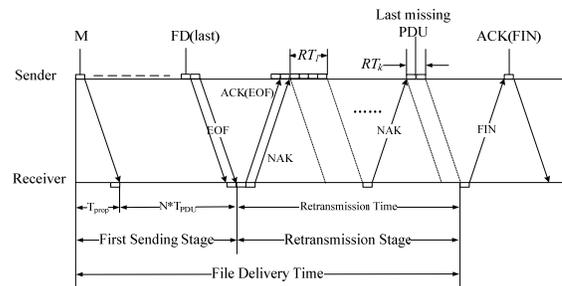
Figure 1.   Delay time diagram of *K*-retransmission-based deferred NAK.

Step I: In first sending stage, sender entity transmits a meta-PDU and all file data PDUs in sequence to receiver entity, after finishing sending all the data PDUs, the sender entity sends a EOF PDU to the receiver entity, then waits ACK(EOF) and NAK. Upon receiving the ACK(EOF) and NAK, if there is no missing PDU, the transaction is closed, details refer to [3][4], otherwise goes to Step II.

Step II: In retransmission stage, upon receiving a NAK, the sender entity immediately retransmits each missing PDU K times consecutively. If missing PDUs still remain, the receiver entity issues another NAK. This process repeats until the receiver entity receives all necessary PDUs successfully, then the transaction is closed.

Here, K is a meaningful parameter, which is counted by a counter at the sender entity and influences the overall performance and helps increase the probability of data delivery at cost of power waste (proportional with K).

B. Theoretical Analysis

Before the theoretical analysis of *K*-retransmission-based deferred NAK, we address the setting rules on EOF timer and NAK timer. Considering CFDP over a direct hop-to-hop link, in order to prevent unnecessary duplicate retransmission, the time-out value of the EOF timer should be $2T_{prop}$. The time-out value of the NAK timer set upon issuance of the NAK that causes the kth retransmission spurt in the retransmission stage, should be $2T_{prop} + RT_k$, where RT_k denotes the transmission time of the PDUs requested by the receiver for the kth retransmission spurt in the retransmission stage [4]. From the aforementioned assumptions, the EOF delivery time can be T_{prop} because it is transmitted successfully at a time.

Now let us focus on analysis of the retransmission stage, we denote random variable H_i^K the number of transmissions of the ith PDU up to and including its first successful transmission during the period of *K*-retransmission-based. Under our channel assumption, H_i^K has a geometric distribution, and H_i^K is equal or larger than 0 because some possible PDUs have been successfully delivered to the receiver during the first sending stage. The retransmission spurts will reoccur until all PDUs are

delivered to the receiver, so $\max(H_1^K, H_2^K, \cdots, H_N^K)$ is the number of retransmission spurt. Then, we define random variable H_M^K as $H_M^K = \max(H_1^K, H_2^K, \cdots, H_N^K)$.

Considering the minimum setting value of NAK timer is $2T_{prop} + RT_k$, the expected retransmission time during the first retransmission spurt can be obtained as

$$\sum_{i=1}^{\infty}\left[i(2T_{prop} + RT_1)\right]P_{er}^{i-1}(1 - P_{er}) = \frac{2T_{prop} + RT_1}{1 - P_{er}} \quad (1)$$

So the expected time during the whole retransmission stage is given as

$$E\left(\sum_{k=1}^{H_M^K} \frac{2T_{prop} + RT_k}{1 - P_{er}}\right) = \frac{E(H_M^K) \cdot 2T_{prop}}{1 - P_{er}} + \frac{E\left(\sum_{k=1}^{H_M^K} RT_k\right)}{1 - P_{er}} \quad (2)$$

For the calculation of $E(H_M^K)$, we have

$$\begin{aligned}
E(H_M^K) &= \sum_{m=1}^{\infty} P(H_M^K \geq m) \\
&= \sum_{m=1}^{\infty}\left[1 - P(H_M^K < m)\right] \\
&= \sum_{m=1}^{\infty}\left[1 - \prod_{i=1}^{N} P(H_i^K < m)\right] \\
&= \sum_{m=1}^{\infty}\left[1 - \left(1 - P_{ef}^{K(m-1)+1}\right)^N\right]
\end{aligned} \quad (3)$$

Note that $E(\sum_{k=1}^{H_M^K} RT_k)$ is the expected total time needed for transmission of the missing PDUs until all of them have been successfully received. Thus we have

$$E\left(\sum_{k=1}^{H_M^K} RT_k\right) = \sum_{i=1}^{N} E(H_i^K)T_{PDU} = (K \cdot N \cdot T_{PDU})\left(\frac{P_{ef}}{1 - P_{ef}^K}\right) \quad (4)$$

So the number of PDUs required in the K-retransmission-based retransmission stage is

$$K \cdot N \frac{P_{ef}}{1 - P_{ef}^K} \quad (5)$$

Therefore, the expected total file delivery time of a transaction can be given as

$$N \cdot T_{PDU} + T_{prop} + \frac{E(H_M^K) \cdot 2T_{prop}}{1 - P_{er}} + \frac{K \cdot N \cdot P_{ef} \cdot T_{PDU}}{(1 - P_{er}) \cdot (1 - P_{ef}^K)} \quad (6)$$

In general, we have a better understanding of K-retransmission-based deferred NAK mode. We will validate the K-retransmission-based scheme in Section IV, and determine the proper K value.

IV. SIMULATION AND DISCUSSION

This section includes both analytical and experimental evaluations of our scheme. To that purpose, we define a ratio expressed in terms of the decrement of file delivery time by one extra retransmitted PDU to validate the performance of K-retransmission-based scheme, and the proper K is determined.

A. Validation

Under the aforementioned analysis, we have constructed our simulation scenarios. We consider a single link file-transfer operation. The typical configurations of parameters are listed in Table II.

We denote T_K the time spent in K-retransmission-based retransmission stage (with T_1 is corresponding to retransmission time needed in standard deferred NAK CFDP), and P_K the number of PDU requested in K-retransmission-based retransmission stage (with P_1 is corresponding to retransmission number of PDU needed in standard deferred NAK CFDP).

We then define a ratio as follows:

$$\Delta(K - 1) = -\frac{T_K - T_1}{P_K - P_1} \quad (7)$$

where $K \geq 2$, refers to the decrement of file delivery time by one extra retransmitted PDU transmitted in the K-retransmission-based stage, relative to the standard deferred NAK CFDP. We utilize this ratio to evaluate the performance of our scheme. Here we only consider the ratio of the Earth to Mars scenario to see general. The tests are performed in order to show how the performance changes with different packet error probability ranging from 0.01 to 0.5 and retransmission times K from 2 to 10. Combining (2), (3), (5) and (7), we plot a figure, as shown in Fig. 2, by using MATLAB tool.

It is straightforward that increasing the times of retransmissions, the values of ratio $\Delta(K - 1)$ decrease. We observe that the values of $\Delta(2 - 1)$ under different PER are all superior to the other cases. If the number of transmissions is further increased, from 3 to 10, when PER ranging from 0.1 to 0.5, the curves will tend to be flat. As highlighted in the previous case, best results are provided, when K is equal to 2, with maximum decrement of file delivery time at the cost of same number of PDUs. More precisely, the contribution to lower the file delivery at the cost of power is maximal in the case of $K = 2$. The peak values under different K and PER appear in the vicinity of $P_{ef} = 0.05$, the reason causing this phenomena will be a contribution in future work.

On the basis of above mentioned, we choose the proper K equal to 2, and name this scheme as Double Retransmission (DR) deferred NAK. We make some simulations in Section IV.B for a supplementary validation.
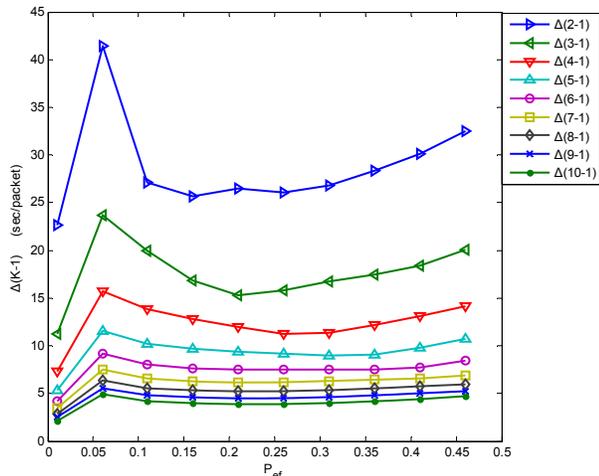
Figure 2.   Validation of *K*-retransmission-based deferred NAK. Earth to Mars scenario: File size = 1MB, transmission rate = 20kb/s, and one way propagation delay = 2.5 a.u. (1 a.u. = 480s).

TABLE II.   CONFIGURATION OF SIMULATION PARAMETERS.

Items	Scenario A	Scenario B	Scenario C
	GEO	*Earth to Moon*	*Earth to Mars*
File size	1MB		
N	1000		
P_{ef}	0.01: 0.05: 0.5, 0.1: 0.1: 0.5		
T_{prop}	0.12 sec	1.352 sec	2.5 a.u. (1200 sec)
T_{PDU}	0.008 sec	0.008 sec	0.4 sec

B.   Simulation and Analysis

From the above Fig. 2, it is clear that increasing the *K*, the decrement becomes smaller and smaller. Thus, we confine the maximum *K* down to 5, and the packet error probability in the range of 0.1 to 0.5, to make a supplementary validation under three typical scenarios which have been listed in Table II.

(a) GEO

(b) Earth to Moon
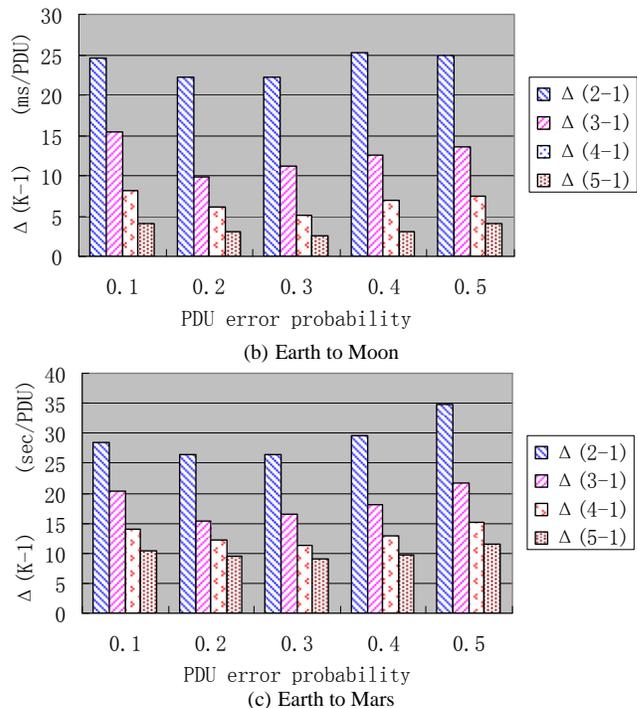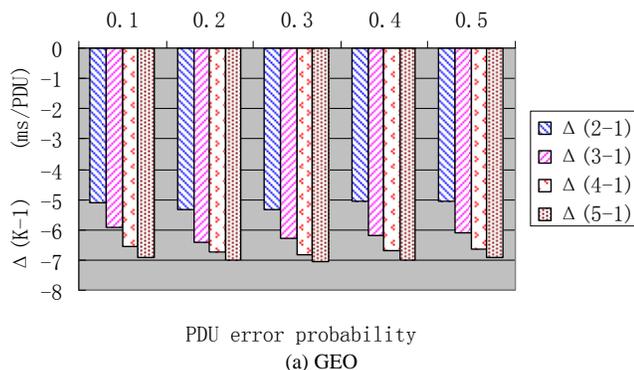
(c) Earth to Mars

Figure 3.   Simulation results under three typical Scenarios:
(a) GEO. (b) Earth to Moon· (c) Earth to Mars

From Fig. 3, it shows that the values of ratio $\Delta(K-1)$ decrease with the increase of *K* under different P_{ef} gradually, and the maximum decrement of file delivery time appears at $K = 2$.Moreover, the maximum values emerge when the packet lost rate is higher, especially in "deep-fading period". For scenario A, we note that the values of ratio are negative in Fig. 3(a). This is easily explained by the fact that the retransmission time is ruled by the number of retransmitted PDUs, in the case of one way propagation time is small. As far as scenario B and C are concerned, when $P_{ef} = 0.4, 0.5$, the value of $\Delta(2-1)$ is 25.3*ms* and 35*s* respectively, which are both the maximum. The longer distance from the earth to destination and the higher PER, the larger decrement of file delivery time can be realized, which can be informed by comparing the scenarios B and C. These are also in line with the previous results shown in Section IV. A.

V.   COMPARISON MEASUREMENTS

This section derives the numerical presentation of DR deferred NAK, and makes a validation between numerical analysis and random simulation. We carry out several tests to compare DR deferred NAK with standard deferred NAK.

A.   Numerical Analysis and Random Simulation

From (3) and (6), the expected total file delivery time of a transaction based on the DR deferred NAK scheme can be easily obtained as

$$T_{prop} + N \cdot T_{PDU} + \frac{E(H_M^2) \cdot 2T_{prop}}{1 - P_{er}} + \frac{2N \cdot P_{ef} \cdot T_{PDU}}{(1 - P_{er}) \cdot (1 - P_{ef}^2)} \quad (8)$$

Where

$$E(H_M^2) = \sum_{m=1}^{\infty} \left[1 - (1 - P_{ef}^{2m-1})^N \right] \quad (9)$$

We compare the numerical evaluation of (8) and the results of random simulation. The simulation results closely match the mathematically derived results, as can be observed form Fig. 4.
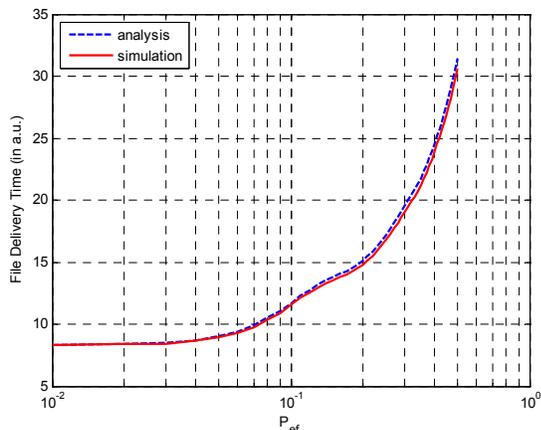
Figure 4.   DR deferred NAK: analytic and simulation results. File size = 1MB, transmission rate = 20kb/s, and one way propagation delay = 2.5 a.u.

B. Deferred NAK vs DR Deferred NAK

Under the simulation parameters presented in Table II, we implement several experiments to compare DR deferred NAK with standard deferred NAK. We only consider the formula (6) in this paper and the formula (7) in [4] (in Section III. C). The comparisons between them under three typical scenarios are shown in Fig 5. From Fig 5, we notice that the file delivery time ascends with the increase of packet loss rate. Although the DR deferred NAK does not fit the GEO scenario, it quite suits the latter two scenarios, especially when the packet loss rate is larger than 0.1. As discussed in [3] (in Section II. C, proposition 1), the $E(H_M^2)$ increases in logarithmic order with N. The expected file delivery time in (8) has a term that increased with N and a term that has the factor $E(H_M^2)$. For very long propagation delay, the product of multiplicative factor $E(H_M^2)$ and one way propagation time is much larger than that of the term linear of N, which is on the order of the PDU transmission time. In such an environment, as the number of PDUs in the file increases, the expected file delivery time is initially ruled by the term logarithmically growing with N, and the order of growth later becomes linear with a small multiplicative factor for larger values of N. Given as N is fixed, the expected file delivery time is absolutely dominated by the multiplicative factor and long one way propagation delay. For a small one way propagation delay (relatively to the PDU transmission time), the order of

growth is always ruled by the term linear of N. Generally, the DR deferred NAK is especially suited to scenarios with very long propagation delay, or high packet error rate, or both in deep space communication.

(a) GEO

(b) Earth to Moon
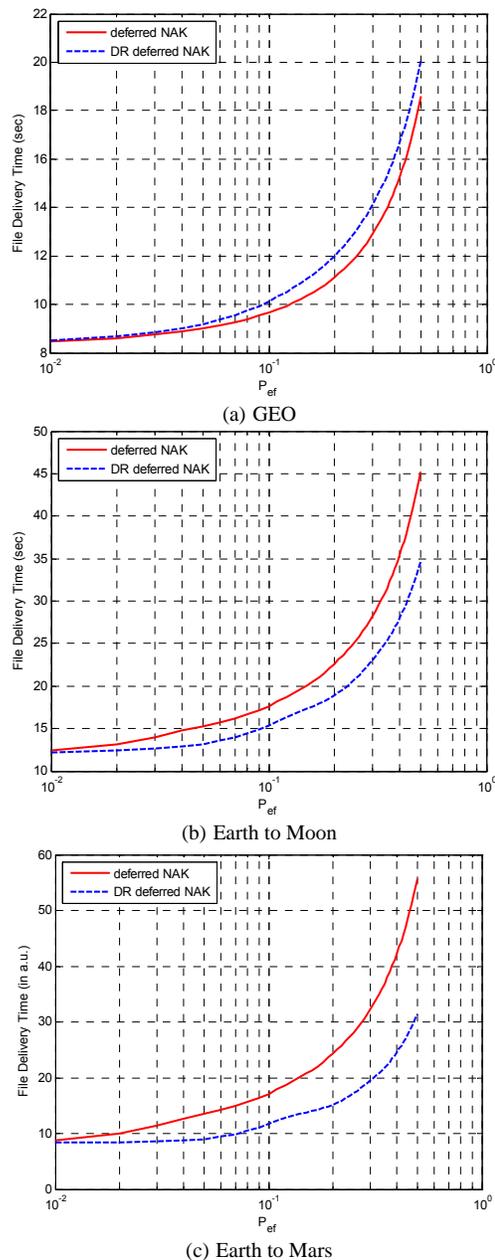
(c) Earth to Mars

Figure 5.   Comparison between deferred NAK and DR deferred NAK: (a) GEO. (b) Earth to Moon. (c) Earth to Mars

VI.   CONCLUSION AND FUTURE WORK

Future space communication needs simple, effective and reliable protocols. A single type of protocol would not meet the needs of all scenarios in future Delay-Tolerant Network (DTN). We have reconsidered the deferred NAK CFDP and integration of the retransmission concept. We have put forward an improved scheme utilizing double retransmission

mechanism to standard deferred NAK CFDP. The protocol make an efficient trade-off between the file-delivery time and the throughput efficiency under several typical scenarios. Our theoretical evaluation and preliminary simulation results reveal that DR deferred NAK outperforms standard deferred NAK mode in terms of file delivery time. This improved scheme favors the scenarios with small connectivity time, long propagation delay and high PER greatly.

Future work will emphasize on the joint use of our scheme with erasure coding particularly. A deeper analysis of channel model combined with rain attenuation will also be tackled in our future work.

REFERENCES

[1] T. de Cola, E. Paolini, G. Liva, and G. P. Calzolari, "Reliability Options for Data Communications in the Future Deep-Space Missions," Proceedings of the IEEE, Vol. 99, Nov. 2011, pp. 2056-2074.

[2] T. de Cola, M. Marchese, "Reliable data delivery over deep space networks- Benefits of long erasure codes over arq strategies" IEEE Wireless Communications, Vol 17, Apr. 2010, pp. 57-65.

[3] Consultative Committee for Space Data Systems, CCSDS File-delivery Protocol (CFDP), Recommendation for Space Data System Standards, CCSDS 727.0-B-4, Blue Book, Jan. 2007.

[4] D. Lee, W. Baek, "Expected file-delivery time of deferred NAK ARQ in CCSDS file-delivery protocol," IEEE Transactions on Communications, Vol. 52 , issue 8 , Aug. 2004, pp. 1408-1416.

[5] W. Baek, D. Lee, "Analysis of CCSDS file-delivery protocol-immediate NAK mode", IEEE Transactions on Aerospace and Electronic Systems, Vol. 41, Apr. 2005, pp. 503-523.

[6] T. de Cola, H. Ernst, M. Marchese, "Performance analysis of CCSDS File Delivery Protocol and erasure coding techniques in deep space environments," Computer Networks, Vol. 51, Oct. 2007, pp. 4032-4049.

[7] G.Papastergiou, I.Psaras, V.Tsaoussidis, "Deep-Space Transport protocol: A Novel Transport Scheme for Space DTNs", Computer Communications, Vol.32, Oct. 2009, pp. 1757-1767.

A Weather-forecasting Improved CFDP over Ka-band Channel

Qiuping Zheng, Zhihua Yang, Jian Jiao, Qinyu Zhang

Department of Electronics & Information Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China

zhengqiuping1987@126.com, yangzhihua@hitsz.edu.cn, hellojesus@sina.com, zqy@hitsz.edu.cn

Abstract—A novel file delivery scheme over a Ka-band channel is proposed, which is heavily dependent on weather around the earth's atmosphere. The proposed scheme based on CFDP is designed to resist the influence of weather conditions by the weather-forecasting. In this work, the condition-forecasting model that the weather sequence is forecasted through the first two states is constructed, demonstrated by the theoretical analysis and furthermore, implement of forecasting algorithm is also carefully researched and designed. Simulating results show that the file delivery time of selection-waiting CFDP with weather-forecasting is reduced effectively compared with traditional CFDP.

Keywords-Ka-band; weather-forecasting; CFDP; file delivery time; NAK

I. INTRODUCTION

As a result of long round trip times and frequent packet errors, the CFDP (CCSDS File Delivery Protocol) for deep-space communications was proposed by the CCSDS (Consultative Committee for Space Data Systems) [1][2]. On the one hand, deeper space exploration, further probing, and even larger amount of data transferred, not only lead to more and more CFDP retransmissions, but also increases of the file transfer delay; thus, the validity of the data transmission was seriously affected. On the other hand, as the Ka-band is strongly influenced by weather conditions, CFDP has no relevant functions to resist the state's high error rate of bad weather. Although Ka-band links can be available to a large range of bandwidth, which will greatly increase the capabilities and capacity of the future space networks, yet they are highly vulnerable to fluctuations in weather, such as the rainfall. In this paper, the selection-waiting CFDP in order to meet the change of Ka-band weather conditions is proposed due to the original CFDP was highly improved in that the file transfer delay was decreased efficiently.

Some works had been done in order to improve the performance of CFDP by researchers. In [3], an automatic repeat-request (ARQ) scheme of the Consultative Committee for Space Data Systems file-delivery protocol for the single-hop file-transfer operation was presented, and the CFDP performance was in detail analysed in [4-7], which came down to the cislunar, geo-stationary Earth orbit (GEO) and low Earth orbit (LEO) link environments.

A novel file delivery mechanism over a Ka-band channel is proposed in this paper, which is heavily dependent on weather of the earth.

This paper is organized in the following way. Firstly, the related works of our study are briefly explained, which contain deferred NAK CFDP transmission mechanism and Gilbert-Elliot channel. Secondly, the weather-forecasting algorithm is thoroughly demonstrated. Thirdly, the selection-waiting CFDP is designed and the simulation of the novel file delivery mechanism is illustrated and satisfied results are gained.

II. PROBLEM FORMULATION

The Deferred NAK CFDP Transmission Mechanism is produced in the background that TCP/IP uses less efficiently in space communication, but the CFDP is a transmission protocol for the application layer [3]. Then CCSDS proposed four different ways for NAK in CFDP; they are Immediate NAK, Prompted NAK, Asynchronous NAK and Deferred NAK. In this paper, we take Deferred NAK CFDP as an example of analysis simplicity. As is shown in Figure 1, when file transmission states, the object file is divided into a set of protocol data units (PDUs) before being sent. At the first "spurt", the sending entity of CFDP core procedure issues all PDUs of file in order. As soon as the EOF PDU is received correctly, the receiving entity acknowledges it with ACK (EOF) and then checks the integrality and validity of object file by file information contained in the Meta PDU. Those sequence numbers of missing PDUs are listed and reported back to sending entity by a NAK message. According to the information in NAK, sending entity initiates the second "spurt" consisted of the PDUs required by receiving entity, and so on. Once all of object file PDUs are received successfully, the receiving entity returns back a FIN PDU to the sending entity, which requires an acknowledgement with ACK (FIN). When ACK (FIN) is successfully delivered to the receiving entity, this transmission is finished and closed.
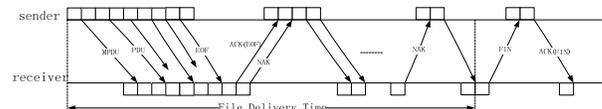
Figure 1. Two error state AWGN channel Gilbert-Elliot model

In the study of channel model, according to the Gilbert-Elliot channel [8], a variation of the additive white Gaussian noise (AWGN) channel is used to model different bit error rate (BER) in each weather state. This type of

channel, called the Gilbert-Elliot (GE) channel [8], has two weather states, a good and a bad weather state, which are separated by the threshold value. During good weather conditions, most of the transmitted packets will be received successfully, during bad weather conditions, however, most of the transmitted packets will generate some errors due to the high noise temperature at receiver. Therefore, two different BERs are applied to each good and bad weather state. In our new model, we define the relatively high BER value for the bad weather state and relatively low BER value for the good weather state.

To capture the weather correlation, the Gilbert-Elliot channel with two weather states are shown in Figure 2. The transition from one state to another state is defined by the transition matrix *P*, which completely characterizes the channel behavior. In this model, the current state is determined by the previous state, and the λ_G and λ_B are the transition probabilities from good to bad and from bad to good state, respectively.
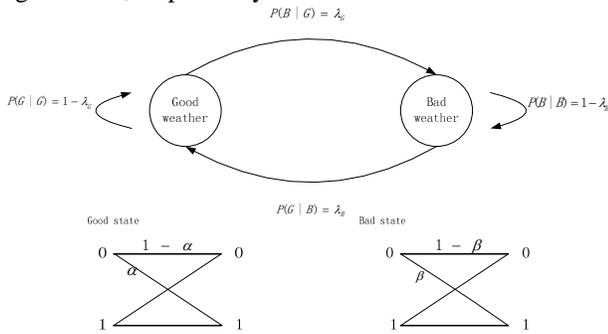
Figure 2.   Gilbert-Elliot channel

Then the transition matrix is

$$P = \begin{bmatrix} P(G \ / \ G) & P(B \ / \ G) \\ P(G \ / \ B) & P(B \ / \ B) \end{bmatrix}$$
$$= \begin{bmatrix} 1 - \lambda_G & \lambda_G \\ \lambda_B & 1 - \lambda_B \end{bmatrix}$$

(1)

The difference of the file transfer delay between Gilbert-Elliot (GE) channel and AWGN channel is illustrated in Figure 3, which depicts MRO (Mars Reconnassance Orbiter) sending files using the CFDP over Ka-band. Under the same condition (files are at most 10MB in size and transmission rate is 200kb/s) the simulation was carried out and the result was compared. Figure 3 shows that the file transfer delay over GE channel is longer than AWGN. It is said that the performance of CFDP is poor over Gilbert-Elliot channel. As a result, it is necessary to improve the CFDP over Ka-band.
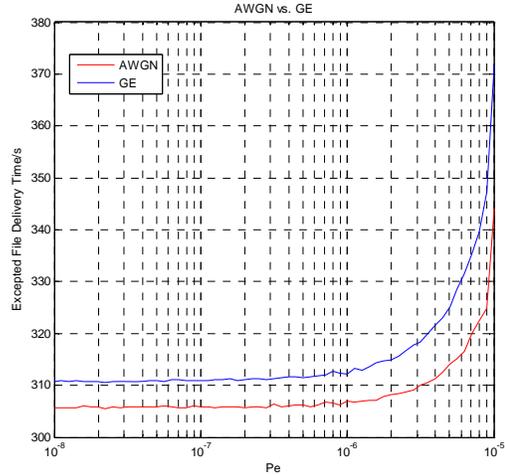
Figure 3.   The file transfer delay over Gilbert-Elliot channel and AWGN channel

III.   WEATHER-FORECASTING ALGORITHM

A. Condition-forecasting Model and Assumption

It is known that the one-way packet transmission time is $T = 20 \ \text{min}$. Mars sends data packets to the earth, and then the earth returns the control commands to the Mars. The time spent in this process is $2T$. We consider the bit-error-rate of the receiver (P_e) as threshold, when $P_e < 10^{-5}$, we believe that the weather state is good. Otherwise, it is bad.

$$\begin{cases} P < P_e, good \ state \\ P > P_e, bad \ state \end{cases}$$

(2)

As is shown in Figure 4, assume that the initial channel state is good ('1'), we can forecast the next state according to the Ka-band channel model. Then, we forecast the third state due to the first two states. While forecasting the fourth state, the true value of the third can be determined because of the NAK feedback mechanism. We set the third state as the new initial state to predict the next two states. In this way, the so-called two-step forecast is carried out.
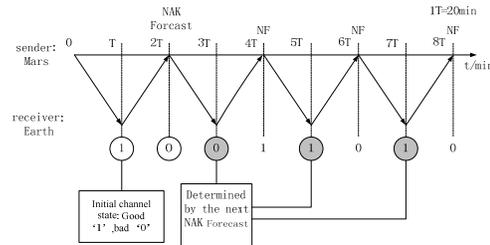
Figure 4.   State of channel forecasting theory according to Ka-band

Step 1, When the file starts to transmit, the receiver records the error rate every time when the file reaches the

earth. According to the error rate at this moment, the receiver can estimate the weather condition: a good state denoted by 1 and a bad state denoted by 0.

Step 2, condition-forecasting model is worked by the written '0' or '1' into the received words of the NAK, and the sender could get this weather condition when the NAK reaches Mars.

Step 3, when the sender receives the NAK with the value '1', it judges that the weather condition is good, otherwise bad, as is shown in Figure 5.
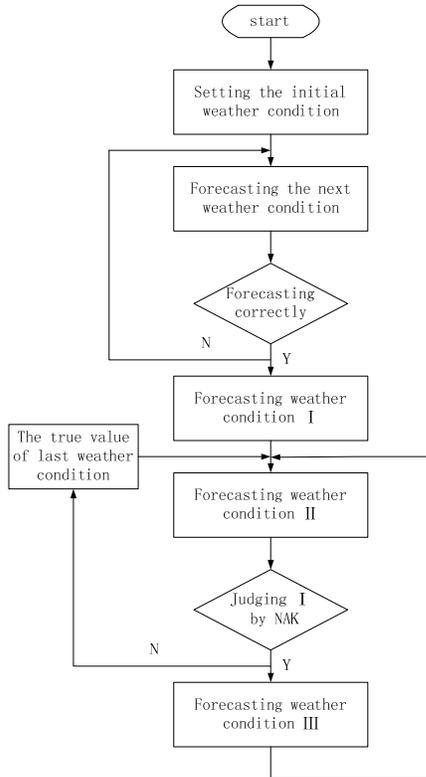
Figure 5.   The condition-forecast flow chart

B. Error Analysis

According to the GE channel model mentioned above, the steady probability can be attained when the probability of good weather is P_G and that of bad is P_B. If the initial state is bad('0')，one-step prediction error probability is

$$P_{e0} = 2\lambda_B(1 - \lambda_B)$$ (3)

Similarly, when the initial state is good ('1'), one-step prediction error probability is

$$P_{e1} = 2\lambda_G(1 - \lambda_G)$$ (4)

So, the sum of the one-step prediction probability is

$$P_{error} = P_B \cdot P_{e0} + P_G \cdot P_{e1}$$ (5)

The two-step state transition matrix is

$$P^2 = \begin{bmatrix} (1 - \lambda_G)^2 + \lambda_G\lambda_B & 2\lambda_G - \lambda_G^2 - \lambda_G\lambda_B \\ 2\lambda_B - \lambda_B^2 - \lambda_G\lambda_B & (1 - \lambda_B)^2 + \lambda_G\lambda_B \end{bmatrix}$$ (6)

Similarly, two-step prediction probability is

$$P'_{error} = P_B \cdot 2(2\lambda_B - \lambda_B^2 - \lambda_G\lambda_B)\left[(1 - \lambda_B)^2 + \lambda_G\lambda_B\right]$$
$$+P_G \cdot 2\left[(1 - \lambda_G)^2 + \lambda_G\lambda_B\right](2\lambda_G - \lambda_G^2 - \lambda_G\lambda_B)$$ (7)

C. Implement of the Forecasting Algorithm

CFDP NAK [2] packet format originally has reserved 3-bit (as shown in Figure 6). The reserved words can be used to carry the value of the initial state of the two-step prediction. Selecting the first reserved word, when the initial state is good, the value '1' will be written, otherwise we will write '0'.

Figure 6.   Fixed PDU header of CFDP

IV.   DESIGN OF SELECTION-WAITING CFDP

According to the analysis of the predictive model about the Ka-band channel, the selection-waiting CFDP mechanism is chosen to overcome shortcomings in the light of the change of the weather conditions (as is shown in Figure 7).
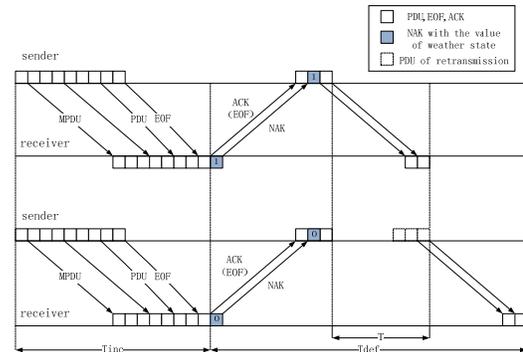
Figure 7.   Selection-waiting CFDP transmission mechanism

TABLE Ⅰ NOTATION

Symbol	Definition
W	Total File Delivery Time

T	Once-Waiting Time
T_{prop}	One-way propagation delay
T_{inc}	Transmission Detection Time at First Time
T_{def}	Time of Retransmission
P_e	Bit Error Rate
P_{ePDU}	Pocket-loss-rate of PDU
P_{eEOF}	Pocket-loss-rate of EOF
L_{PDU}	Length of PDU
$T_{ACK(EOF)}$	Transmission Time of ACK(EOF)

Based on the tow-step forecasting of NAK, when the channel condition is good, the sender transmits the file to the receiver normally. However, when the channel is bad, the sender will make the following judgments: (1) if the time of the file-transmission time>T(20min), it will wait for the time value of T, and then transmit sequentially；(2) if the time of the file-transmission time<T, it transmits the file directly.

The relationship between the pocket-loss-rate of PDU and the bit error rate is

$$P_{ePDU} = 1 - (1 - P_e)^{L_{PDU}} \qquad (8)$$

The pocket-loss-rate of the EOF is

$$P_{eEOF} = 1 - (1 - P_e)^{L_{EOF}} \qquad (9)$$

The time value of the EOF timer is set as

$$time\_EOF = 2T_{prop} + T_{ACK(EOF)} \qquad (10)$$

The mathematical expectation of the transmission of EOF is

$$E(EOF) = \frac{P_{ePDU}(2T_{prop} + T_{EOF} + T_{ACK(EOF)})}{1 - P_{ePDU}} \\ + T_{prop} + T_{EOF} \qquad (11)$$

We can get the Transmission Detection Time at first time is

$$T_{inc} = NT_{PDU} + E(EOF) + T_{ACK(EOF)} \\ = NT_{PDU} + \frac{P_{ePDU}(2T_{prop} + T_{EOF} + T_{ACK(EOF)})}{1 - P_{ePDU}} \\ + T_{prop} + T_{EOF} + T_{ACK(EOF)} \qquad (12)$$

The time value of NAK is

$$time\_NAK_i = 2T_{prop} + R_i \qquad (13)$$

and，

$$E(\sum_{i=1}^{M_N} R_i) = NT_{PDU}(\frac{1}{1 - P_{ePDU}} - 1) \qquad (14)$$

$$E(M_N) = 1 + \sum_{m=1}^{\infty} [1 - (1 - P_{ePDU}^m)^N] \qquad (15)$$

The time of retransmission at ith time is

$$T_{def_i} = \begin{cases} n_i \cdot T_{pdu} + 2 \cdot T_{prop}, good\_weather \\ n_i \cdot T_{pdu} + 2 \cdot T_{prop} + T, bad\_weather \end{cases} \qquad (16)$$

So the time of retransmission is

$$T_{def} = \sum_{i=1}^{M_N} T_{def_i} \qquad (17)$$

And the total of file delivery time is

$$W = T_{inc} + T_{def} \qquad (18)$$

V. SIMULATION AND DISCUSSION

The Earth-to-Mars communication is taken as an example to testify the rationality of selection-waiting CFDP proposed in this paper. Some experiments parameters are shown in Table Ⅱ and Table Ⅲ.

TABLE Ⅱ PARAMETERS IN SIMULATION

Parameters	Parameters Descriptions	Value
P_e	Bit-error-rate of Direct Point-to-point Communication Between Mars and Earth	10^{-4}、10^{-5}、10^{-6}、10^{-7}、10^{-8}
Tx(kb/s)	Transmission Rate	200
N	Number of PDUs	10^5、$2*10^5$...
L_{pdu}(Kbyte)	Length of PDU	2
d(km)	The Distance between Mars and Earth	4×10^8

TABLE Ⅲ VALUE OF TRANSITION MATRIX

transition probability	Sampling by averaging	Sampling by choosing the max	Sampling by choosing the min
P(G/G)	0.9773	0.9656	0.9853
P(B/G)	0.0227	0.0344	0.0147
P(G/B)	0.1667	0.1618	0.1799
P(B/B)	0.8333	0.8382	0.8201

Figure 8 compares the performance of selection-waiting CFDP and traditional CFDP as the numbers of PDU varies from 1×10^5 to 1×10^6 under different bit error rates. The results show that the file delivery time of both selection-waiting CFDP and the traditional CFDP increases in line ratio as the number of PDU increase. Apparently, as the bit-error-rate decreases, the file delivery times both increase based on two CFDPs. In Figure 8, the performance of CFDP with weather-forecasting mechanism is better than that of the Gilbert-Elliot channel. It is said that Selection-waiting CFDP could reduce approximately 5×10^3s time at least, compared with the traditional CFDP. As the result, it means that Selection-waiting CFDP is more applicable than that of

the traditional CFDP in the space communication over Ka-band heavily influenced by the weather conditions, which is significant for cutting down the transmission time for long-distance space communication links such as the Earth-to-Mars communication.
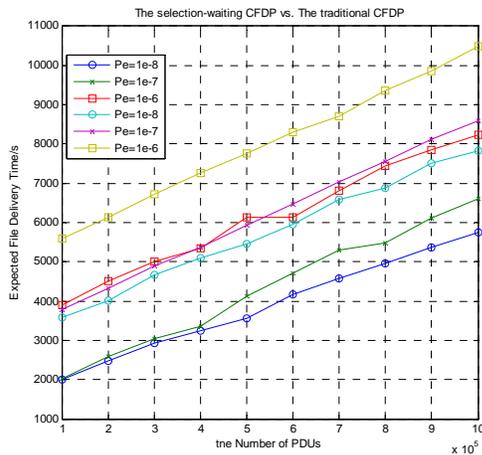
Figure 8.   The file delivery time of the selection-waiting CFDP vs. The traditional CFDP under the different bit-error-rates

We choose the transition matrix which is sampling by averaging in above simulation. In fact, the transition probability from good to bad states can be derived by 1/(average duration of good weather) and the transition probability from bad to good state is 1/(average duration of bad weather). The transition probabilities of different sampling methods with the threshold of *20K,*are shown in the Table Ⅲ. These transition probabilities will be used for simulating weather patterns.

VI.   CONCLUSION

In this study, the weather-forecasting algorithm is presented to evaluate the CFDP performance under good and bad weather conditions, and a novel file delivery protocol, inserting the weather-forecasting algorithm into traditional direct point-to-point link, is proposed. Based on the condition-forecasting model eatablished, comprehensive experiments are carried out and discussed under different conditions of parameters involved in the proposed protocol. Results indicate that the proposed protocol perform much better than that of the traditional Deferred NAK CFDP in terms of file delivery time.

Further work will consider the selection of transmission rate into the performance of the proposed protocol over Ka-band. And another algorithm about the perception of channel will be proposed and will be compared with the protocol proposed in this work.

VII.   ACKNOWLEDGMENT

REFERENCES

[1]   CCSDS File Delivery Protocol (CFDP) Part 1: Introduction and Overview. CCSDS 720.1-G-3[EB/OL]. April 2007. http: // ccsds. org. pp. 20-30

[2]   CCSDS File Delivery Protocol (CFDP) Part 2: Implementers Guide. CCSDS 720.2-G-3[EB/OL]. April 2007. http: // ccsds. org. pp. 1-20

[3]   Lee D C, and Baek W. Expected file-delivery time of deferred NAK ARQ in CCSDS file-delivery protocol [J]. IEEE Transactions on Communications. 2004, 52(8): pp.1408-1416.

[4]   Ruhai W, Shrestha B L, and Xiaoli M. Channel Delay Impact on CCSDS File   Delivery Protocol(CFDP) over Space Communications Links[A].ICC'07[C]. Glasgow, U.K., 2007. pp. 5201-5205.

[5]   Ruhai W, Rudraraju D L, and Rapet P K V, et al. Performance Evaluation of CCSDS File Delivery Protocol (CFDP) in Deferred NAK mode over Geostationary Earth Orbit (GEO)-Satellite Links[A].ICC'07[C]. Glasgow, U. K., 2007. pp. 4421-4425.

[6]   Ruhai W, Shrestha B L, and Xuan W, et al. Experimental Investigation of CCSDS File Delivery Protocol (CFDP) over Cislunar   Communication   Links   with   Intermittent Connectivity.ICC'08.Beijing,China,2008. pp. 1910-1914.

[7]   Wang R H, Shrestha B L, and Wu X, et al. Unreliable CCSDS File Delivery Protocol (CFDP) over Cislunar Communication Links[J]. IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS. 2010, 46(1): pp. 147-169.

[8]   Sung I. U, and Jay L. Gao. "CFDP Performance over Weather-Dependent Ka-band Channel", in: Proceeding of SpaceOps 2006, Jun. AIAA-2006-5968

Optimum Selective Beam Allocation Scheme for Satellite Network with Multi-spot Beams

Unhee Park, Hee Wook Kim, Dae Sub Oh and Bon Jun Ku

Satellite Handheld Transmission Research Team

ETRI

Daejeon, Korea

{unipark, prince304, trap, bjkoo}@etri.re.kr

Abstract—**We are living in the world in which resources are limited compared to the traffic demands, which are increasing rapidly. In this environment, it is essential for the resource management techniques to maximize the effectiveness of the resource utilization. In particular, since the satellite resource such as on-board power, bandwidth even beam itself are expensive and limited due to its inherent characteristics, an effort to enhance the efficiency on their utilization is crucial. In this paper, we consider beam resource management in the N multi-beam satellite system. The simulation result demonstrates that the proposed scheme selecting only K(≤N) active beams assigning the beam power can achieve the increase of a total throughput compared to the non-selective beam allocation.**

Keywords-beam selection; resource management; multi-beam satellite; wireless communication

I. INTRODUCTION

A resource management is one of the parts of important techniques in the wireless communications. In this system, it is possible to achieve a improve efficiency in a volatile environment depending on how to manage the resources. Since the satellite resources such as power, bandwidth and the use of a spot beam are expensive and scarce, the effort to enhance their efficiency is crucial. And also, recently, the multi-beam antenna technique plays important role in the satellite communications as it can be flexible network configuration. In previous research [1]-[4], the beam power allocation schemes were proposed using advantages of parallel multi-beams which is monotonically increasing the capacity gain with the number of beams. In particular, Jihwan et al. [1] addressed the issue of long term average gain in terms of Shannon capacity and power efficiency. It emphasized the mathematical formulation and analytic solutions of the optimum resource allocation problem as well as explained the trade off problem between the total system capacity and fairness among all spot beams with traffic demands. It noted that if we want to maximize the total system capacity, we should allocate a fixed amount of power regardless the traffic demand which exceeds the some threshold for each beam and sacrifice the loss of proportional fairness. It is well known that the maximum total capacity can be achieved by water-filling approach [5].
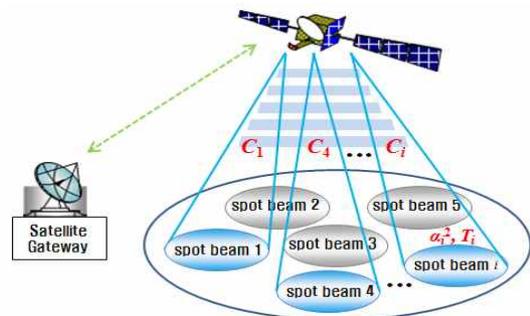
Figure 1. Multi-beam satellite system applying active beam selection scheme

On the other hand, it needs to allocate more power to beam with higher traffic demand and channel condition to achieve the proportional fairness over spot beams, but loses some total capacity. Referring to this, we want to allocate the beam power and small number of K active beams among the $N(>K)$ multiple beams which can be illuminated simultaneously to achieve more total system capacity.

The rest of the paper is organized as follows. Section II describes the background of optimum selective beam allocation scheme, and presents how to give the highest priority to the active beam for multi-beams within a satellite coverage area. Section III presents simulation result to compare the proposed scheme with the case of non-selective beam allocation scheme. Finally, the conclusion is drawn in Section IV.

II. OPTIMUM SELECTIVE BEAM ALLOCATION SCHEME FOR MULTI-BEAM SATELLITE SYSTEM

A. Background of K active beam selection over N multi-beam satellite system

Fig. 1 shows the system configuration of a multi-spot-beam satellite selecting the active beams to allocate the resource. In the network, a multi-spot-beam satellite in geostationary orbit and an ensemble of spot beam cells are deployed in the network. Each spot beam has different traffic demand T_i and signal attenuation α_i^2 by channel conditions. The concept of active beam selection scheme is to allocate the capacity C_i subject to the chosen $K(<N)$ active beams. The purpose of the beam selecting

algorithm is to maximize the effectiveness of the resources utilization, and then total system capacity can be improved. As parts of this effort, we consider to make the reasonable beam power allocation for each spot beam with different traffic demand and channel condition [1]. Since the real traffic is non-uniform and time is varying, the resource management must reflect the different traffic distribution and channel conditions across all spot beams. Using the time sharing for Gaussian broadcast channels [5], we can obtain the Shannon bounded beam capacity of $C_i = W \log_2(1 + \alpha_i^2 P_i / WN_0)$, where N_0 is the noise power density and W is the allocated bandwidth. P_i is the beam power to be allocated and adjusted. In this paper, we mainly consider that downlink channels have the property of rainfall attenuation, which are slow fading events.

B. The review of optimum power allocation(OPA) scheme

As one of the metrics to evaluate the system performance of resource allocation over satellite downlinks, Jihwan et al. [1] addressed some tradeoff between different objects for system optimization. They derived the downlink multi-beam capacity optimization problem and proposed a schematic method. It focused on the best case when available capacity matches demand under the assumption of a simplified model. Minimization of difference between the supported capacity C_i and traffic demand T_i for ith beam can provide a reasonable objective for resource allocation, considering trade-off between the maximization of total capacity and the support of fairness among beams with traffic demand. In view of this, a square deviation cost function between capacities and traffic demands was adopted as given below.

$$\arg \min_{C_i} \sum_{i=1}^{N} (T_i - C_i)^2.$$ (1)

where

$$C_i \leq T_i \quad \text{for } i=1,2,\ldots,N$$ (2)

$$\sum_{i=1}^{N} P_i \leq P_{total}$$ (3)

First constraint in (2) means that the capacity allocated to each beam cannot exceed the traffic demand generated in each beam to prevent the unnecessary waste of resources. The condition in (3) implies the power for whole spot beams should be allocated within total system power and is called as aggregate power constraint. Applying the Lagrangian function as $L(P_i, \Lambda) = \sum_i (T_i - C_i)^2 + \Lambda(\sum_i P_i - P_{total})$, it can obtain the optimum beam power profile P_i, which should satisfy as follow equation (4).

$$T_i - W \log\left(1 + \frac{\alpha_i^2 P_i}{WN_0}\right) = \frac{\Lambda N_0 \ln 2}{2}\left(\frac{1}{\alpha_i^2} + \frac{P_i}{WN_0}\right),$$ (4)

where Λ is a Lagrange multiplier that is determined by the total power constraint of (3). If Λ is a nonnegative value, it means that the determined bandwidth by (4) satisfies the constraint $C_i \leq T_i$ of (2).

As we mentioned, this OPA scheme needs to allocate more power to beam with higher traffic demand and channel condition to achieve the proportional fairness among the spot beams, but loses some total capacity. Therefore, we present a proper beam selection algorithm to achieve more total system capacity based on OPA scheme.

C. OPA with selective beam allocation scheme

In order to derive the selective beam allocation algorithm, we adopt another constraint as below.

$$P_i \geq 0$$ (5)

The constraint (5) is added to see which users should be served with non-zero power and to consider the beam selecting algorithm [6]. Applying the Lagrangian function as $L(P_i, \Lambda, s_i) = \sum_i (T_i - C_i)^2 + \Lambda(\sum_i P_i - P_{total}) + \sum_i s_i(-P_i)$ and differentiating with respect to P_i, we can derive the following (6).

$$\frac{\partial L(P_i, \Lambda, s_i)}{\partial P_i} = \frac{\partial \sum_i (T_i - C_i)^2}{\partial P_i} + \Lambda - s_i$$ (6)

where, Lagrangian multipliers $s_i(\geq 0)$ are for $-P_i \leq 0$, and Λ for the total power constraint. The Kuhn-Tucker condition [7] can yield $s_i = 0$ if $P_i > 0$, and $s_j \geq 0$ if $P_j = 0$. First, applying this condition to (6), we have the $s_i = 0$ and $\partial L/\partial P_i|_{P_i = P_{i*}} = 0$ at the optimum beam power $P_i^* > 0$. From (6), we obtain as follow.

$$-\frac{\partial \sum_i (T_i - C_i)^2}{\partial P_i}\bigg|_{P_i = P_i^*} = \Lambda < -\frac{\partial \sum_i (T_i - C_i)^2}{\partial P_i}\bigg|_{P_i = 0}$$ (7)

where the inequality is valid by concavity of capacity C_i. Next, when the optimal power is $0 (P_j^* = 0)$ for jth beam, we have the $s_j \geq 0$ and $\partial L/\partial P_j|_{P_j=0} = 0$. Then, (6) can be expressed as in (8).

$$-\frac{\partial \sum_i (T_i - C_i)^2}{\partial P_j}\bigg|_{P_j^* = 0} + s_j = \Lambda \geq -\frac{\partial \sum_j (T_j - C_j)^2}{\partial P_j}\bigg|_{P_j = 0}$$ (8)

By comparing the right sides of (7) and (8) with respect to common Λ, we can find the optimum policy to select K active beam with highest value of

$$-\frac{\partial \sum_i (T_i - C_i)^2}{\partial P_i}\Bigg|_{P_i=0} = -\frac{\partial \sum_i \left(T_i^2 - 2T_i C_i + C_i^2\right)}{\partial P_i}\Bigg|_{P_i=0} \quad (9)$$

Inserting $C_i = W\log_2(1+\alpha_i^2 P_i/WN_0)$ to (9),

$$-\left[-2T_i W \frac{\frac{\alpha_i^2}{WN_0}}{\left(1+\frac{\alpha_i^2 P_i}{WN_0}\right)\ln 2} + W^2 \frac{2\left(\frac{\alpha_i^2}{WN_0}\right)\log_2\left(1+\frac{\alpha_i^2 P_i}{WN_0}\right)}{\left(1+\frac{\alpha_i^2 P_i}{WN_0}\right)\ln 2}\right]\Bigg|_{P_i=0}$$

$$= \frac{2T_i \alpha_i^2}{N_0 \ln 2} \quad (10)$$

In other words, in order to select the K active beams among the N spot beams with traffic demand, we only need to investigate the value of (10) across all spot beams. Whereas if we look at it, we can find that the factor affecting to value of (10) is only term of $T_i \alpha_i^2$. It can deem that the higher traffic demand and better channel condition for whole beams give the highest priority to the active beam. It can achieve the improvement of system capacity, but lose the fairness for some of the beams. We can compare the performance when used with beam selection algorithm in the next section.

III.   SIMULATION RESULTS

This section presents the simulation result. For the purpose of performance comparison, we create a simplified simulation model as follows. We evaluate the performance of beam selection scheme in an S-band GEO satellite channel. Also, we assume non-uniform signal attenuation α_i^2 (≤ 1) across each spot beam, and generate randomly for the simulation. Table 1 represents the system parameters used in our simulation model for fig. 2.

Fig. 2 shows the capacity distributions of spot beams that are allocated by water-filling method and optimum power allocation scheme and OPA with beam selection scheme, respectively. It is well known that the maximum total capacity can be achieved by water-filling approach. As confirmed in Table 2, indeed, a water-filling scheme can achieve greater total system capacity than the conventional OPA scheme. On the other hand, the selective beam allocation scheme results in total throughput gain compared to non-selective allocation scheme. However, in some cases, it cannot be selected as active beam based on worse channel condition despite the beam with high traffic demand by considering the total system capacity (e.g., ith beams for i= 10, 14, 15, 18, 20 in the Fig. 2). From this result, we can infer that we can achieve more total system capacity, but lose the fairness for some of the beams. In other words, we face a trade-off problem between the maximum total capacity and fairness among the spot beams.

TABLE I. SYSTEM PARAMETERS IN THE SIMULATION FOR FIG. 2

Parameters	Values
Number of spot beams, N	20
Number of active beams, K	15
Total system bandwidth	500 MHz
On board EIRP	80.33dBW
Free-space path loss [8], f_s=2.5GHz	191.53 dB
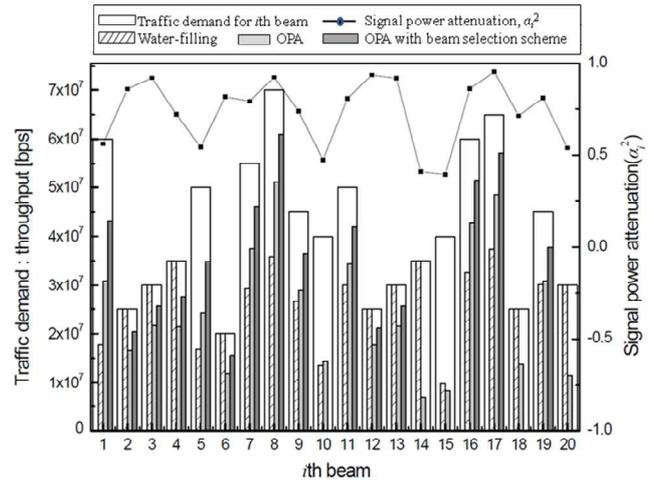
Figure 2. The comparison of throughput of ith beam for water-filling, OPA scheme and OPA scheme with selective beam allocation

In this regard, we focused on the minimization of the gap between supported C_i and T_i, and can confirm that the proposed scheme coincides more closely to the objective of this resource allocation through a comparison of the total sum of the gaps shown in Table 3.

TABLE II. THE COMPARISON OF TOTAL THROUGHPUTS

Schemes	Total throughput
Water-filling in [5]	534.64 Mbps
OPA scheme in [1]	494.75 Mbps
Proposed selective scheme	545.76Mbps

TABLE III. COMPARISON OF THE TOTAL SUM OF $\sum_i (T_i-C_i)^2$

Schemes	$\sum_i (T_i-C_i)^2$
Water-filling in [5]	7.0797E15
OPA scheme in [1]	5.6435E15
Proposed selective scheme	1.1480E15

IV.   CONCLUSION AND FUTURE WORK

In this paper, we proposed an optimum selective beam allocation scheme based on the traffic distributions and channel conditions for the parallel multi-beam satellite system. This paper studies how to allocate the active beam

considering trade-off problem between maximum total capacity and proportional fairness among beams with traffic demand by minimizing difference between beam traffic demands and allocated beam capacities assuming a simplified model and minimization problem. The simulation results demonstrate that the proposed beam selection scheme sacrifices the fairness for some of the beams, but nevertheless, can achieve more proportional fairness compared with water-filling method. In addition, it can achieve the best performance of total system capacity. In order to implement the interaction of the algorithm such as ACM, the practical simulation condition with weather impairments should be considered in future.

REFERENCES

[1] Jihwan P. Choi and Vincent W. S. Chan, "Optimum Power and Beam Allocation based on Traffic Demand and Channel Conditions over Satellite Downlinks," IEEE Trans.on wireless Communications, vol. 4, no. 6, pp. 2983-2993, Nov. 2005.

[2] Michael. J. Neely, Eytan Modiano, and Charles E. Rohrs, "Power Allocation and Routing in Multibeam Satellites with Time-Varying Channels," IEEE/ACM Trans. On Networking, vol. 11, no. 1, pp. 138-152, Feb. 2003.

[3] Kwang Jae Lim, Soo Young Kim and Ho Jin Lee, "Adaptive Radio Resource Allocation for a Mobile Packet Service in Multibeam Satellite Systems," ETRI Journal, vol. 27, no. 1, pp. 43-52, Feb. 2005.

[4] Yang Hong, Anand Srinivasan, Brian Cheng, Leo Hartman and Peter Andreadis, "Optimal Power Allocation for Multiple Spot Beam Satellite Systems", Proceedings of IEEE Radio and Wireless Symposium (RWS 2008), pp. 823-826, Orlando, FL, Jan. 2008.

[5] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*. New York: Wiley, 2006.

[6] Jihwan P. Choi and Vincent W.S. Chan, "Resource Management for Advanced Transmission Antenna Satellites," IEEE Trans. on wireless Communications, vol. 8, no. 3, pp. 1308-1321, Mar. 2009.

[7] Dimitri P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[8] ITU-R, "*Calculation of Free-Space Attenuation,*" Rec. P. 535-2, International Telecommunication Union, 1994.

Investigating the Effect of Bit Depth Rescaling on Spectral Bands Cross-Correlation in Hyperspectral DataCube

Ayman Mahmoud Ahmed

Egyptian Space Program

National Authorty for Remote Sensing and Space science

Cairo, Egypt

ayman.mahmoud@narss.sci.eg

Salwa ElRamly

Department of Electronics and Communication Engineering

Ain-Shams University

Cairo, Egypt

sramlye@netscape.net

Abstract— **Hyperspectral data analysis nowadays is one of the most common activities in the scientific community regarding data fusion and processing; most available free data through NASA websites are very helpful for researchers and postgraduate studies; Data has different formats band interleaved by pixel (BIP), Band sequential (BSQ) and Band-Interleaved-by-Line (BIL); also, it differs in bit depth from 12 to 16; as well as processed or raw format, headers included or without header; many researchers have to reformat the data for further use with either standard software like Environment for Visualizing Images software (ENVI), or custom developed software. This paper investigates the effect of rescaling pixel bit depth of hyperspectral data, on the spectral correlation between successive bands of the datacube. On the other hand, all other reformatting processes like band reordering and headers removal will not affect the spirit of the hyperspectral cube itself.**

Keyword - Hyperspectral Datacube; Spectral correlation; Bit depth; Spectral Analysis.

I. Introduction

Hyperspectral images typically have a high degree of spectral and spatial correlation. Consequently, data compression can significantly reduce hyperspectral data volumes to more manageable size for storage and communication.

Compressing hyperspectral data cube with Component Of The Shelf (COTS) coders has became an interesting area of research. A standard video coder has been used to compress hyperspectral data in many researches. A video is a sequence of still images, a hybrid video coder use the correlation in time by doing block based motion compensated prediction between images. In principle only the differences are transmitted. This method of coding is used on hyperspectral data if we consider one of the three dimensions as the time axis [1].

Hyperspectral data is distinguished by its unique spatial-temporal and spectral correlation for data in three dimensions, manipulating the data may affect one of its exclusive properties.

Investigating the spectral correlation between hyperspectral bands showed that, on one sample, there is a strong spectral correlation between different bands that are up to $40 - 50$ bands away [2].

Bit depth is the number of bits used to represent a pixel; dynamic range is closely related with bit depth which gives more accurate representation of the point being imaged.

For most free available hyperspectral data samples from airborne or space-borne, large bit depth is used (10-16 bit per pixel); on the other hand, many researchers have carried out their research using COTS program for the hyperspectral data; while COTS software available more likely to operate in 8 bits per pixel.

This paper begins with an introduction about the need of bit depth rescaling in hyperspectral data compression using video codec in the first section; while in the second section, bit depth rescaling is defined and explained; section three discusses the cross-correlation and equation to estimate it; fourth section shows how the experiment was carried out and data samples parameters; finally, results are shown and discussed.

II. Bit depth rescaling

Bit depth is one of the main attributes of any image. The more bits assigned to each pixel, the more possible intensities states the pixel can take. Conventional bit-depth scaling methods such as spatial 2D filtering and dithering algorithms [4] suffer from false contours or dithered pattern artifacts, which significantly deteriorate image quality in display devices.

Based on linear bit depth rescaling [5]; this experiment has been carried out, converting the bit depth from 12-16 bit/ pixel to 8bit/pixel.

Linear rescaling sets a minimum and maximum input value to 0 and 255 respectively and all other values in between are linearly aligned to intermediate output values.

In digital grayscale images one has a finite number of different intensities available, usually between 256 and 4096 levels, depending on the application. The number of uncompressed bits per pixel is the bit depth of the image.

Bit depth 8 allows 256 intensity levels; bit depth 12 allows 4096 intensity levels.

Downscaling of bit depth are investigated by researcher in order to enhance the process of rescaling and minimize the losses [6]; rescaling process makes each pixel intensity down scaled to another value while this should not affect visual sense of the human eye; in the field of hyperspectral imaging rescaling may affect the spectral correlation between

successive bands negatively, which may consequently result in different behavior when using a compression algorithm depend on inter-band correlation.

III. SPECTRAL CROSS-CORRELATION

A correlation is a number between -1 and +1 that measures the degree of similarity between two variables in our case it is bands or images. A positive value for the correlation implies a positive association (large values of X tend to be associated with large values of Y and small values of X tend to be associated with small values of Y). A negative value for the correlation implies a negative or inverse association (large values of X tend to be associated with small values of Y and vice versa).

Cross-correlation is used usually to describe the relation between two probably similar signals, hyperspectral datacube is distinguished with the significant highly correlated bands; since the same image is taken in different spectral bands; there is a high probability of having many regions within the image that reflects the light with the same intensity.

Measuring the inter band similarity with Cross-correlation using "(1)" indicated how these bands can be predicted, or generated from each other, this concept is used in hyperspectral data compression.

Cross-correlation is estimated using the following formula:

$$cor_{i,j} = cov(x_i, x_j)/\sqrt{(var(x_i)*var(x_j))} \qquad (1)$$

where $cor_{i,j}$ is the correlation value between bands i and j; $cov(x_i, x_j)$ is the covariance between bands i and j; $var(x_i)$ and $var(x_j)$ is the variance of the individual bands, i and j, respectively.

The normalization and mean values of the correlation between bands are implemented according to Fast Normalized Cross-Correlation [7].

The proposed algorithm assumes a template of one image then represents the template as sum of rectangular basis function, the correlation is then implemented to each basis function instead of the whole template, and the result is weighted sum of the correlation function of the basic functions.

IV. IMPLEMENTATION

The implementation was carried out by using ENVI [9] for rescaling of hyperspectral data down to 8-bits per pixel; bands Cross-correlation has been estimated using MATLAB [8] program, to generate Cross-correlation curves for both the original data and the downscaled data.

In Figures 1, 2, 3, 4, 5, 6 and 7, x-axis is the bands number starting from 0 till 224, while Y-axis is the Cross-correlation values with maximum of 1 represents identical highly correlated successive bands, while zero indicates no correlation between the two successive bands.

Cross-correlation between bands is estimated for the original data cube; the hyperspectral datacube is then linearly rescaled and the inter-band Cross-correlation is estimated for the new datacube with the same way and same order.

Hyperspectral data used is downloaded from Airborne Visible Infrared Imaging Spectrometer (AVIRIS) website [10]; the hyperspectral data cube has the following characteristics:

The files are in raw format (no header). All data samples are stored as: 2-byte integers, big-endian byte order, band-interleaved-pixel (BIP).

Each file is a 512-line with 224 bands.
- "hawaii_sc01.raw" is 512 lines x 614 samples x 224 bands, instrument bit depth = 12 bits
- "maine_sc10.raw" is 512 lines x 680 samples x 224 bands, instrument bit depth = 12 bits
- "Aviris_sc0" is 512 lines x 680 samples x 224 bands, instrument bit depth = 16 bits.
- "Aviris_sc3" is 512 lines x 680 samples x 224 bands, instrument bit depth = 16 bits.
- "Aviris_sc10" is 512 lines x 680 samples x 224 bands, instrument bit depth = 16 bits.
- "Aviris_sc18" is 512 lines x 680 samples x 224 bands, instrument bit depth = 16 bits.
- "f960705t01" is 512 lines x 680 samples x 224 bands, instrument bit depth = 16 bits.

In all the figures, solid RED line represents the inter-band correlation in the original data; while dotted Blue line is the downscaled data.

V. RESULTS

Seven files downloaded from the AVIRIS website, each figure below indicates the name of each file; in Figures 1, 2, 3 and 4, it can be seen that the highly correlated bands have lost their correlation after the downscaling process.
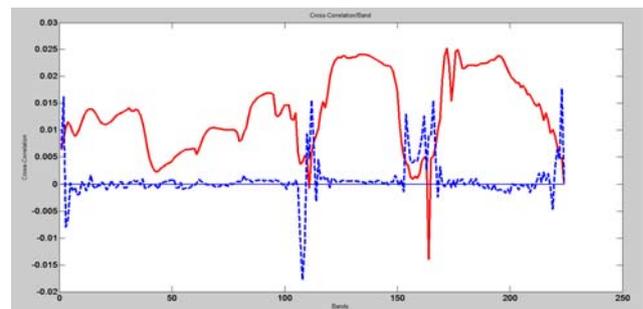
Figure 1, "Aviris_sc0".

In Figure 1, Inter band Cross-correlation has been lost after the process of bit depth rescaling.
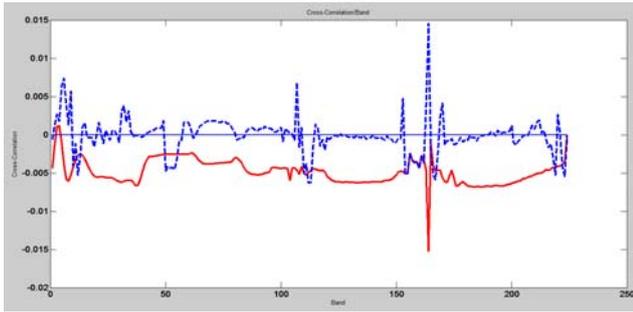
Figure 2, "aviris_sc3".

In Figure 2, Inter band Cross-correlation has been lost after the process of bit depth rescaling.

Figure 3, "f960705t01".

In Figure 3, Inter band Cross-correlation has been lost after the process of bit depth rescaling.

Figure 4, aviris_sc18.

In Figure 4, Inter band Cross-correlation has been lost after the process of bit depth rescaling.

On the other hand, in Figures 5, 6 and 7, bands kept their property in different -uncontinuous- manner.

Figure 5, "aviris_sc10".

In Figure 5, Inter band Cross-correlation has been preserved after the process of bit depth rescaling, but in a difference discrete manner.

Figure 6, hawaii.

In Figure 6, Inter band Cross-correlation has been preserved after the process of bit depth rescaling, but in a difference discrete manner.
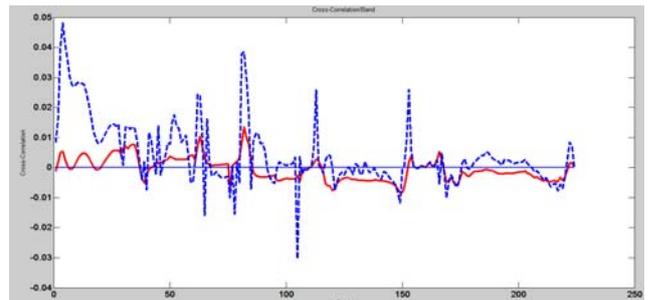
Figure 7, maine_sc10.

In Figure 7, inter band Cross-correlation has been preserved after the process of bit depth rescaling, but in a difference discrete manner.

Losing about 50% of pixel`s value, even in with a linear relative manner, have shown that, this is not a linear process;

and correlation cannot be kept even if the ratio between values across bands is preserved.

Losing Cross-correlation is significantly observed in the results, which gives a conclusion about the similarity behavior between bands.

A significant point can be noted in almost all curves; that the downscaling process reverses the correlation behavior between bands.

CONCLUSION

The process of bit depth rescaling for hyperspectral data cube takes a significant time of processing, as well as estimating the Cross-correlation between successive bands in the datacube.

Processing one datacube takes about 420-460 second using Matlab 2010 running in Intel core2 due Quad (3 GHz) processor.

Surprisingly, inter-bands Cross-correlation has been lost in the downscaling process; From the graphs, it can be concluded that the process of bit depth rescaling for the hyperspectral datacube significantly affects the cross-correlation between hyperspectral data bands, in most cases the Cross-correlation is missed during the process of linear rescaling.

It is not recommended to downscale the bit depth of hyperspectral data cube if further data fusion or spectral profile analysis is needed.

Many researches that used a COTS decoder [1] have to reconsider this issue and indicated whether this process can affect their results or not.

However; COTS codec should be modified -if it is an open source- to handle data with more intensity ranges, i.e., 10-12-16 bit per pixel.

One important issue explains why the using video codec that depend on successive bands similarities, usually gives a non satisfactory results in compressing hyperspectral data.

Losing these similarities-correlation- between bands has a great influence on the performance of the Video codec.

Further studies can be carried out in different types of bit depth rescaling methods such as Gaussian and square root; which may have a significant improvement on inter band Cross-correlation.

REFERENCES

[1] A. Linderheda, N. Wadströmera, K. Stenborga, and H. Nautschb, "Compression of Hyperspectral Data for Automated Analysis" Image and Signal Processing for Remote Sensing Berlin, Germany, Proc. of SPIE Vol. 7477 74770P-1, April 2009.

[2] D. Manolakis R. Lockwood, and T. Cooley "On The Spectral correlation structure of Hyperspectral Imaging Data" Geoscience and Remote Sensing Symposium, 2008, pp. 908-925, doi: 10.1147/rd.523.0275

[3] M. Chiarella, D. Fay, A. Waxman, and R. Bomberger, "Multisensor image fusion and mining: from neural systems to COTS software with application to remote sensing AFE " Advances in Techniques for Analysis of Remotely Sensed Data, IEEE Workshop, 2003, pp. 100 - 107, doi: 10.1109/WARSD.2003.1295180

[4] Y. Kihyun and H. Kim "An adaptive image bit-depth scaling method for image displays " Consumer Electronics, IEEE Transactions February 2010, pp. 141 - 146, doi: 10.1109/TCE.2010.5439137

[5] M. Winken, D. Marpe, H. Schwarz, and T. Wiegand "Bit Depth Scaleble Video coding" Image Processing IEEE International Conference 2007, pp. I - 5 - I - 8 doi : 10.1109/ICIP.2007.4378877

[6] A. Sevcenco and W. Lu "Adaptive Down-Scaling Techniques for JPEG-Based Low Bit-Rate Image Coding " IEE Signal Processing and Information Technology 2006, pp 349 – 354, doi: 10.1109/ISSPIT.2006.270824

[7] J. Lewis, ``Fast Template Matching'', Vision Interface, Image Processing, IEEE Volume: 16 Issue:8 1995, pp. 2139 - 2149, doi: 10.1109/TIP.2007.901243

[8] http://www.mathworks.com/index.htm [retrieved: Jan. 2011]

[9] http://www.exelisvis.com/language/en-us/productsservices/envi.aspx [retrieved: Jan. 2011]

[10] http://compression.jpl.nasa.gov/hyperspectral [retrieved: Dec. 2011]

New Factor Graph Based Multiuser Detector for Spectrally Efficient CPM

Nele Noels and Marc Moeneclaey

TELIN Department

Ghent University

Gent, Belgium

Email: {nnoels, mm}@telin.UGent.be

Abstract—This paper presents a new iterative multiuser detection algorithm for asynchronous spectrally-efficient M-ary continuous-phase modulation in additive white Gaussian noise. This detection algorithm is closely related to another algorithm that was recently proposed by the same authors, but it follows from applying the sum-product algorithm to a different factor graph of the same multiuser detection problem. This, in turn, results in a different way to approximate the marginal bit a-posteriori probabilities that are used to perform minimum bit error rate multiuser detection. The girth of the factor graph considered in this contribution is twice as large, which is known to be potentially beneficial for the accuracy of the a-posteriori probabilities. The size of the largest factor graph variable alphabets also multiplies with M, rendering the straightforward application of the sum-product algorithm more complex. Through approximating a suitable set of sum-product messages by a Gaussian distribution, this complexity is significantly reduced. For a set of system parameters yielding high spectral efficiency, the resulting algorithm significantly outperforms the previously proposed solution.

Keywords— Factor Graph; Multiuser Detection; Spectral Efficiency; Continuous Phase Modulation.

I. Introduction

Continuous-phase modulation (CPM) is a modulation method commonly used in wireless modems [1]. It is attractive because of its high power and spectral efficiency, and because of its robustness to non-linearities. Although the optimal detection [2] of a bit interleaved coded CPM signal is prohibitively complex, there exist approximate iterative detectors with reasonable complexity that yield a very good performance. Such practical detectors can be derived from the sum-product (SP) algorithm and the factor graph (FG) framework [3]. In the past few years, several advanced techniques have been proposed for multiuser (MU) detection of *spectrally efficient* (SE) CPM systems [4], [5], [6]. In such systems, the spacing between the carrier frequencies that are assigned to different users is kept small, such that the leakage of the neighboring signal energy into the desired frequency band cannot be ignored. This leakage signal is referred to as inter-user interference (IUI).

In this paper, we derive a new FG based MU detector for SE CPM. The proposed receiver algorithm will be referred to as g6-MU-FG-GA. Numerical results indicate that, for a set of system parameters yielding high spectral efficiency, the novel g6-MU-FG-GA algorithm significantly outperforms the existing solutions in terms of reliability and/or complexity.

The structure of the paper is as follows. Section II reports on the state-of-the-art in MU bit detection of SE CPM systems. Section III describes the considered system model. The proposed g6-MU-FG-GA detector is derived in Section IV. Its computational complexity and memory requirements are adressed in Section V. Section VI presents numerical results on the packet error rate (PER) performance of the SU-FG, the g10-MU-FG-2IU, the g3-MU-FG-GA and the g6-MU-FG-GA detectors. The main conclusions are summarized in Section VII.

II. State-of-the-art

In [6], an intuitive approach to MU detection of SE CPM is proposed. The considered ad-hoc receiver executes a practical form of MU detection by iterating between conventional FG-based single user (SU) detectors and a separate module for IUI cancellation. A similar solution is also considered as a reference system in [5]. This receiver is further referred to as SU-FG.

A more fundamental approach to MU detection is to consider a FG of the actual MU detection problem. Such a FG is not unique. It represents one particular way to factorize the joint a-posteriori probability (APP) of the bit sequences from all the users, given the observed received MU signal. Running the SP algorithm on this graph yields an approximation of the marginal bit APPs. Using the latter APPs to perform maximum a posteriori (MAP) bit detection results in minimum bit error rate MU detection. Choosing a FG is a matter of trading off the accuracy of the obtained marginal bit APPs against the computational complexity associated with their evaluation. An important parameter for the accuracy of the SP algorithm is the minimum cycle length (i.e., girth) of the graph, while the complexity of the procedure primarily depends on the number of edges and the size of the variable alphabets in the graph.

Optimal MU detection is obtained when the employed FG corresponds to a tree (cycle-less graph). In that case, running the SP algorithm is straightforward and yields the exact marginal bit APPs. Using these APPs to perform MAP bit detection results in minimum bit error rate performance. Unfortunately, the complexity of this optimal MU detector is extremely high and exponential in the number of users [5]. It is therefore not suited for use in practice and one has to resort to approximations.

The practical MU detector proposed in [5] is based on another FG, which contains cycles. The complexity of this receiver is lower than that of the optimal MU detector, but still increases exponentially with the number of users, which becomes impractical as the number of users increases. As an option to further decrease the complexity, the authors propose to apply the simplifying assumption that only the two most adjacent users significantly contribute to the IUI. This results in a FG with girth 10, with a significantly reduced the degree of connectivity. The corresponding receiver will be referred to as g10-MU-FG-2IU. It has a good performance and a linear complexity in the number of users, but it still involves quite a large number of computations and it also requires a considerable amount of memory.

Another low-complexity MU FG-based detection algorithm was proposed in [4]. It results from approximating, in a FG with girth only 3, a suitable set of SP messages by Gaussian distribution functions. The computational complexity and memory requirement of this detector is considerably lower than for the g10-MU-FG-2IU detector from [5]. A significant performance improvement over the SU-FG detector was demonstrated for a particular SE CPM scheme, in [4]. This receiver algorithm will be referred to as g3-MU-FG-GA.

In this paper, we derive a new variant of the algorithm from [4]: g6-MU-FG-GA. The idea of approximating a suitable set of SP messages by a Gaussian distribution is applied to a different FG of the same detection problem. This FG is slightly more complex than the one used to derive the g3-MU-FG-GA detector from [4], but it has cycles of minimum length 6, which is twice as long as the shortest cycles in [4]. This is potentially beneficial for the accuracy of the marginal bit APPs that result from applying the SP algorithm.

III. SYSTEM MODEL

The transmitter of user u encodes N_b information bits $\mathbf{b}^{(u)} = \left\{ b_k^{(u)} \right\}$ into N_c coded bits $\mathbf{c}^{(u)} = \left\{ c_l^{(u)} \right\}$. These coded bits are subsequently interleaved and mapped to N symbols $\mathbf{a}^{(u)} = \left(a_0^{(u)}, ..., a_{N-1}^{(u)} \right)$, with $a_n^{(u)}$ taking values from the M-ary alphabet $\Omega_M = \{\pm 1, \pm 3, ..., \pm (M-1)\}$. The resulting symbols are then used to generate the complex envelope $s_{SU}^{(u)}(t)$, for $0 \le t < NT$, of the CPM signal from user u:

$$s_{SU}^{(u)}(t) = e^{j\psi\left(t; \mathbf{a}^{(u)}\right)}, \tag{1}$$

$$\psi\left(t; \mathbf{a}^{(u)}\right) = 2\pi h \sum_i a_i^{(u)} q(t - iT). \tag{2}$$

Here, T is the symbol period, $h = K/P$ is the modulation index and $q(t)$ is the phase-smoothing response. Taking into account that $q(t) = 0$ for $t \le 0$ and $q(t) = 0.5$ for $t \ge LT$, we can rewrite (2) for $nT \le t \le (n+1)T$ as:

$$\psi\left(t; \mathbf{a}^{(u)}\right) = \Psi\left(t - nT; \mathbf{S}_n^{(u)}\right), \quad nT \le t \le (n+1)T, \tag{3}$$

where $n = 0,1, ..., N - 1$. The quantity $\mathbf{S}_n^{(u)}$ in (3) describes the CPM state transition during the nth symbol interval

$[nT, (n+1)T[$ of the transmitter signal $s_{SU}^{(u)}(t)$ from user u:

$$\mathbf{S}_n^{(u)} = \left(\boldsymbol{\sigma}_n^{(u)}, a_n^{(u)} \right), \tag{4}$$

with $\boldsymbol{\sigma}_n^{(u)}$ an L element vector $\left(\sigma_{n,0}^{(u)}, \sigma_{n,1}^{(u)}, ..., \sigma_{n,L-1}^{(u)} \right)$ denoting the CPM state at time instant n and

$$\Psi\left(\mathtt{t}; \mathbf{S}_n^{(u)}\right) = \sigma_{n,0}^{(u)} + 2\pi h \sum_{i=1}^{L-1} \sigma_{n,i}^{(u)} q\left(\mathtt{t} - (L - i)T\right) \tag{5}$$
$$+ 2\pi h a_n^{(u)} q\left(\mathtt{t}\right), \quad 0 \le \mathtt{t} < T.$$

Given the symbol vector $\mathbf{a}^{(u)}$ and starting from a given initial CPM state $\boldsymbol{\sigma}_0^{(u)}$, the vectors $\boldsymbol{\sigma}_n^{(u)}$, with $n = 1, 2, ..., N$, can be computed recursively according to the following equations:

$$\sigma_{n,0}^{(u)} = \left[\sigma_{n-1,0}^{(u)} + \pi h \sigma_{n-1,1}^{(u)} \right]_{2\pi}, \tag{6}$$

$$\sigma_{n,i}^{(u)} = \sigma_{n-1,i+1}^{(u)}, \quad 1 \le i \le L - 2, \tag{7}$$

$$\sigma_{n,L-1}^{(u)} = a_{n-1}^{(u)}, \tag{8}$$

where $[x]_{2\pi}$ denotes modulo-2 reduction of x to the interval $[0, 2\pi[$. At each time instant n the CPM transmission scheme has PM^{L-1} possible states. The complete set of CPM state vectors $\boldsymbol{\sigma}_n^{(u)}$, $u = 1, 2, ..., U$, is grouped in the vector $\boldsymbol{\sigma}_n = \left(\boldsymbol{\sigma}_n^{(1)}, \boldsymbol{\sigma}_n^{(2)}, ..., \boldsymbol{\sigma}_n^{(U)} \right)$. The complete sequence of CPM state transition identifiers $\mathbf{S}_n^{(u)}$, $n = 0, 1, ..., N - 1$, is grouped in the vector $\mathbf{S}^{(u)} = \left(\mathbf{S}_0^{(u)}, \mathbf{S}_1^{(u)}, ..., \mathbf{S}_{N-1}^{(u)} \right)$.

A group of U users are simultaneously active. They asynchronously transmit frequency-shifted versions of their signals $s_{SU}^{(u)}(t)$, $u = 1, 2, ..., U$ over an additive white Gaussian noise (AWGN) channel, which is typical for satellite communications. The short-hand notations $\mathbf{b}$, $\mathbf{c}$, $\mathbf{a}$, $\boldsymbol{\sigma}$ and $\mathbf{S}$ respectively collect the information bit sequences, the coded bit sequences, the data symbol sequences, the CPM state vector sequences and the CPM state transition identifier sequences from all the users: $\mathbf{b} = \left\{ \mathbf{b}^{(1)}, \mathbf{b}^{(2)}, ..., \mathbf{b}^{(U)} \right\}$, $\mathbf{c} = \left\{ \mathbf{c}^{(1)}, \mathbf{c}^{(2)}, ..., \mathbf{c}^{(U)} \right\}$, $\mathbf{a} = \left\{ \mathbf{a}^{(1)}, \mathbf{a}^{(2)}, ..., \mathbf{a}^{(U)} \right\}$, $\boldsymbol{\sigma} = \left\{ \boldsymbol{\sigma}^{(1)}, \boldsymbol{\sigma}^{(2)}, ..., \boldsymbol{\sigma}^{(U)} \right\}$ and $\mathbf{S} = \left\{ \mathbf{S}^{(1)}, \mathbf{S}^{(2)}, ..., \mathbf{S}^{(U)} \right\}$. The phase offset, frequency shift and time delay associated with user u ($u = 1, 2, ..., U$) are equal to $\theta^{(u)}$, $f^{(u)}/T$ and $\tau^{(u)}T$, with $f^{(u)}$ and $\tau^{(u)}$ denoting the dimensionless relative (to the symbol interval T) frequency shift and time delay, and are assumed to be known at the receiver. We further assume that, within a given group of users, all signals are received with the same power. The complex baseband representation $s(t)$ of the received signal component, which aggregates the contributions of all users from a same group is given by $s(t) = \sum_{u=1}^{U} s^{(u)}(t)$, with

$$s^{(u)}(t) = s_{SU}^{(u)}\left(t - \tau^{(u)}T\right) e^{j2\pi f^{(u)} \frac{t}{T}} e^{j\theta^{(u)}}. \tag{9}$$

The latter contribution differs from zero only for $t \in \left[\tau^{(u)}T, \left(\tau^{(u)} + N \right)T \right[$. It is assumed that $s(t)$ is band-limited (although this is not strictly true in the case of CPM signals whose spectrum has an infinite support) with bandwidth lower than $R_s/2T$, where R_s is a proper integer value (determined by the spectral shape of the considered CPM scheme and the

number of users in a group). The received baseband signal is applied to a low-pass anti-aliasing filter and sampled at R_s samples per symbol interval; the corresponding sampling interval is $T_s = T/R_s$. It is assumed that the spacing between groups of U users is sufficiently large, such that the leakage of the signal energy from neighboring groups into the considered frequency band $\left[\frac{R_s}{2T}, \frac{R_s}{2T}\right]$ can be safely ignored. The resulting samples r_k can then be modeled as follows:

$$r_k = \sum_{u=1}^{U} s_k^{(u)} + n_k, \quad (10)$$

where $s_k^{(u)}$ are samples of the receiver signal $s^{(u)}(t)$ from (9), taken at $t = kT_s$ and n_k are zero-mean complex AWGN samples with variance equal to $N_0 R_s/E_s$, with N_0 the noise power spectral density and E_s the energy per symbol period. The samples r_k are conveniently grouped into vectors of the type $\mathbf{r}_l = \left(r_{lR_s}, r_{lR_s+1}, ..., r_{(l+1)R_s-1}\right)^T$ and $\mathbf{r} = \left(\mathbf{r}_0^T, \mathbf{r}_{R_s}^T, ..., \mathbf{r}_{(N_s-R_s)}^T\right)^T$.

IV. PROPOSED G6-MU-FG-GA RECEIVER: DERIVATION

When an information bit is detected erroneously at the receiver, a bit error occurs. Optimal detection, which minimizes the bit error probability is achieved by performing MAP bit detection [2]:

$$\hat{b}_k^{(u)} = \arg \max_{b \in \{0,1\}} p\left(b_k^{(u)} = b \,|\, \mathbf{r}\right), \forall k. \quad (11)$$

The APPs $p\left(b_k^{(u)} \,|\, \mathbf{r}\right)$ involved in (11) are the marginals of $p(\mathbf{b} \,|\, \mathbf{r})$, with $p(\mathbf{b} \,|\, \mathbf{r})$ the probability of $\mathbf{b}$, given $\mathbf{r}$. An efficient way to jointly compute these marginals is to apply the SP algorithm to a FG representing a suitable factorization of the joint probability $p(\mathbf{b}, \mathbf{x} \,|\, \mathbf{r})$ of $\mathbf{b}$ and $\mathbf{x}$, where $\mathbf{x}$ represents a convenient set of additional variables [3].

In the following, we derive the proposed g6-MU-FG-GA detector.

We construct a FG for deriving $p\left(b_k^{(u)} \,|\, \mathbf{r}\right)$ by employing $(\mathbf{c}, \mathbf{a}, \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_N, \mathbf{S}, \mathbf{s})$ as additional variables. The vector $\mathbf{s}$ results from stacking the U vectors $\mathbf{s}^{(u)}$, $u = 1, 2, ..., U$, where $\mathbf{s}^{(u)}$ consists of all the samples (both zero and non-zero valued) of $s^{(u)}(t)$ from (9), i.e., taken at instants kT_s. The vector $\mathbf{s}^{(u)}$ itself results from stacking the vectors $\mathbf{s}_x^{(u)}$, where $\mathbf{s}_x^{(u)}$ contains the R_s samples $s^{(u)}(iT_s + xT)$, $i = 0, 1, ..., R_s - 1$, that correspond to the xth symbol interval $[xT, (x+1)T[$ of the receiver signal $s^{(u)}(t)$ from (9). In general, these samples are spread over the $\left(x - N_\tau^{(u)} - 1\right)$th $\left[\left(x - N_\tau^{(u)} - 1\right)T, \left(x - N_\tau^{(u)}\right)T\right[$ and the $\left(x - N_\tau^{(u)}\right)$th $\left[\left(x - N_\tau^{(u)}\right)T, \left(x - N_\tau^{(u)} + 1\right)T\right[$ symbol interval of the signal $s_{SU}^{(u)}(t)$ from (1), with $N_\tau^{(u)} = \lfloor \tau^{(u)} \rfloor$, being the smallest integer value smaller than or equal to $\tau^{(u)}$. The joint APP $p(\mathbf{b}, \mathbf{c}, \mathbf{a}, \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_N, \mathbf{S}, \mathbf{s} \,|\, \mathbf{r})$ can be factorized as follows, by taking into account the specific structure of the transmitted
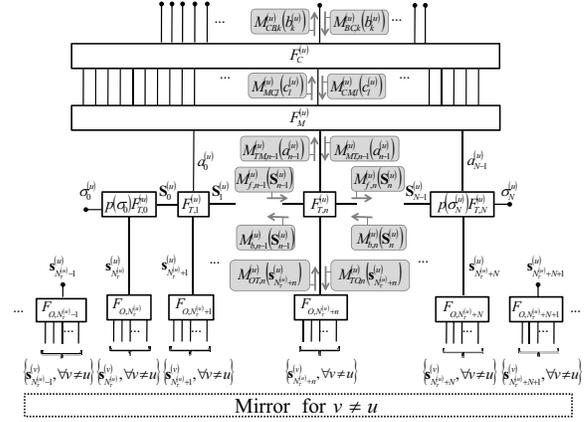
Figure 1. FG representing the factorization of $p(\mathbf{b} \,|\, \mathbf{r})$ from (12).

CPM signal:

$$p(\mathbf{b}, \mathbf{c}, \mathbf{a}, \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_N, \mathbf{S}, \mathbf{s} \,|\, \mathbf{r}) \quad (12)$$

$$\propto \prod_x F_{O,x}(\mathbf{s}_x) \prod_{u=1}^{U} F_C^{(u)}\left(\mathbf{c}^{(u)}, \mathbf{b}^{(u)}\right) F_M^{(u)}\left(\mathbf{a}^{(u)}, \mathbf{c}^{(u)}\right)$$

$$\cdot p\left(\boldsymbol{\sigma}_0^{(u)}\right) p\left(\boldsymbol{\sigma}_N^{(u)}\right) F_T^{(u)}\left(\boldsymbol{\sigma}_0^{(u)}, \boldsymbol{\sigma}_N^{(u)}, \mathbf{S}^{(u)}, \mathbf{s}^{(u)}\right),$$

with

$$F_T^{(u)}\left(\boldsymbol{\sigma}_0^{(u)}, \boldsymbol{\sigma}_N^{(u)}, \mathbf{S}^{(u)}, \mathbf{s}^{(u)}\right) = F_{T,0}^{(u)}\left(\boldsymbol{\sigma}_0^{(u)}, \mathbf{S}_0^{(u)}, \mathbf{s}_{N_\tau^{(u)}}^{(u)}\right) \quad (13)$$

$$\cdot \prod_{n=1}^{N-1} F_{T,n}^{(u)}\left(\mathbf{S}_{n-1}^{(u)}, a_{n-1}^{(u)}, \mathbf{S}_n^{(u)}, \mathbf{s}_{N_\tau^{(u)}+n}^{(u)}\right)$$

$$\cdot F_{T,N}^{(u)}\left(\mathbf{S}_{N-1}^{(u)}, a_{N-1}^{(u)}, \boldsymbol{\sigma}_N^{(u)}, \mathbf{s}_{N_\tau^{(u)}+N}^{(u)}\right)$$

and

$$F_{O,x}(\mathbf{s}_x) = e^{\frac{E_s}{N_0 R_s}\left[2\sum_{u=1}^{U}\Re\left\{\mathbf{r}_x^H \mathbf{s}_x^{(u)}\right\} - \sum_{k=0}^{R_s-1}\left|\sum_{u=1}^{U} s_{xR_s+k}^{(u)}\right|^2\right]}. \quad (14)$$

Again $F_C^{(u)}\left(\mathbf{c}^{(u)}, \mathbf{b}^{(u)}\right) = p\left(\mathbf{c}^{(u)} \,|\, \mathbf{b}^{(u)}\right)$, $F_M^{(u)}\left(\mathbf{a}^{(u)}, \mathbf{c}^{(u)}\right) = p\left(\mathbf{a}^{(u)} \,|\, \mathbf{c}^{(u)}\right)$ and $F_T^{(u)}\left(\boldsymbol{\sigma}_0^{(u)}, \boldsymbol{\sigma}_N^{(u)}, \mathbf{S}^{(u)}, \mathbf{s}^{(u)}\right)$ impose the constraints from the encoder, the mapper and the CPM state transitions. The factors $F_{T,n}^{(u)}(\cdot)$ in (13) equal one when their arguments satisfy equations (1)-(8) and (9), and zero otherwise. The notation $\mathbf{s}_x$ is used to denote the set of vectors $\mathbf{s}_x^{(u)}$, $u = 1, 2, ..., U$. This set contains the contributions from all the users from a same group to the samples of the received signal $r(t)$ taken in a given symbol interval $[xT, (x+1)T[$.

The FG representing the function factorization (12) is shown in Figure 1. Only the part related to user u is detailed. The upper three rows of nodes need to be repeated for every user $v \neq u$ and suitably connected to the observation nodes $F_{O,x}$, in order to obtain the complete graph. The cycles with minimum length are between the trellis constraint nodes of two interfering users; they are of length 6. This is relatively short, but twice as long as for the FG employed to derive the low-complexity MU detector from [4].

Algorithm 1 SP initialization and scheduling strategy.

Initialization. $i = 0$; $\left\{ M_{CM,l}^{(u)} \right\} = \left\{ M_{MT,n}^{(u)} \right\} = \left\{ M_{TO,n}^{(u)} \right\} = 1$.

Iterative procedure. For $i = 1, 2, ... i_{max}$:

 -**Multiuser processing.** Compute messages $\left\{ M_{OT,n}^{(u)} \right\}$.

 -**Per user processing.** For $u = 1, 2, ..., U$, compute

- forward/backward messages $\left\{ M_{f,n}^{(u)} \right\}$ and $\left\{ M_{b,n}^{(u)} \right\}$.
- upward messages $\left\{ M_{TM,n}^{(u)} \right\}$, $\left\{ M_{MC,l}^{(u)} \right\}$ and $\left\{ M_{CB,k}^{(u)} \right\}$.
- downward messages $\left\{ M_{CM,l}^{(u)} \right\}$, $\left\{ M_{MT,n}^{(u)} \right\}$, $\left\{ M_{TO,n}^{(u)} \right\}$.

 -The messages $\left\{ M_{CB,k}^{(u)} \right\}$ provide an estimate of the marginal information bit APPs. Use these estimates to compute $\left\{ \hat{b}_k^{(u)} \right\}$ using (11). **If** all detected bits are error-free: end iterations.

The notation for the SP messages traveling along the FG edges is also introduced in the figure. The rules for computing the FG messages are the SP rules from [3], followed by a normalization step such that all messages communicated along the edges of the FG can be interpreted as probability mass functions. Because the graph contains cycles (paths from a node to itself), the SP algorithm is an iterative procedure that, after convergence, yields only an approximation of the APPs $p\left(b_k^{(u)} | \mathbf{r} \right)$. The initialization and message passing scheduling strategy is outlined in Algorithm 1. In most practical scenarios and without further approximations, the multiuser processing step in Algorithm 1 is prohibitively complex. This follows directly from applying the SP algorithm and it is an inevitable consequence of the large amount of variable edges entering the observation function node and the large size of the associated variable alphabet. It follows from (1)-(4), (6)-(8) and (9) that, for a given value of $\left(f^{(u)}, \theta^{(u)}, \tau^{(u)} \right)$ and for $x = N_\tau^{(u)} + 1, N_\tau^{(u)} + 2, ..., N_\tau^{(u)} + N - 1$, $\mathbf{s}_x^{(u)}$ is fully determined by the consecutive state transitions $\mathbf{S}_{x-N_\tau^{(u)}-1}^{(u)}$ and $\mathbf{S}_{x-N_\tau^{(u)}}^{(u)}$, with $\mathbf{S}_{x-N_\tau^{(u)}}^{(u)}$ itself fully determined by $\mathbf{S}_{x-N_\tau^{(u)}-1}^{(u)}$ and $a_{x-N_\tau^{(u)}}^{(u)}$ (see (4) and (6)-(8)). The vectors $\mathbf{s}_x^{(u)}$, $x = N_\tau^{(u)} + 1, N_\tau^{(u)} + 2, ..., N_\tau^{(u)} + N - 1$, can, therefore, take on $P M^{L+1}$ different values. It is further easily verified that

- $\mathbf{s}_x^{(u)} \equiv \mathbf{0}_{R_s \times 1}$, if $x < N_\tau^{(u)}$ or $x > N + N_\tau^{(u)}$.
- $\mathbf{s}_{N_\tau^{(u)}}^{(u)}$ ($\mathbf{s}_{N+N_\tau^{(u)}}^{(u)}$) is fully determined by the state transition $\mathbf{S}_0^{(u)}$ ($\mathbf{S}_{N-1}^{(u)}$), and, therefore, the vector $\mathbf{s}_{N_\tau^{(u)}}^{(u)}$ ($\mathbf{s}_{N+N_\tau^{(u)}}^{(u)}$) can take on only $P M^L$ different values.

A common approach to simplify the SP algorithm is to approximate messages by a canonical distribution. We will apply this approach to the FG from Figure 1, which will lead to a novel receiver structure. Based on the discussion in [7], we propose to approximate the messages $\left\{ M_{TO,n}^{(u)} \right\}$ in Figure 1 by the product of R_s univariate Gaussian distribution functions of complex-valued circularly symmetric random variables with

means $\left\{ \mu_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} \right\}$ and variances $\left\{ v_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} \right\}$:

$$M_{TO,n}^{(u)} \left(\mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right) \approx e^{ - \sum_{k=0}^{R_s-1} \frac{\left| s_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} - \mu_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} \right|^2}{v_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)}} } \tag{15}$$

with

$$\mu_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} \tag{16}$$
$$= \sum_{\mathbf{s}_{N_\tau^{(u)}+n}^{(u)}} s_{\left(N_\tau^{(u)} + n \right) R_s + k}^{(u)} M_{TO,n}^{(u)} \left(\mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right),$$

$k = 0, 1, ..., R_s - 1$, and

$$v_l^{(u)} = \begin{cases} 1 - \left| \mu_l^{(u)} \right|^2 & , \tau^{(u)} R_s \le l < \left(\tau^{(u)} + N \right) R_s \\ 0 & , \text{otherwise} \end{cases} \tag{17}$$

The approximation (15) significantly simplifies the multiuser processing step of Algorithm 1. Exploiting the property that the sum of independent Gaussian variables (signal samples from different users $v \ne u$) is Gaussian with mean equal to the sum of the means and variance equal to the sum of the variances, the SP messages $M_{OT,n}^{(u)}$ are easily obtained in closed-form. Applying the SP rules from [3], we find:

$$M_{OT,n}^{(u)} \left(\mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right) \propto e^{ 2\Re \left\{ \left(\mathbf{y}_{N_\tau^{(u)}+n}^{(u)} \right)^H \mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right\} }, \tag{18}$$

where $\mathbf{y}_x^{(u)}$ is a size-R_s column vector with components $y_q^{(u)}$, for $q = xR_s, xR_s + 1, ..., xR_s + R_s - 1$, given by:

$$y_q^{(u)} = \frac{r_q - \left(\mu_{MU,q} - \mu_q^{(u)} \right)}{\frac{N_0}{E_s} R_s + \left(v_{MU,q} - v_q^{(u)} \right)}, \tag{19}$$

where $\mu_{MU,q} = \sum_{u=1}^{U} \mu_q^{(u)}$ and $v_{MU,q} = \sum_{u=1}^{U} v_q^{(u)}$. Expression (19) indicates that a soft interference estimate $\left(\mu_{MU,q} - \mu_q^{(u)} \right)$ is subtracted from the observation r_q, and the estimation error variance $\left(v_{MU,q} - v_q^{(u)} \right)$ is added to the noise variance $N_0 R_s / E_s$. The particular structure of (18) indicates that the proposed multiuser detector can be decomposed into an equivalent set of U iterative single user detectors, with a separate module for IUI parameter estimation (means and variances). The single user detectors are operated in parallel, with the uth detector accepting the samples $\left\{ y_q^{(u)} \right\}$ from (19) as equivalent observations. At each iteration, new values for the messages $\left\{ M_{TO,n}^{(u)} \right\}$ are computed at the end of the per user processing step in Algorithm 1, according to [3]:

$$M_{TO,n}^{(u)} \left(\mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right) \tag{20}$$
$$= M_{f,n-1}^{(u)} \left(\dot{\mathbf{S}}_{n-1}^{(u)} \right) M_{MT,n-1}^{(u)} \left(\dot{a}_{n-1}^{(u)} \right) M_{b,n}^{(u)} \left(\dot{\mathbf{S}}_n^{(u)} \right)$$
$$\cdot F_{T,n}^{(u)} \left(\dot{\mathbf{S}}_{n-1}^{(u)}, \dot{a}_{n-1}^{(u)}, \dot{\mathbf{S}}_n^{(u)}, \mathbf{s}_{N_\tau^{(u)}+n}^{(u)} \right),$$

with $\left(\dot{\mathbf{S}}_{n-1}^{(u)}, \dot{a}_{n-1}^{(u)}, \dot{\mathbf{S}}_{n}^{(u)}\right)$ the only possible value of $\left(\mathbf{S}_{n-1}^{(u)}, a_{n-1}^{(u)}, \mathbf{S}_{n}^{(u)}\right)$ for which $F_{T,n}^{(u)}\left(\mathbf{S}_{n-1}^{(u)}, a_{n-1}^{(u)}, \mathbf{S}_{n}^{(u)}, \mathbf{s}_{N_{\tau}^{(u)}+n}^{(u)}\right)$ differs from zero. These messages are subsequently used to update the IUI parameters $\left\{\mu_{q}^{(u)}\right\}$ and $\left\{v_{q}^{(u)}\right\}$, which in turn are used to update the equivalent observation samples $\left\{y_{q}^{(u)}\right\}$. The messages $\left\{M_{OT,n}^{(u)}\right\}$ can then efficiently be obtained from (18), using (1)-(3), (9).

V. PROPOSED G6-MU-FG-GA RECEIVER: COMPLEXITY

The proposed g6-MU-FG-GA algorithm involves the same steps as the g3-MU-FG-GA detection algorithm from [4], but the execution of these steps it is slightly more complex. In particular, the messages $M_{f,n}^{(u)}$, $M_{b,n}^{(u)}$, $M_{TO,n}^{(u)}$ and $M_{OT,n}^{(u)}$ computed in the g6-MU-FG-GA detector take M times more values than the corresponding messages computed in the g3-MU-FG-GA detector from [4]. All these values have to be re-computed and stored at each iteration. The number of operations that is required to update the messages $M_{TM,n}$ according to the SP rules and the number of additions that is needed to evaluate the means $\mu_{q}^{(u)}$ from (16) is also M times larger for g6-MU-FG-GA than for g3-MU-FG-GA.

The SU-FG detection algorithm from [5], [6] has about the same computational load and memory requirement as the g3-MU-FG-GA detector. The structure of the g10-MU-FG-2IU detector from [5] differs substantially from that of the proposed detectors g3-MU-FG-GA and g6-MU-FG-GA. Per user, per iteration and per symbol period, 2 additional messages have to be evaluated and stored. Both these additional messages take as much values as the messages $M_{OT,n}^{(u)}$ in the proposed g6-MU-FG-GA detector. Besides, the number of operations involved in the computation of these additional messages is large and contributes significantly to the overall complexity of the g10-MU-FG-2IU detector.

Table I compares the memory requirements (expressed in number of real values to be stored, MEM) and the computational complexity (expressed in the number of elementary operation between two real arguments, OP) of the proposed detection algorithm to those of the g3-MU-FG-GA and the g10-MU-FG-2IU detector. Only the contributions to the required memory and the number of operations that result from parts in which the g6-MU-FG-GA, the g3-MU-FG-GA and the g10-MU-FG-2IU detector differ from each other are taken into account. These contributions provide a solid basis for comparing the considered algorithms because they dominate the detector's total memory requirements and total computational complexity, respectively. Operations that are executed only once, at the start of the iterative process, are also not taken into account. General closed form expressions are provided, as well as numerical results for the simulation set-up in Section VI. For the g10-MU-FG-2IU detector, we distinguish between memory that needs to be allocated dynamically (first term) and static memory (second term). For the g6-MU-FG-GA and the g3-MU-FG-GA detectors, the amount of static memory that is needed is negligibly small as compared to the dynamic memory resources they require, and therefore only the latter is reported.

The g10-MU-FG-2IU detector requires a considerably larger amount of memory than the proposed g6-MU-FG-GA detector, which in turn requires about M times as much memory as the g3-MU-FG-GA detector. The complexity of the g6-MU-FG-GA detector is also significantly less complex than the g10-MU-FG-2IU detector. As opposed to the complexity of the g10-MU-FG-2IU detector, the complexity of the proposed g6-MU-FG-GA detector increases less than proportional with the number of users in a group. Finally, the complexity of the g6-MU-FG-GA detector is about M times as large as that of the g3-MU-FG-GA detector.

A fair complexity comparison of the different algorithms must also take into account the number of iterations that actually needs to be performed by the receiver (in order to meet some given performance specifications). The required number of iterations for the different algorithms will be considered in Section VI.

VI. NUMERICAL PERFORMANCE RESULTS

Computer simulations have been run to evaluate the performance of the proposed g6-MU-FG-GA detector. For comparison, the performance of the simple ad-hoc SU-FG detector (also used as a reference system in [5]), the performance of the g3-MU-FG-GA algorithm from [4] and the performance of the overly complex graph-based g10-MU-FG-2IU receiver from [5] are evaluated as well. We consider one of the MU BIC-CPM schemes with the highest asymptotic (for $U \rightarrow \infty$) spectral efficiency reported in [8]. Each user asynchronously transmits an information bit vector of size 1024. Gray mapping and pseudo-random bit interleaving are employed. The outer code is a (128,115) extended BCH code. The CPM parameters are $M = 4$, $L = 2$, $h = 1/3$ and $q(t) = \frac{t}{2LT}\left(1 - \cos\left(\frac{2\pi t}{LT}\right)\right)$, $t \in [0, LT]$ (raised-cosine pulse shaping). Equal normalized nominal frequency spacings of 0.8 are used such that $f^{(u)} = 0.8\left(u - \frac{U+1}{2}\right)$. Groups of $U = 17$ active users are considered. In each simulation new normalized time delays $\tau^{(u)}$, $u = 1, 2, ..., 17$, are taken independently from a random uniform distribution over $[0, 8.5]$. The influence of adjacent user groups is neglected. The received signal is sampled $R_s = 16$ times per symbol period. In this case, the complexity (per user, per iteration and per transmitted symbol) of the proposed g6-MU-FG-GA detector is about 4 times as large as that of the g3-MU-FG-GA detector from [4], but only about 1/8 times as large as that of the g10-MU-FG-2IU detector from [5]. At every iteration, hard decisions about the information bits are made from the corresponding a posteriori information bit probabilities, after which a genie checks for bit errors; the receiver stops iterating after a maximum number of iterations $i_{\max}$, or when all information bits have been detected correctly. Figure 2 shows the PER of the middle user ($u = 9$) for several values of $i_{\max}$, at a given value of E_b/N_0, with

	g3-MU-FG-GA	g6-MU-FG-GA	g10-MU-FG-2IU
MEM	$2NUPM^{L-1}(1+M)$	$2NUPM^{L}(1+M)$	$[2NUPM^{L}(1+3M)] +$ $[2\left(R_s PM^{(L+1)}\right)^2]$
MEM	1 197 480	4 789 920	14 253 792 + 18 874 368
OP	$15R_s + 8R_s PM^{L} + 13PM^{L}$	$15R_s + 8R_s PM^{L+1} + 13PM^{L+1}$	$13PM^{L+1} + 8P^2 M^{2(L+1)}$
OP	7008	27312	297408

Table I
MEM: Number of real values to be stored, assuming the transmission of 1024 information bits. OP: number of elementary operations between two real arguments, per user, per iteration and per transmitted symbol.
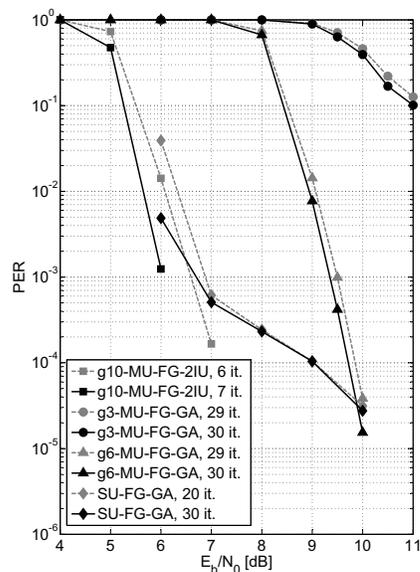
Figure 2.   PER versus E_b/N_0.

$E_b = NE_s/N_b$ (with 1 packet containing 1024 information bits).

Values of $i_{\max}$ up to 30 are considered for SU-FG, g3-MU-FG-GA and g6-MU-FG-GA. The latter two algorithms are reasonably close to convergence after 29 iterations. The SU-FG detector converges faster; for E_b/N_0 above 6.5 dB, the PER of this algorithm hardly decreases between the 20th and the 30th iteration. For this particular simulation set-up, the proposed g6-MU-FG-GA performs much better than g3-MU-FG-GA from [4]. The gain in terms of PER amounts to several dB. The g6-MU-FG-GA also outperforms the ad-hoc SU-FG for PERs smaller than 3.10^{-5}. Above $PER = 3.10^{-5}$, SU-FG yields a better performance. Values of $i_{\max}$ up to 7 are considered for g10-MU-FG-2IU. This is too little to reach convergence. This overly complex algorithm results in a lower PER than g6-MU-FG-GA and SU-FG, after significantly less iterations. However, 5 or more iterations of g10-MU-FG-2IU require considerably more computation time than 20 iterations of SU-FG or 30 iterations of g6-MU-FG-GA.

VII. Conclusion

In this paper, we have derived a new MU detection procedure for asynchronous spectrally efficient CPM. In practice, the technique employs simple IUI cancellation, yet its derivation from the SP algorithm and a FG of the MU detection problem with girth 6 is theoretically sound. It is characterized by a relatively low computational complexity. For a set of system parameters yielding very high spectral efficiency, it is shown to outperform the existing solutions either in terms of PER performance (SU-FG from [6] and g3-MU-FG-GA from [4]) or computational complexity (g10-MU-FG-2IU from [5]). Overall, the proposed detection algorithm can be a valuable alternative to more complex algorithms for systems where the main concerns are the computational complexity and the memory requirements.

Acknowledgment

References

[1] T. Aulin, N. Rydbeck and C.W. Sundberg, "Continuous Phase Modulation", IEEE Trans. Commun., vol. 29, no. 3, Mar. 1981, pp. 196-225.

[2] H. Meyr, M. Moeneclaey and S. Fechtel, "Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing," Wiley Series in Telecommunications and Signal Processing, JohnWiley & Sons, NY, NY, USA, 1998.

[3] A.H. Loeliger, "An introduction to Factor Graphs," IEEE Sig. Proc. Mag., vol. 21, no. 1, Jan. 2004, pp. 28-41.

[4] N. Noels and M. Moeneclaey, "Spectrally Efficient CPM: Suboptimal FG-based Multiuser Detection," accepted for publication in Proc. of IEEE Wireless Communications and Networking Conference (WCNC12), Apr. 1-4 2012, Paris, France.

[5] A. Piemontese and G. Colavolpe, "A novel Graph-Based Suboptimal Multiuser Detector for FDM-CPM Transmissions," IEEE Trans. Wireless Commun., vol. 9, no. 9, Sep. 2010, pp. 2812-2819.

[6] A. Perotti, S. Benedetto and P. Remlein, "Spectrally efficient multiuser continuous-phase modulation systems," Proc. IEEE International Conference on Communications (ICC10), 23-27 May 2010, Cape Town, South-Africa, pp.1-5.

[7] J. Boutros and G. Caire, "Iterative Multiuser Joint Decoding: Unified Framework and Asymptotic Analysis," IEEE Trans. Inform. Th., vol. 48, no. 7, Jul. 2002, pp. 1772-1793.

[8] A. Piemontese, A. Grael i Amat and G. Colavolpe, "Information-Theoretic Analysis and Practical Coding Schemes for Spectrally Efficient FDM-CPM Systems," Proc. 2010 6th International Symposium on Turbo Codes and Iterative Information Processing (ISTC), 6-10 Sep. 2010, Brest, France, pp. 275-279.

Dependency of SAR Image Structure Descriptors with Incidence Angle

Corneliu Octavian Dumitru

Remote Sensing Technology Institute (IMF)
German Aerospace Center (DLR)
Oberpfaffenhofen - Wessling, Germany
corneliu.dumitru@dlr.de

Mihai Datcu

Remote Sensing Technology Institute (IMF)
German Aerospace Center (DLR)
Oberpfaffenhofen - Wessling, Germany
mihai.datcu@dlr.de

Abstract— **The interpretation of the structure in Synthetic Aperture Radar images depends by the used parameters and incidence angle. The evaluation is done on the high resolution SAR data and the interpretation is realized automatically. In this paper, we propose to study and asses the behavior of the primitive feature extracted methods for images of the same scene with 2-3 look angles covering the min-max range of the sensor. The tests are done on TerraSAR-X products High Resolution Spotlight mode at 3 m resolution and two sites were found that are appropriate for this. To identify the best features and appropriate incidence angle for them the Support Vector Machine and as a measure of the classification accuracy the precision –recall were considered. The precision-recall was computed first for all investigated features and after that the best were taken into account for the incidence angle evaluation.**

Keywords - TerraSAR-X products; inicidence angle; patch; features; semantic; classification; precision-recall.

I. INTRODUCTION

The specific information in High Resolution (HR) SAR (Synthetic Aperture Radar) images acquired in single polarization is mainly in the "structure", e.g. textures, objects, or scattering signatures. The "spatial context" becomes very important rather then the "pixel based" descriptors which are less informational. The adopted solution is to analyze image patches corresponding to ground areas of ca. 200x200m. Experiments and tests carried recently confirmed the usefulness of the concept, however further analysis is needed to asses the behavior of the method for the indexing of very large SAR data sets as the case in Image Information Mining (IIM).

There are few publications available [1] ÷ [5] where the images are tile into patches. In [1], the patch size is 256x256 m in order to ensure that the extracted information capture the local characteristics within a patch rather the global features across the entire image.

In [2], the TerraSAR-X (TSX) HR Spotlight products (resolution of ~1 m) were tiled into patches of 200x200m in order to characterize the large and relatively small structures available in the urban scene. The images covered different region: Las Vegas, Venice, Gizah, and Gauting.

In [3], the original images are tiled into patches of 16x16 pixels or 128x128 pixels. The results of the classification (city, forest, and sea) were better for the patch size of 128x128 pixels. The same authors propose in [4] a patch

contextual approach for HR satellite images (resolution of 0.6 m) where the patch size is 200x200 pixels.

In our previous work [5], a pyramid with different resolutions (1m, 2m, 2.9m, 4m, and 8m) was considered for TSX HR Spotlight where each image was tiled into patches at different size in order to have the same area cover on the ground. The patch sizes vary from 400x400m (for 1m resolution) to 25x25m (for 8m resolution).

The paper structure is the following. Section 2 presents the TSX products used for tests, while Section 3 explains the actual state-of-the-art of the feature extraction methods and shortly describe the applied methods. Section 4 provides the details about the experiments and points the conclusion. The references end the paper.

II. TERRASAR-X PRODUCTS

TerraSAR-X is the German radar satellite launched on June 2007. It operates in the X-band and is a side-looking SAR based on active phased array antenna technology. It does supply high quality radar data for purposes of scientific observation of the Earth [6].

The basic products are available in a huge diversity of modes (Stripmap, Spotlight, ScanSAR), types (complex, detected, geocoded), and configurations (Spatially Enhanced Products or Radiometrically Enhanced Products) [6].

For our investigation, we considered TSX products, geocoded product, high resolution spotlight mode, and radiometrically enhanced. Two sites are downloaded from the TSX EOWEB portal [7], one covering the Berlin area and the second one the Ottawa area. For these two sites the parameters extracted from the metadata of each product are: Berlin-the ground range resolution is about 2.9m, the orbit direction with ascending looking, and the incidence angles are 30° and 42°, and Ottawa-the ground range resolution is similar, but the orbit direction is descending and the incidence angles are 27°, respectively 41°.

The number of looks depends by the incidence angle and varies from 5 for an incidence angle of 20° to 9 for an incidence angle of 55°.

III. FEATURES EXTRACTION METHODS

Many feature extraction methods have been proposed in the past several decades but few authors are compare these feature for satellite images.

On a conceptual level we decide which features can be extracted in general and on a practical level, we apply the: gray level co-occurrence feature extraction [33] for texture

analysis, Gabor filtering [18] to extract any geometrical or neighbourhood relationships, quadrature mirror filters [30] for texture analysis, and non-linear short time Fourier transform [32] for spectral characteristics of the image.

We can divide the features in two categories: statistical and spectral.

A. Statistical

1) Gray level co-occurrence matrix

a) State of the art

The gray level co-occurrence matrix (GLCM) is a second order statics of how often different combinations of pixel brightness values (gray levels) occur in an image [13].

Haralick et al. [8] compute gray level co-occurrence matrix for a distance of one with four directions (0°, 45°, 90°, and 135°). For a seven-class classification problem, they obtained approximately 80% classification accuracy using texture features in remote sensing images application.

Rignot and Kwok [9] have analysed SAR images using texture features computed from GLCM. However, they supplement these features with knowledge about the properties of SAR images. For example, image restoration algorithms were used to eliminate the specular noise present in SAR images in order to improve classification results.

Schistad and Jain [10] compare different methods for texture computation in ERS SAR imagery. One of the used and computed methods was GLCM with four directions like in [8]. The *angular second moment, contrast, entropy, cluster shade, inertia, and inverse difference moment* [13] were computed as texture features from the GLCM. A five class classification problem was considered and 29% (an average) classification error using GLCM was obtained.

Randen and Husoy [11] consider GLCM as a reference method and they compared this with other filtering methods (like: QMF, Gabor, discrete cosine transform, etc) for texture extraction. The size of the gray levels in the image is 8 x 8 (also chosen by Ohanian and Dubes [12]). On the one hand, if the value is large, the number of pixel pairs contributing to each element in image will be low, and the statistical significance poor. On the other hand, if the gray levels are low, much of the texture information may be lost in the image quantization. The *angular second moment, contrast, correlation, and entropy* were computed as texture features for each orientation. The average of the classification error was 32%.

b) Applied method

The GLCM is created from a gray scale image by selecting either horizontal (0°), vertical (90°), or diagonal (45° or 135°) orientation.

The size of GLCM depends on the number of gray values available in the image. For example, in [29], they obtain for an input image of 8 bits, i.e., 256 values, a GLCM of 256x256 elements.

In our case, we scale the radiometric range of the input images to 16 steps and obtain a GLCM size of 16x16 elements.

The texture parameters [33] computed from the GLCM are: *mean, variance, entropy, contrast, energy, correlation,* *homogeneity, autocorrelation, dissimilarity, cluster shade, cluster prominence,* and *maximum probability*.

B. Spectral

1) Gabor filters

a) State of the art

A Gabor filter (GAFS) is a linear filter used in image processing.

Randen and Husoy [11] review the major filtering approaches to texture feature extraction and performed a comparative study by comparing with two classical non-filtering approaches (GLCM which is a statistical method and autoregressive which is model based method). The dyadic Gabor filter bank (i.e. Gaussian shaped band-pass filters, with dyadic coverage of the radial spatial frequency range and multiple orientations) proposed by Jain and Farrokhnia [14] was considered for the experiments in [11]. Five radial frequency were used (proposed by [14] for images of size 256 x 256 pixels) and four orientations (0°, 45°, 90°, and 135°). The average error on the classification was 31%.

Du [15] used texture features derived from Gabor filters to segment SAR images. He successfully segmented the SAR images into categories of water, new forming ice, older ice, and multi-year ice. Lee and Philpot [16] also used spectral texture features to segment SAR images.

Shu et al. [17] extract the information at four directions (0°, 45°, 90°, and 135°) by using Gabor filters and then computing the mutual information of each corresponding image pair. The experiments show that the method can work very well even if the SAR image is not filtered; this indicates that the method is robust to speckle noise.

In Manjunath and Ma [18] a Gabor wavelet based texture analysis method is proposed and its application to image databases is demonstrated on Brodatz texture database but also considering the current work related to the idea of browsing large satellite images database. The experiments results demonstrate that these Gabor features are robust. Rotation and scale invariance are important in many applications and the preliminary results obtained by [18] using Gabor features are very promising.

In [19] ÷ [22], the Gabor filters are applied to Brodatz texture database with very good results.

b) Applied method

Frequency and orientation representations of a Gabor filter are similar to those of the human visual system, and it has been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave [18]. The Gabor filters are self-similar - all filters can be generated from one mother wavelet by dilation and rotation.

We have chosen the Gabor filters designed by Manjunath and Ma at Vision Research Lab, University of California.

The texture parameter results computed from the Gabor filter are *mean* and *variance* for different *scales* and *orientations*.

2) Quadrature mirror filters

a) State of the art

Quadrature Mirror Filter (QMF) banks are multirate (i.e. with variable sampling rate throughout the system) digital filter banks, introduced by Croisier, [23], Esteban and Galand [24]. During the last two decades since the inception of QMF banks, they have been extensively used in speech signal processing, image processing and digital transmultiplexers [25]. QMF banks are used to split a discrete-time signal into a number of bands in the frequency domain to process each sub-band in independent manner.

QMF was used for texture analysis by Randen and Husoy [11] as extended classes of filters which include among others Gabor filters, discrete cosine transform, etc. This is a large class of filters which incorporate both infinite impulse response (IIR) and finite impulse response (FIR) filters. In their experiments the average of the classification error was between 26% and 33%.

b) Appplied method

As proposed in [30], statistical features obtained from the filtered images using QMF banks in synergy with some other features can be used for image (satellite image) indexing.

The number of features which can be obtained from the presented algorithm depends upon the level selected for the QMF sub-band decomposition like a wavelet. Features are nothing but the mean and variance of the four filtered and sub-sampled images in the QMF sub-band pyramid.

There are many techniques available to design QMF banks. We have chosen the QMF banks designed by Simoncelli and Adelson at the Vision Science Group, The Media Laboratory, MIT [34].

The parameters computed from the QMF banks (QMFS) are *mean* and *variance* of the *low pass sub-band*, *horizontal sub-band*, *vertical sub-band*, and *diagonal sub-band*.

3) Non-linear short time Fourier transform

a) State of the at

Much work on extraction of features based on short time Fourier transform is done in speech and audio processing.

The method proposed in [26] was investigated by Li and Ogihara [32] for music information retrieval. They are using short time Fourier transform feature extraction method to extract the timbral texture witch is not capture by the popular method in speech and music processing, the Mel-frequency cepstral coefficients. The derived features computed from STFT are: spectral centroid, spectral Rolloff, spectral flux, low energy, and zero crossings.

The goal of Popescu et al. paper [26] is to define an analysis model for High Resolution Spotlight SAR imagery, which is able to integrate the radiometric, as well as geometric and texture properties of the SAR data, in order to facilitate large data-base queries by informational content indexing of the images. The proposed model use the information contained in the spectra of the SAR signal.

The Short Time Fourier Transform (STFT) was considered in order to extract the features necessary for the Bayesian Support Vector Machine classifier. The features are: spectral centroid, spectral flux, cepstral coefficients, and first and second statistic measures. Using this method a

number of 30 classes were recognized from the 9,000 patches of SAR images acquired with TerraSAR-X satellite.

b) Appplied method

This method of SAR image feature extraction and complex image information retrieval was first proposed in [31]. This non-parametric analysis is a form of time frequency analysis where the cutting of a spectrum allows the study of the phase responses of scatterers seen from different viewing angles.

The STFT extracts six non-linear features: the first two features are based on statistical properties of the spectrum and the next four features are timbre features used for music genre classification [32].

Non-linear STFT (NLFT) features were initially proposed mainly for feature extraction from complex-valued SAR images, but experiments showed that they give very encouraging results also for real-valued images.

Our proposed algorithm is an implementation of the non-linear STFT feature extraction. The features parameters computed from the STFT are: *mean* of the STFT coefficients, *variance* of the STFT coefficients, *spectral centroid in range*, *spectral centroid in azimuth*, *spectral flux in range*, and *spectral flux in azimuth*.

IV. PERFORMANCE EVALUATION

Presently Earth Observation (EO) satellites acquire huge volumes of high resolution images, very much over-passing the capacity of the users to access the information content of the acquired data. In addition to the existing methods for EO, data and information extraction are needed new methods and tools to explore and help to discover the information hidden in large EO image repositories.

For our investigation two sites were considered: Berlin-Germany and Ottawa – Canada. For the evaluation of the best features the Berlin site was considered in order to compute the precision-recall of GLCM, GAFS, QMFS and NLFT features. After the best features were identified these are used for answering to the question "Which is the best incidence angle". In this case both sites, Berlin and Ottawa were process.

To evaluate the feature extraction methods and the best incidence angle a tool based on Support Vector Machine with relevance feedback (SVM – RF) was built.

The SVM – RF tool supports users to search images (patches) of interest in a large repository. The Graphical User Interface of this tool allows Human-Machine Interaction to rank the automatically suggested images which are expected to be grouped in the class of relevance. Visual supported ranking allows enhancing the quality of search results by giving positive and negative examples as right and left click respectively.

The size of the images covering the area of: Berlin is 5549x3368 pixels and Ottawa the size is 4783x3381 pixels.

In our case, the product-image is tile in patches with the size of 220x220m, and after that, sub-sampled to 110x110m for better performances (see in [5] the comparison results).

The feature vector for GLCM has a fix number of features for each orientation equal to 12 features, but in our

experiments all the orientations (from 1 to 4) were considered obtaining a feature vector of 48 features (denoted by GLCM_1_2_3_4). In the case of Gabor filters, 4 scales and 6 orientations (48 features denoted by GAFS 4_6) were considered. For QMFS, the number of levels of wavelet decomposition was equal to 1 this means a vector of 8 features was obtained (denoted by QMFS 1), while for NLFT the number of features was fixed to 6.

All the features are normalised before being used in the SVM-RF tool. For normalisation the Z-score normalisation method was selected from the many available and used [33].

We define a number of semantic classes and group the patches accordingly, using the SVM-RF tool and the human expertise. In our approach, for assigning the patches into classes, one patch was assigned only to one class based on the dominant content of the patch.

During the evaluation, the number of classes retrieved for Berlin area is 11 classes and for Ottawa area the number of classes is 6 (some examples are shown in Figure 1).
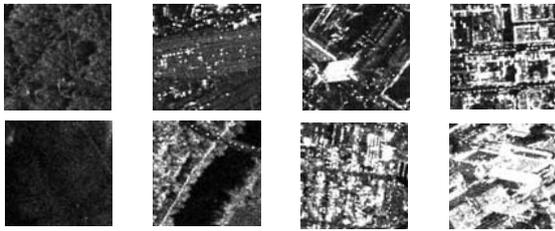
Figure 1. Typical classes extracted from the Berlin image (first line) and from Ottawa (second line).

For each feature extraction method, we tried to detect the classes among the number of identified patches of our database. For each class, we give 20% of the patches of each class for the training as positive examples and one patch from the rest of the classes as a negative example and we tray to detect the similar patches during 7-10 training iterations. The evaluations stop when the classified patches which are displayed by the Search Engine (SVM - RF tool) remain in a stable result. The procedure is repeated two times for the same class, giving the same positive and negative examples in the same order.

For the quantitative assessment, we compared the classification results with the annotated database. We purpose for our evaluation the Precision-Recall that will be computed for each class, feature, and incidence angle.

The precision is defined as the fraction of the retrieved images which are relevant, while the recall is defined as the fraction of relevant images which have been retrieved.

For the evaluation of the best feature that are intend to be used for the evaluation of the incidence angles, in Table I and Table II are displayed (for Berlin site with 30° of the incidence angle) the precision-recall for all four features, each class separately. With red colour is marked the best result obtained for each class, with blue colour is represented the average of the precision or recall for each class or feature algorithm, and with green colour is represented the global average of the precision or recall for entire product-mode-patch (this means for all investigated classes and feature algorithms).

After the investigation and comparison between the features is done the following observation arise:

a) The Gabor filters perform better than the other features especially when the precision is computed.

b) Regarding the recall, the best performance is obtained for quadrature mirror filters.

c) The quadrature mirror filters has the advantage of being faster (in required run time for feature computation) than the Gabor filters.

Based on the previous remarks, for evaluating the best incidence angle the GAFS and QMFS are taking into account. The two selected feature extraction methods were applied for this investigation to our dataset (Berlin and Ottawa).

On the TerraSAR-X archive [6], we identified two sites with different incidence angles and orbit direction: Berlin 30° and 42° with ascending looking and Ottawa 27° and 41° with descending looking.

In the next tables (Table III –IV), for these two sites the precision-recall was computed and the results are displayed. With blue color is represented the average of the precision or recall for each class or feature algorithm and with green color is represented the global average of the precision or recall for entire mode-incidence angle. The best incidence angle was obtained for both sites in the case of bigger value of the incidence angle.

We are focus only to recall because is more relevant than the precision for our investigation.

TABLE I.   THE PRECISION- COMPARISON BETWEEN ALL PF ALGORITHMS (GEC-RE PRODUCT, SPOTLIGHT MODE AND PATCH SIZE 110x110) - BERLIN

Semantics	Class No	GAFS 4_6	GLCM 1_2_3_4	NLFT	QMFS 1	Average features - class
Forest	class00	100.00%	100.00%	73.33%	81.25%	88.65%
Forest + other objects	class01	84.62%	75.76%	76.79%	79.69%	79.22%
Channel	class02	100.00%	92.59%	70.73%	63.16%	81.62%
Train lines type 1	class03	100.00%	100.00%	100.00%	100.00%	100.00%
Urban type 1	class04	86.37%	86.67%	90.91%	68.97%	83.23%
Train lines type 2	class05	100.00%	100.00%	100.00%	100.00%	100.00%
Building reflection	class06	80.10%	65.45%	53.09%	61.25%	64.97%
Urban type 2	class07	100.00%	60.00%	36.84%	65.22%	65.52%
Street plus building	class08	58.33%	75.00%	20.00%	55.00%	52.08%
Urban type 3	class09	81.82%	71.43%	68.75%	90.00%	78.00%
Sport fields	class10	100.00%	100.00%	100.00%	100.00%	100.00%
Average all class / features		**90.11%**	**84.26%**	**71.86%**	**78.59%**	*Total: all classes and features: 81.21%*

TABLE II.   THE RECALL - COMPARISON BETWEEN ALL PF ALGORITHMS (GEC-RE PRODUCT, SPOTLIGHT MODE AND PATCH SIZE 110x110) - BERLIN

Semantics	Class No	GAFS 4_6	GLCM 1_2_3_4	NLFT	QMFS 1	Average features - class
Forest	class00	71.43%	71.43%	78.57%	92.86%	78.57%
Forest + other objects	class01	66.67%	75.76%	65.16%	77.27%	71.22%
Channel	class02	51.35%	67.57%	78.38%	64.86%	65.54%
Train lines type 1	class03	58.33%	66.67%	66.67%	58.33%	62.50%
Urban type 1	class04	63.33%	50.00%	66.67%	66.67%	61.67%
Train lines type 2	class05	27.27%	27.27%	63.64%	63.64%	45.46%
Building reflection	class06	68.06%	50.00%	59.72%	68.06%	61.46%
Urban type 2	class07	32.35%	26.47%	20.59%	44.12%	30.88%
Street plus building	class08	14.58%	31.25%	25.00%	22.92%	23.44%
Urban type 3	class09	32.15%	35.71%	39.29%	32.15%	34.83%
Sport fields	class10	55.56%	55.56%	44.44%	55.56%	52.78%
Average all class / features		**49.19%**	**50.70%**	**55.28%**	**58.77%**	*Total: all classes and features: 53.49%*

TABLE III.   THE PRECISION / RECALL - COMPARISON BETWEEN DIFFERENT INCIDENCE ANGLES (GEC-RE PRODUCT WITH SPOTLIGHT MODE) - BERLIN

Semantics	Class No	Precision		Recall	
		Incidence angle = 30°	Incidence angle = 42°	Incidence angle = 30°	Incidence angle = 42°
		Average features - class		Average features - class	
Forest	class00	100.00%	90.63%	64.29%	82.15%
Forest + other objects	class01	95.17%	82.16%	81.82%	71.97%
Channel	class02	98.08%	81.58%	60.81%	58.11%
Train lines type 1	class03	100.00%	100.00%	41.67%	58.33%
Urban type 1	class04	88.46%	77.67%	58.34%	65.00%
Train lines type 2	class05	100.00%	100.00%	41.67%	45.46%
Building reflection	class06	83.09%	70.68%	42.36%	68.06%
Urban type 2	class07	82.36%	82.61%	32.35%	38.24%
Street plus building	class08	66.59%	56.67%	29.17%	18.75%
Urban type 3	class09	89.59%	85.91%	32.15%	32.15%
Sport and other fields	class10	100.00%	100.00%	44.45%	55.56%
Total all classes and features		*91.21%*	*84.35%*	*48.10%*	*53.98%*

TABLE IV.   THE PRECISION/RECALL - COMPARISON BETWEEN DIFFERENT INCIDENCE ANGLES (GEC-RE PRODUCT WITH SPOTLIGHT MODE) - OTTAWA

Semantics	Class No	Precision		Recall	
		Incidence angle = 27°	Incidence angle = 41°	Incidence angle = 27°	Incidence angle = 41°
		Average features - class		Average features - class	
Water	class00	100.00%	92.19%	70.97%	79.04%
Channel	class01	97.62%	89.58%	65.63%	67.19%
Building reflection	class02	87.97%	79.79%	59.53%	80.96%
Urban type 1	class03	98.17%	96.24%	66.67%	76.07%
Urban type 2	class04	100.00%	68.00%	59.38%	62.50%
Field	class05	88.10%	85.16%	88.10%	66.67%
Total all classes and features		*95.31%*	*95.31%*	*68.38%*	*72.07%*

V.   CONCLUSION

Based on the presented results and the parameters of the TSX products extracted from the XML file, a general conclusion can be drawn that, for value of the incidence angle closer to the upper bound of the sensor range (for TerraSAR-X High Resolution Spotlight mode products the bounds are around 20° for lower value and 55° for upper value) combined with orbit orientation (ascending or descending looking) give better results that in the case when the value of the incidence angle is closer to the lower bound of the sensor range.

For Berlin this is 42° and for Ottawa is 41°. For these two sites the number of retrieved classes is equal to 11 for Berlin area and to 6 for Ottawa area.

The good classes in recall are: for Berlin – forest, forest plus other objects, building reflection and urban; for Ottawa – water, channel, building reflection, urban, field. (2) The bad classes in recall are only for Berlin - street plus building class.

REFERENCES

[1] C.-R. Shyu, M. Klaric, G. Scott, A. Barb, C. Davis, and K. Palaniappan, "GeoIRIS: Geospatial Information Retrieval and Indexing System – Content Mining, Semantics Modeling, and Complex Queries", IEEE Trans. Geoscience and Remote Sensing, vol. 45, Issue 4, pp. 839-852, 2007.

[2] A. Popescu, I. Gavat, and M. Datcu, "Image Patch Contextual Descriptors for Very High Resolution SAR data: A Short Time Fourier Transform Non-Linear Approach", "in press".

[3] P. Birjandi and M. Datcu, "ICA based visual words for describing under meter high resolution satellite images", Proc. of. IGARSS 2009, Cape Town, 2009.

[4] P. Birjandi and M. Datcu, "Patch Contextual Descriptors for Very High Resolution Satellite Images: A Topographic ICA Approach", "in press".

[5] C.O. Dumitru, J. Singh, and M. Datcu, "Selection of relevant features and TerraSAR-X products for classification of high resolution SAR images", EUSAR 2012, May 2012.

[6] TerraSAR-X: "Basic Products Specification Document", Issue: 1.6 (TX-GS-DD-3302), April 2012.

[7] https://centaurus.caf.dlr.de:8443/eoweb-ng/template/default/welcome/entryPage.vm, April 2012.

[8] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification", IEEE Trans. Systems, Man, and Cybernetics, SMC-3, pp. 610-621, 1973.

[9] E. Rignot and R. Kwok, "Extraction of Textural Features in SAR Images: Statistical Model and Sensitivity", Proc. International Geoscience and Remote Sensing Symposium, Washington, DC, pp. 1979-1982, 1990.

[10] A.S. Solberg and A. Jain, "Texture Fusion and Feature Selection Applied to SAR Imagery", IEEE Trans. Geoscience and Remote Sensing, vol. 35, no. 2, pp. 475-478, 1990.

[11] T. Randen and J.H. Husoy, "Filtering for Texture Classification: A Comparative Study", IEEE Trans.Pattern Analysis and Machine Intelligence, vol. 21 no.4, pp. 291-310, 1990.

[12] P.P. Ohanian and R.C. Dubes, "Performance Evaluation for Four Classes of Textural Features", Pattern Recognition, vol. 25, no. 8, pp. 819-833, 1992.

[13] http://www.fp.ucalgary.ca/mhallbey/orderliness_group.htm, April 2012

[14] A.K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters", Pattern Recognition, vol. 24, no. 12, pp. 1167-1186, 1991.

[15] L. J. Du, "Texture Segmentation of SAR Images Using Localized Spatial Filtering", Proc. International Geoscience and Remote Sensing Symposium, Washington, pp.1983-1986, 1990.

[16] J. H. Lee and W. D. Philpot, "A Spectral-Textural Classifier for Digital Imagery", Proc. International Geoscience and Remote Sensing Symposium, Washington, pp. 2005-2008, 1990.

[17] L. Shu, T. Tan, M. Tang, and C. Pan, "A Novel Registration Method for SAR and SPOT Images", Proc. IEEE International Conference on Image, pp. II.213-II.216, 2005.

[18] B. S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data". IEEE Trans. Pattern Analysis and Machine Intelligence, vol.18, pp.837–842, 1996.

[19] P. Porter and N. Canagarajah, "Robust Rotation-Invariant Texture Classification: Wavelet, Gabor filter and GMRF based Schemes", Proc. Visual Image Processing, vol. 144, no. 3, pp. 180-188, 1997.

[20] S.E. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of Texture Features based on Gabor Filters", IEEE Trans. Image Processing, vol. 10, no. 11, pp. 1160-1167, 2002.

[21] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based Image Retrieval Using Gabor Texture Features", IEEE Trans. Pattern Analysis and Machine Intelligence, pp. 13-15, 2000.

[22] M. Torres-Torriti and A. Jouan, "Gabor vs. GMRF Features for SAR Imagery Classification", Proc. Int. Conference on Image Processing, Thessaloniki, vol. 3, pp. 1043 – 1046, 2001.

[23] A. Croisier, D. Esteban, and C. Galand, "Perfect Channel Splitting by use of Interpolation/Decimation/Tree Decomposition Techniques", Proc. International Conference on Information Science and Systems, Patras Greece, 1976.

[24] D. Esteban and C. Galand, "Application of quadrature mirror filters to split-band voice coding schemes", Proc. IEEE Int. Conf. ASSP, Hartford, Connecticut, pp. 191-195, 1977.

[25] P. Vaidyanathan, "Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Techniques", IEEE ASSP Magazine, vol. 4, no. 3, pp. 4-20, 1987.

[26] A. Popescu, C. Patrascu, I. Gavat, J. Singh, and M. Datcu, "Spotlight TerraSAR-X Data Modeling using Spectral Space-Variant Measures, for scene Targets and Structure Indexing", Proc. The 8th European Conference on Synthetic Aperture Radar, Aachen, Germany, 2010.

[27] T. Zou, W. Yang, D. Dai, and H. Sun, "Polarimetric SAR Image Classification Using Multifeatures Combination and Extremely Randomized Clustering Forests", EURASIP Journal on Advances in Signal Processing vol. 2010, ID 465612, 9 pages, 2010.

[28] M. Fauvel, J. Chanussot, J.A. Benediktsson, and J.R. Sveinsson, "Spectral and Spatial Classification of Hyperspectral Data using SVMs and Morphological Profiles", IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, pp. 4834-4837, 2002.

[29] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification", IEEE Trans. Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610–621, 1973.

[30] M. Campedel, E. Moulines, and M. Datcu, "Feature Selection for Satellite Image Indexing", ESA-EUSC: Image Information Mining – Theory and Application to EO, 2005.

[31] A. Popescu, I. Gavat, and M. Datcu, "Complex SAR image characterization using space variant spectral analysis", Proc. IEEE Radar Conference, pp. 1-4, 2008.

[32] Z. Li and M. Ogihara, "Towards Intelligent Music Information Retrieval", IEEE Trans. Multimedia, vol. 8, no. 3, pp. 564-574, 2006.

[33] N. Karthikeyani Visalakshi and K. Thangavel, "Impact of Normalization in Distributed K-Means Clustering". Int. Journal of Soft Computing, vol. 4, no. 4, pp. 168-172, 2009.

[34] http://persci.mit.edu/pub_pdfs/simoncelli_subband.pdf, April 2012.

Evaluation of Complexity Versus Performance for Turbo Code and LDPC Under Different Code Rates

Alaa Eldin.S. Hassan

Dept. Space Science

National Authority for Remote sensing and Space Science (NARSS)

Cairo, Egypt.

E-mail: alaa_eldin@narss.sci.eg

Moawad Dessouky, Atef Abou Elazm and Mona Shokair

Dept. Electrical Communication

Menoufiay University

Menouf, Egypt

atef_abuelazm@menofia.edu.eg,
i_shokair@yahoo.com, Dr_moawad@yahoo.com

Abstract— **Turbo codes play a major role in channel error correction schemes used in wireless communications. Turbo codes emerged in 1993 and since that time they dominate the research in error control coding together with low-density parity-check codes. Due to their remarkable performances, turbo code and low density parity check code have been accepted to a number of standards by many organizations which decide to include turbo code and low density parity check into their new standards after these codes were proven successful in a many of missions. In this paper, the calculation and comparison of performance versus complexity for those two techniques of channel coding was done. For a fair comparison, the performance and complexity should be compared together. The complexity was calculated by counting the number of clock cycles need to complete the decoding algorithm. This comparison is used as a guiding lines of using either turbo code or the low density parity check in specific communication applications, The performance comparison of turbo code and low density parity check were computed for rates 1/2 and 7/8. The complexity for the two codes were calculated for different code rates like (1/2, 1/3, 3/4, 7/8), the evaluation study concludes that the turbo code was recommended for moderate rate, while the LDPC is recommended for higher code rates.**

Keywords-turbo code; LDPC; complexity.

I. INTRODUCTION

Turbo code is a powerful error correcting code used in wireless communication. Turbo code emerged in 1993 [1] and since this year, it becomes a popular area of communications research. Turbo code has a performance near to Shanon limit, and it is stable for long time and now being accepted as standard forward error correction technique by many organizations such as CCSDS, but turbo code still facing high complexity problem, on the other hand Low-density parity-check (LDPC) codes are forward error-correction codes, first proposed in the 1962 by Gallager in his dissertation at MIT [2], at that time it was unpractical to be implemented, but then largely neglected for over 35 years. After that, it is rediscovered again by MacKay and Neal in their work [3]. Because LDPC shares the main concept of message passing algorithm as the turbo code and its performance is also very close to Shannon limit. However, in the last few years, the advances in low-density parity-check

codes have prove that the LDPC beat turbo code in terms of error floor and performance for the higher code rates. In this paper, we analyze the decoding algorithm for turbo code and calculating its complexity under different code rates, and then the same scenario was applied for the LDPC to calculate its complexity and performance for different code rates.

A similar work was made at Stravanger University [4], but the comparison was made for rate 1/2 only, in this paper a complete comparison of turbo code and LDPC was made for different code rates. Also, because the complexity of LDPC is a function of code rate, the complexity was studded for different code rates. It is concluded from this research work that the Turbo code is recommended for moderate code rate because of its better performance, while the LDPC is recommended for higher code rates because of its better performance besides lower complexity compared with turbo code.

This paper is organized as follows; the turbo code decoding algorithm is reviewed in Section II and Section III. The complexity calculations are made in Section IV. The LDPC code decoding algorithm is reviewed in Section V. its complexity calculation is made in Section VI. The comparison of performance and complexity was elaborated in Section VII and Section VIII. Tradeoff between performance and complexity was compared in Section IX; finally, conductive conclusions are done in Section X.

II. TURBO CODE SCHEME

Conventional turbo code consists of two (or more) convolutional codes connected in serial or in parallel via some pseudo-random interleavers.
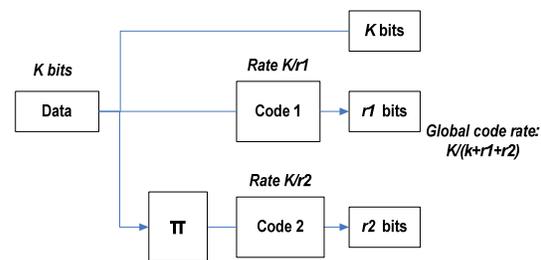
Figure 1. Basic structure of the turbo encoder

Figure 1 presents a block diagram of an encoder of a systematic turbo code with an information block length K. The turbo code is composed of two Recursive Systematic Convolutional codes (RSC) [1].

The information bits are fed to the first RSC and after being interleaved are passed through the second constituent encoder. The resulting codeword consists of the systematic bits, k (i), and two parity check streams, r1(i), r2(i), i = 1, 2, . . . , K . The coding rate of this code is 1/3. Higher code rates can be achieved by puncturing some of the parity check bits, using more constituent codes result in codes with rates lower than 1/3

III. TURBO CODE DECODING ALGORITHM

A turbo decoder consists of two concatenated decoders, each using the received systematic stream and the corresponding received parity stream. Each decoder provides a soft output of the transmitted bits by using the received data and the information provided by the other decoder. The soft output is the a posteriori probability (APP) and consists of three components: the intrinsic information which is a function of the received signal for the corresponding bit position, the a priori (AP) probability of that bit position and the extrinsic information which comes from the received signal for other bit positions and their a priori probabilities, In each iteration the extrinsic information produced by the other constituent decoder is used to evaluate the a priori probabilities in that iteration. Repeating this procedure improves the estimation of the bit probability values and hence, reduces the probability of error. One efficient algorithm for soft output decoding, based on the trellis diagram of the code known as the BCJR algorithm, is presented in [5]. The suboptimal decoder introduced in [6] Finds the extrinsic information on the transmitted bits by one of the constituent decoders and passes it to the other decoder through the interleaver. The decoder can decode the received vector only if the iterative decoding converges. The output of the "symbol-by-symbol" Maximum a posteriori Probability (MAP) decoder is defined as the a posteriori log-likelihood ratio, that is, the logarithm of the ratio of the probabilities of a given bit being "+1"or"-1' given the observation y, as in equation (1). The Max-log MAP algorithm for decoding the turbo code was used as it presented in [7] and it is based on the trellis of a convolutional encoder in Figure 3.
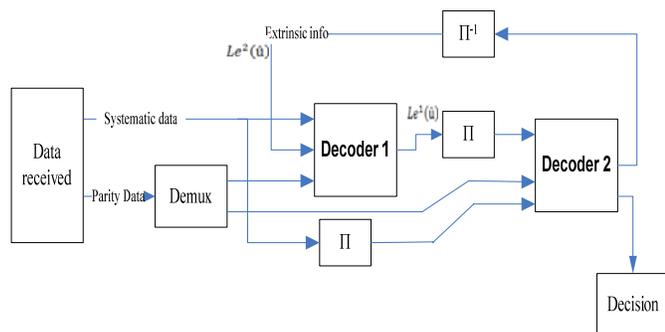
Figure 2. Basic structure of an iterative turbo decoder
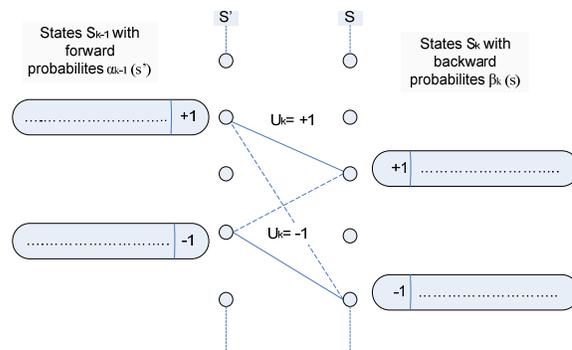
Figure 3. Trellis structure of systematic convolutional codes

The log-likelihood ratio of (u_k) is defined by [7]

$$L(\hat{u}_k) = ln \frac{P(u_k = +1|y)}{P(u_k = -1|y)} = ln \frac{\sum_{u_k=+1}^{(s',s)} P(s',s,y)}{\sum_{u_k=-1}^{(s',s)} P(s',s,y)} \quad (1)$$

$$L(\hat{u}_k) = ln \frac{\sum_{u_k=+1}^{(s',s)} \alpha^*_{k-1}(s').\gamma^*_k(s',s).\beta^*_k(s)}{\sum_{u_k=-1}^{(s',s)} \alpha^*_{k-1}(s').\gamma^*_k(s',s).\beta^*_k(s)} \quad (2)$$

where u_k is the information bit at time k, and α^*_k: is the Probability of path in the trellis going from state S' (k-1) and terminating at state S (k)

$$\alpha^*_k(s) = \sum_{(s,s')} \gamma^*_k(s',s) \alpha^*_{k-1}(s') \quad (3)$$

and β^*_k is the Probability of path in the trellis going from state S(k-1) and terminating at state S'(k)

$$\beta^*_{k-1}(s') = \sum_{(s,s')} \gamma^*_k(s',s) \beta^*_k(s) \quad (4)$$

and γ^*_k is the branching Probability of path in the trellis going from state S (k-1) and terminating at state S'(k)

$$\gamma^*_k(s',s) = \frac{1}{2}.x_k.L_a.(x_k) + x_k.L_c.x'_k + p_k.L_c.p'_k \quad (5)$$

IV. COMPLIXITY OF TURBO DECODING

In this section, we have to have a specific formula for the complexity needed for Turbo code implementation by counting the number of processor cycles for mathematical operations needed for decoding, which related to the time needed to decode a frame of information encoded by Turbo code. From Equations (3), (4), and (5) for the Max-Log-MAP algorithm, the α^*, β^*, and γ^* have to be calculated. The calculation of α^*_k (u) requires two multiplications and one addition for each state. Assuming the encoder with memory length M, so we have to multiply by the number of states 2^M so, $(2*2^M)$ multiplications is needed and 2^M addition. The same number of calculation is needed for β^*_k (u). For the γ^*_k we need $(3* 2^M)$ multiplications and $(2 *2^M)$ additions plus $(1* 2^M)$ divisions.

Let us take turbo code with M = 4 so, 2^M = 16. The decoder has two operations per iteration, so, all number should be multiplied by two.

The total number of additions is (4 * 2 * 16) = 128;

Total number of multiplications is (7 * 2 * 16) = 224;

Total number of divisions for branch metric is (2*16) = 32;

Also, here we assume that max* (a,b) ≈ max (a,b) because the correction factor is small and will reduce the complexity, but its effect on the performance is less than 0.2 dB [7].

So, Max* operation need two cycle per state per iteration.

The total operations per bit is $(2*2 *2^M)$ = 64;

For the logarithm operation, a look up table was used which need one operation per state per iteration.

The total operations per bit is $(1*2 *2^M)$ = 32

Finally, the divisions needed to calculate final L (u_k) is per bit.

Here, it is assumed that 3 clock cycles needed for an addition, 5 clock cycles needed for multiplication operation, and 17 clock cycles for a division operation as in the case for a typical Pentium processor [8]. The ratio of calculation complexity is assumed to be as follows: Addition: multiplication: division= 1:1.5:5.

$$224 + 224 * 1.5 + 33 * 5 = 725 \text{ cycles/iteration/ frame length} \quad (6)$$

Table I summarizes the processor cycles for decoding of Turbo code for different code rates, the number of iterations per frame length "K" is 10; as we mention before, the complexity of turbo code is the same as the mother code rate, e.g., rates 1/2 and 3/4 can be obtained from the mother code 1/3, so all of them has the same complexity as the mother code 1/3.

TABLE I.   COMPLEXITY OF TURBO CODE FOR DIFFERENT CODE RATES

Code Rate	Processor cycles/ iteration/Frame length	For 10 iterations /Frame length
Rate 1/3	425/ Frame length	7250
Rate 1/2	725/ Frame length	7250
Rate 3/4	725/ Frame length	7250

V.   LDPC DECODING PRINCIPLES

Decoding of LDPC used message passing algorithms, these algorithms interpreted by bipartite graph representation of the LDPC code [7], where variable nodes and check nodes are connected through edges. The variable nodes and check nodes exchange the messages along their edges in an iterative fashion, thereby cooperating with each other in the decoding process.

The operations in an LDPC decoder comprises of two steps; first, is how the messages are generated at the check nodes and variable nodes (called "node processing"); second, which determine how the generated messages are passed between each other (called "scheduling"). These two operations determine the decoding complexity of LDPC.

A. Node processing

Node processing consists of variable node update (VNU) and check node update (CNU). In the VNU, incoming messages from the check nodes are processed at each VN, and the outgoing messages are generated and passed to the check nodes. Similarly, in the CNU, incoming messages from the variable nodes are processed at each CN, and the outgoing messages are generated and fed back to the variable nodes. Thus, the messages are passed between the variable nodes and check nodes iteratively.

Let C(n) denote the set of check nodes connected to variable node n, and V(m) denote the set of variable nodes connected to check node m, where $0 \leq n \leq N-1$, and $0 \leq m \leq M-1$. C(n)\m refers to exclusion of m from set C(n), and similarly V(m)\n refers to exclusion of n from set V(m).

In the VNU, variable node "n" has messages $R_{m'n}$ coming in from all check nodes m′ connected to it and its channel $L_{ch}(n)$. Hence, the outgoing message ("extrinsic") Q_{nm} on an edge n→m is the sum of all messages except R_{mn}. at iteration i, each variable node "n" calculates messages $Q_{nm}(i)$, which is sent from variable node "n" to each check node m∈C(n). Message $Q_{nm}(i)$ is the LLR of variable node "n" based on all check nodes in C(n)\m, and is calculated as defined by [7]

$$Q_{nm}(i) = L_{ch}(n) + \sum_{m' \in c(n) \backslash m} R_{m'n}(i-1) \quad (7)$$

where $L_{ch}(n)$ is the channel LLR of variable node "n". The computation is shown in Fig. 4-a, assuming that the variable node has degree = 3. The a posteriori LLR for a variable node is obtained by adding all the incoming messages at the variable node.

where "m" can be any check node in C(n). The above expression indicates that the variable-to-check message $Q_{nm}(i)$ in a current iteration can be directly calculated from the previous iteration and the check-to-variable message $R_{mn}(i-1)$ on the same edge from the previous iteration.

$$\sum_{n' \in v(m)} \oplus x_{n'} = 0 \quad (8)$$

$$x_n = \sum_{n' \in v(m) \backslash n} \oplus x_{n'} \quad (9)$$

(a)   (b)
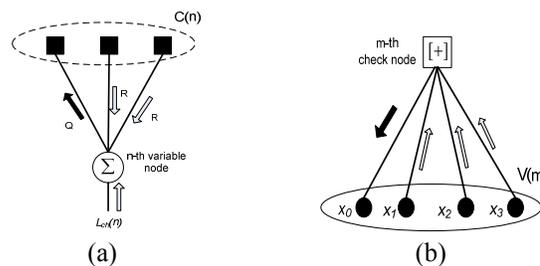
Figure 4. Operations in a belief propagation decoder: (a) Variable Node Update for a degree-3 variable node and (b) Check Node Update for a degree-4 check node. Only message update on one edge (marked with solid arrow) is illustrated but the similar operations are used to update messages on all the edges.

In the CNU, check node "m" has messages $Q_{n'm}$ coming in from all variable nodes "n'" connected to it. Each check node "m" calculates messages $R_{mn}(i)$, which is sent from check node m to each variable node "n" ∈ V(m). Message $R_{mn}(i)$ is the LLR of variable node "n" based on all the variable nodes in V(m)\n.

B. Scheduling

Scheduling involves communicating messages from one node to another as dictated by the edge connections in the bipartite graph. There are two typical schedules of belief propagation: flooding and layered schedule [9]. In flooding, the entire bipartite graph is flooded with messages that are passed back and forth along all the edges, as illustrated in Fig. 5-a. However, this 'flooding' increases the complexity especially for longer block-sizes when the number of edges becomes large.

(a)

(b)
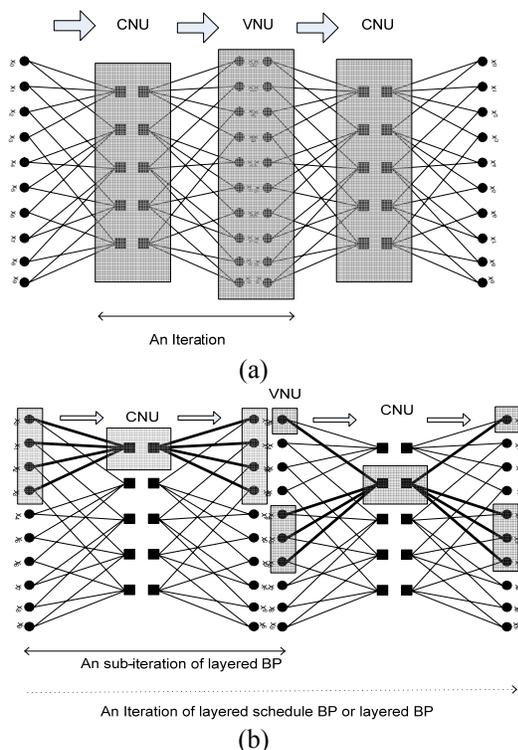
Figure 5.   Message passing in the a) flooding schedule of the belief propagation algorithm, the shaded boxes indicate the CNU and VNU and block arrows indicate direction of message passing. Message passing occurs on a per-iteration basis b) layered BP, Shaded boxes indicate CNU and VNU and block arrows indicate direction of message passing. In the example, one CNU is done per sub-iteration. The edges that are updated in each sub-iteration are shown with thick solid lines.

In the layered algorithm, only a small fraction of the variable nodes and check nodes are updated per sub-iteration, as illustrated in Fig. 5-b. The messages generated in a sub-iteration of a current iteration are immediately used in subsequent sub-iterations within the same iteration. This leads to a faster flow of information and helps improve decoding speeds for the structured LDPC codes.

VI.   LDPC COMPLEXITY CALCULATION

While actual decoding complexity depends on many factors such as hardware architecture, the decoding complexity as estimated in [9], based on operations count, the LDPC decoder computational complexity for Layered BP decoding per-iteration to be as follows

For the Check node update

$$(Nd_v + 2)(N - k) \qquad \textit{Additions}$$

$$(2d_c - 3)(N - k) \qquad \textit{Comparisons}$$

For the Variable node update

$$(N * d_v) \qquad \textit{Additions}$$

So, the Total complexity/ iteration

$$(N - k)(2d_c + 1) + 2Nd_v \qquad \textit{operations}$$

Total complexity

$$K\left(\frac{1}{R} - 1\right) + (4d_c + 1) \qquad (10)$$

where R=K/N is the code rate, N is the decoded frame length, d_c is the average row weight and d_v is the average column weight,

Table II uses the form of equation (10) which manifests the complexity as a function of code rates for different code rates per frame length and 20 iterations.

TABLE II.   COMPLEXITY OF LDPC CODES FOR DIFFERENT CODE RATES FOR FRAME LENGTH "K"

	Code rate	d_v and d_c values	Complexity/ iteration/ frame length	For 20 iterations/frame length K	
Rate increasing	Rate 1/3	dv = 4, dc = 6	2K (4dc+1) = 50 K	50 * 20* K = 1000K	Complexity increasing
	Rate 1/2	dv = 3,dc = 6	K (4dc+1) = 25 K	25* 20* K = 500 K	
	Rate 3/4	dv = 3,dc = 12	0.3K (4dc+1)= 16.3 K	16.3* 20*K = 326 K	

From Table II, it is concluded that the complexity is decreased as the code rate increased.

VII.   LDPC AND TC PERFORMANCE COMPARISON

For the performance of these codes, a simulation is made for 7/8 turbo code and compared with the performance of QC-LDPC 7/8 (8176, 7156) in [10] and the result of comparison is depicted in Fig. 6. The results show that the performance of LDPC is better than the turbo code for higher rates. Another simulation is made for turbo code rate 1/2 and compared with the performance of LDPC rate 1/2 in [11] and the result of comparison is depicted in Fig. 7. The results show that the performance of turbo code is better than the LDPC for moderate code rates. LDPC is better in performance for high code rates (rate 7/8) than the Turbo Code, Beside that the iterations in LDPC can be done in parallel but for turbo code is in serial. Here, the Turbo code is better in performance for moderate code rate than the LDPC.
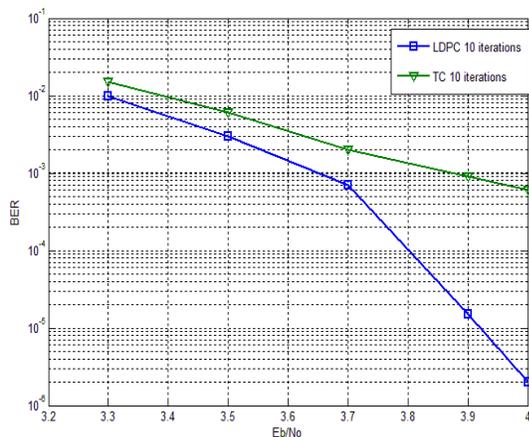
Figure 6. Performance comparison between Turbo code (O) and LDPC (□) for rate 7/8
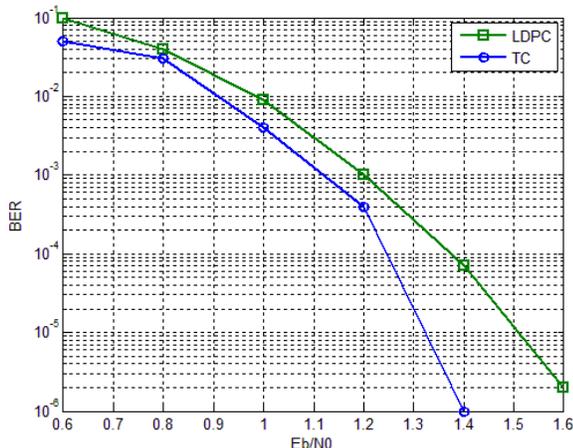
Figure 7. Performance comparison between Turbo code(O) and LDPC (□) for rate 1/2 and coded frame length 4000 bit

So, it is concluded that from the point of view of performance, the LDPC is recommended for higher code rates for communication systems applications, while in the low code rates it is better to use the turbo code. In the next section, a complexity comparison for turbo code and LDPC for different code rates will take place.

VIII. LDPC AND TC COMPLEXITY COMPARISON

For the LDPC codes, the decoding converges within 15 to 20 iterations, while it is well-known that ten turbo-decoding iterations are sufficient for convergence. Therefore, for a fair comparison between LDPC and turbo-decoding algorithms, the number of iterations is chosen to be 20 and 10, respectively. also we have to notice that the LDPC is error detection and correction code, so, when we reach error-free frame before we reach the 20 iteration the decoding will stops, while in Turbo code the decoding has to continue to the total number of iteration even if no more improvement. The operations count of LDPC and turbo decoding algorithms are listed in Table III. The calculation of complexity for Turbo code and LDPC was calculated for

different code rates. the complexity of the turbo code is constant and does not depend on the code rate because different rates comes from the puncturing of the mother code rate, while in the LDPC the complexity is based on the code rate, the higher the code rate the lower the complexity and vice versa.

The calculation is made for rates 1/3, 1/2, and 3/4. The summary of calculations is in Table III.

TABLE III.   COMPLEXITY COMPARISON FOR LDPC AND TURBO CODE FOR DIFFERENT CODE RATES

	LDPC	TC	Complexity reduction Ratio
Number of Iterations	20	10	
Rate 1/3	50*20*K=1000 K	725*10*K= 7250K	13%
Rate 1/2	25*20*K=500K	7250K	7%
Rate 3/4	16.3*20*K= 326K	7250K	5%
	Complexity decreased for higher rates	Complexity is constant	

IX.   COMPLEXITY VS. PERFORMANCE

The simulation results for Turbo code and LDPC are shown in Fig. 6 and 7 for different code rates (1/2 and 7/8), and the complexity calculation is tabulated in Table-III which has the complexities at different code rate. So, it is concluded that for higher code rates, LDPC has better performance and lower complexity, while for rate 1/2 the turbo code has better performance so it should be used even it is more complex because the performance is an important issue. The brief of recommendations for applications for different code rate is summarized in Table IV.

TABLE IV.   RECOMMENDED DECODING ALGORITHM AT DIFFERENT CODE RATES FOR   LDPC AND TURBO CODE

Code Rate	Complexity/ Iteration	Recommended Coding Technique
Low Code Rates	1/2, 1/3, 1/4, 1/6	Turbo code
High Code Rates	2/3, 3/4, 7/8	LDPC

X.   CONCLUSION AND FUTURE WORK

In this paper, a complexity needed to decode a Turbo code and LDPC were calculated; besides, the simulation of their performance was made. A comparison between two codes should compare the complexity and performance before applied in any communication system for specific application. The performance comparison of turbo code and LDPC were computed for rates 1/2 and 7/8. And the complexity for the two codes were calculated for code rates (1/2, 1/3, 3/4). The performance and complexity were based on ten decoding iterations for turbo code, while it is 20 iterations for LDPC.

It is concluded that the turbo code has better performance in moderate code rate (Rate 1/2) while the LDPC is recommended for higher code rates (3/4,7/8) because it has better performance beside less complexity compared with

turbo code. For turbo code, all code rates require the same decoding complexity since all code rates are obtained from the mother code via puncturing. In contrast, the LDPC decoding complexity decreases as the code rate increases.

REFERENCES

[1] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding:Turbo Codes", IEEE Transactions on Communications, vol. 44, pp. 1261-1271, Oct. 1996.

[2] R. G. Gallager, "Low Density Parity Check Codes," Monograph, M.I.T. Press, 1963.

[3] D. MacKay and R. Neal, "Near Shannon Limit Performance of Low Density Parity Check Codes," Electronics Letters, July 1996

[4] K. Fagervik and S. Larssen, "Performance and Complexity Comparison of Low Density Parity Check Codes and Turbo Codes," Stravanger University Website.

[5] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," IEEE Trans. on Information Theory, vol. 20, pp. 284-287, March 1974.

[6] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding Turbo-codes," in IEEE International Conference on Communications, Geneva, Switzerland, May 1993, pp. 1064-1070.

[7] S. Lin and D. Castello, "Error Control Coding" 2nd ed. New Jersey: Pearson Prentice Hall, Chapter 16 and 17, 2004.

[8] T. Kwon, J. Sondeen and J. Draper, "Floating-Point division and square root implementation using a Taylor-series expansion algorithm with reduced look-up tables," in Proceedings of the IEEE Symp. Circuits and Systems, 2008, pp. 954-957.

[9] Y. Blankenship, Stephen Kuffner "LDPC Decoding for 802.22 Standard" IEEE P802.22, 2007.

[10] CCSDS , "Low Density Parity Check Codes for Use in Near –Earth and deep space communications" 131.1-O-2, 2007.

[11] Michael Yang, William E. Ryan and Yan Li, "Design of Efficiently Encodable Moderate-Length High-Rate Irregular LDPC Codes" IEEE Transaction on Communication, Vol. 52, No. 4, April 2004.

Narrowband Uplink Transmission in LTE-based Satellite Radio Interface

Hee Wook Kim, Kunseok Kang, and Bon-Jun Ku

Electronics and Telecommunications Research Institute

305-350 Gajeong-dong Yuseong-gu Daejeion, Korea

{prince304, kskang, and bjkoo}@etri.re.kr

Abstract— **In this paper, we introduce narrowband uplink transmission scheme to adapt orthogonal frequency division multiple access scheme to satellite environments. Maximizing the commonalities with the terrestrial system is one the most important factors for cost-effective service delivery, which will finally bring successful deployment of the system. For this purpose, we adopt an orthogonal frequency division multiplexing based radio interface. At the same time, the interface should be tailored in order to be operable in a power limited uplink scenario of satellite systems. In this paper, we present a narrowband uplink transmission, which was developed for this purpose, in long term evolution based satellite radio interface. Simulation results show that the proposed scheme can increase an uplink spectral efficiency at power-limited handheld terminal.**

Keywords-narrowband transmission; handheld terminal; mobile satellite communications.

I. INTRODUCTION

Considering cost-effective, in a future Mobile Satellite Service (MSS) system, a satellite radio interface needs to be compatible with a maximum degree of commonality with emerging terrestrial standards. Therefore, the techniques adopted for the satellite system have to be similar to or even the same as those of the terrestrial system. The adaptation of a compatible radio interface with maximum commonality will bring possibility to reuse terrestrial part technology to minimize the modification of User Equipment (UE) chipset and network equipment for low cost and fast development.

As emerging terrestrial radio interfaces, the Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) and the Institute of Electrical and Electronics Engineers (IEEE) mobile Worldwide Interoperability for Microwave Access (WiMAX) radio interfaces are being considered [1]. Both two radio interfaces adopted Orthogonal Frequency Division Multiplexing (OFDM) scheme, which is intrinsically able to handle the most common radio frequency distortion without the need for complex equalization techniques and scales easily to fit different bandwidth [2].

Not much attention has been paid to the study on OFDM based satellite radio interfaces due to serious Peak to Average Power Ratio (PAPR) problems, especially for a high cost power amplifier in satellite systems. Nevertheless, recent study results reported the adaption of OFDM technique in the satellite systems to give benefits such as capability of high speed transmission and commonalities with the terrestrial systems [3]-[5]. For example, DVB-SH

adopted OFDM technique [3], which is the same signal format defined in DVB-H for terrestrial systems. The main reason for adopting OFDM stems from the fact the satellite terrestrial transmitters form a Signal Frequency Network (SFN). Furthermore, European Telecommunications Standardization Institute (ETSI) has started feasibility study on OFDM based satellite radio interface, and reported that OFDM based scheme might provide better performance that the conventional Wideband Code Division Multiple Access (WCDMA) based scheme [4]. Recently, a new satellite radio interface is being developing based on the 3GPP LTE standard to maximize the commonalities with the terrestrial system using OFDM [6]. Some satellite specific technologies were proposed including frequency reuse techniques, uplink synchronization techniques, random access techniques, and interleaving technique combined with adaptive modulation and coding.

In this paper, we consider physical layer modifications to terrestrial LTE to make it operable over a satellite channel, which characteristically is power constrained. Following introduction, in Section II, we propose a narrowband uplink transmission in LTE based satellite radio interface. In Section III, we show a few simulation results and then, draw conclusion in Section IV.

II. NARROWBAND UPLINK TRANSMISSION IN LTE BASED SATELLITE RADIO INTERFACE

Mobile satellite system would be power-limited system as well as hand-held terminal has a limited maximum transmitted power. Considering that total transmitted power would be distributed over whole subcarriers in one Resource Block (RB), the large RB size of 180 kHz in terrestrial LTE makes sufficient power not be allocated on one subcarrier at handheld terminals. In this case, high modulation and coding rate scheme may not be supported. Therefore, in this paper, a narrowband RB transmission is defined in order to solve this problem. For high layer commonality, the size of transport block in the RB is same as in terrestrial LTE.

Figure 1 shows Uplink Physical Uplink Shared Channel (PUSCH) structure in order to support narrowband transmission. The PUSCH carries data from the uplink shared channel transport channel [7]. In satellite systems the available bandwidth is constrained due to power limited environments, particularly in uplink. This means that the bandwidth that can be dedicated to one transport block also should be constrained. The constraint can be in the form of fewer subcarriers. Because the size of the transport block for narrowband transmission should be maintained for no
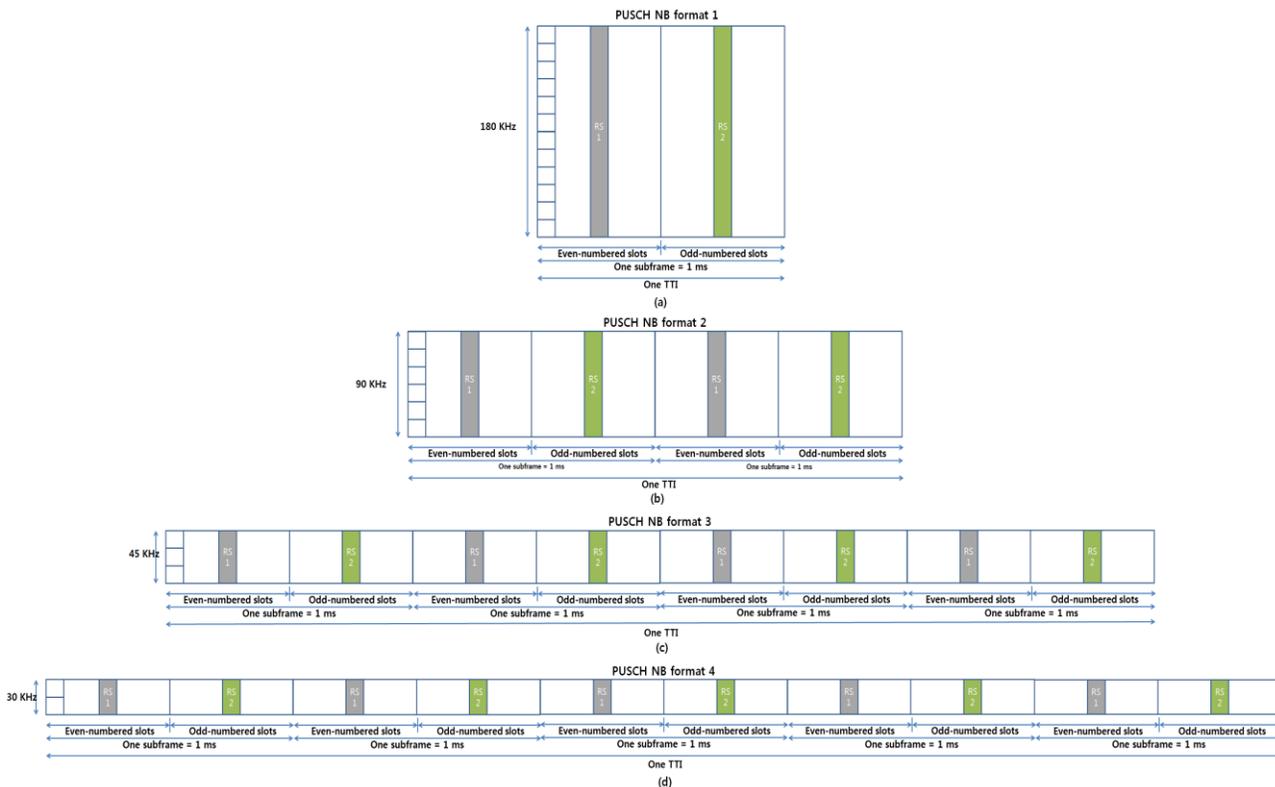
Figure 1. Uplink PUSCH channel structure for narrowband transmission.

modification on terrestrial LTE MAC layer, the data in the transport block is better inserted in such a way that it occupies a larger number of symbols compared to the terrestrial LTE system. For this, LTE physical layer should be modified in order to reduce the size of RB and increase of the length of Transmission Time Interval (TTI) of terrestrial LTE. In terrestrial LTE, 1 ms of TTI is considered in order to reduce latency of service delivery and make fast resource adaptations. However, considering that a satellite system has already a few hundred miliseconds of very long round trip delay and mainly suffers from slow channel fading effects, the 1ms of short TTI doesn't give any advantages in the satellite systems and prevents to get a time diversity gain to compensate slow channel fading effects. Therefore, the increase of the length of TTI in the satellite system will be under a reasonable adaptation of terrestrial LTE to satellite environment.

Figure 1(a) shows the conventional PUSCH structure with the large RB size of 180 kHz, which has 12 subcarriers and 2 slots. The information bits are first channel-coded with a turbo code of mother code rate with 1/3, which is adapted to a suitable final code rate by a rate-matching process. This is followed by symbol-level channel interleaving, which follows a simple 'time-first' mapping – in order words, adjacent data symbols end up being mapped first to adjacent Single Carrier-Frequency Division Multiple Access (SC-FDMA) symbols in the time domain, and then across the subcarriers. The coded and interleaved bits are then

scrambled by a length-31 Gold code prior to modulation mapping, DFT-spreading, subcarrier mapping and OFDM modulation. For channel estimation and data demodulation, a Reference Signals (RSs) 1 and 2 are transmitted in each even-numbered and odd-numbered slots of a TTI, respectively.

Figure 1(b) shows the proposed narrowband PUSCH structure with the RB size of 90 kHz, which has 6 subcarriers and 4 slots. Its channel bandwidth is decreased to the half and TTI is increased to twice, compared to those of the conventional PUSCH. A simple 'time-first' mapping is also made for efficient transmission in power-limited satellite uplink. Within one TTI, adjacent data symbols end up being mapped first to adjacent SC-FDMA symbols in the time domain, and then across the subcarriers. On the other hand, in order to reuse the conventional RSs 1 and 2 in narrowband PUSCH as well as get a time diversity gain, the first half of RSs 1 and 2 are transmitted in the first and second slots, respectively, and then the rest half of RSs 1 and 2 are mapping in the third and fourth slots, respectively.

In a same way to Fig. 1(b), Fig. 1(c) and 1(d) also show the proposed narrowband PUSCH structure with the RB sizes of 45 and 30 kHz, respectively. As we can see, the number of subcarriers in the narrowband PUSCHs is limited to products of 2, 3 and 5 for commonality with terrestrial LTE because DFT size in terrestrial LTE is limited to those for low complexity of DFT implementation.
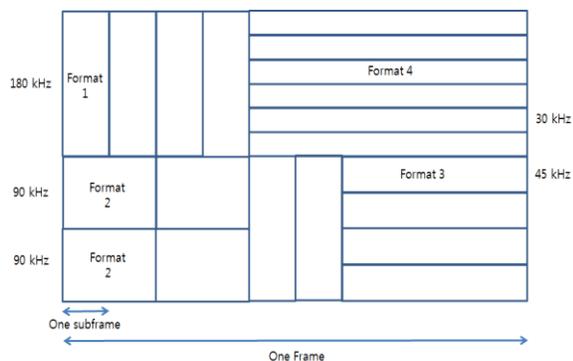
Figure 2. Uplink PUSCH channel structure for narrowband transmission.
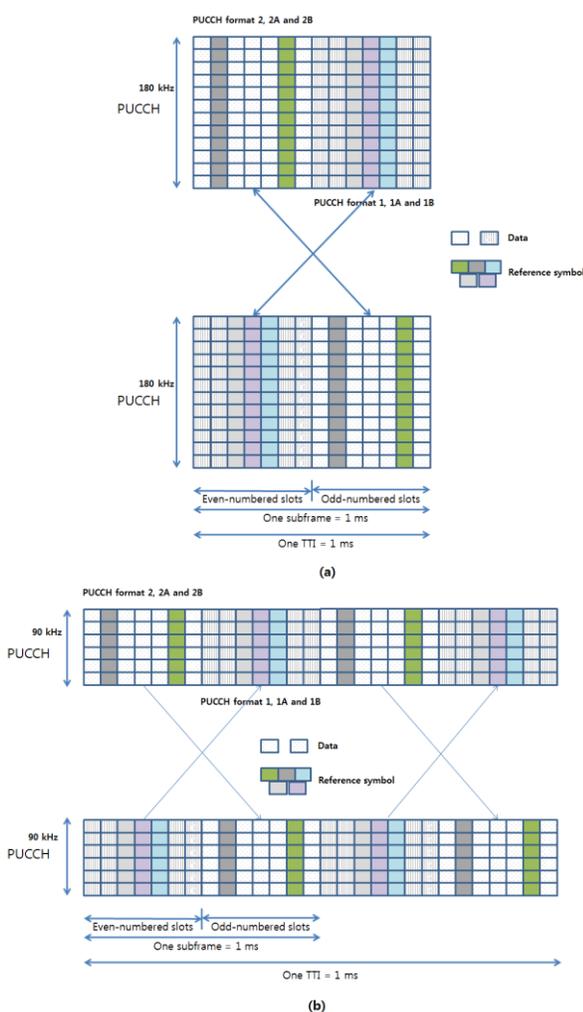
Figure 3. Uplink PUSCH channel structure for narrowband transmission.

Figure 2 illustrates a method to support the narrowband PUSCH transmission in conventional terrestrial LTE frame
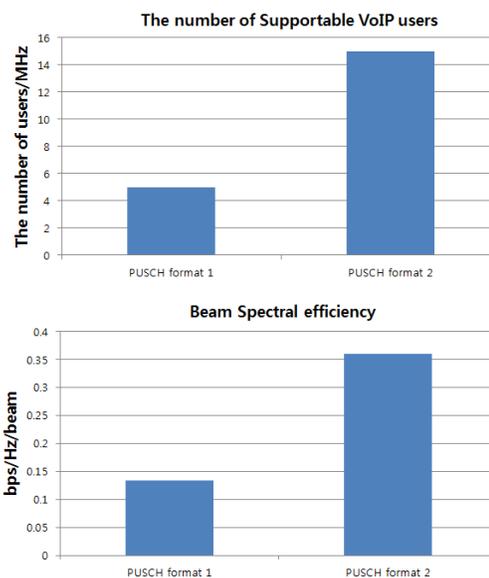
Figure 4. System level simulation results

structure. We assume that one TTI corresponds to one frame duration in a satellite radio interface. Because the terrestrial LTE is operated with 180 kHz size of RB, narrowband PUSCHs should be grouped with 180 kHz bandwidth size within one TTI in order to be compatible with terrestrial LTE. For example, Formats 1 and 4 are transmitted in the above 180 kHz bandwidth and formats 1, 2 and 3 are mapped in the below 180 kHz in Fig. 2.

In the same principle, Physical Uplink Control Channel (PUCCH) structure will be shown in Fig. 3. PUCCH is used by a UE to transmit any necessary control signaling only in subframes, in which the UE has not been allocated any RBs for PUSCH transmission. The control signaling on the PUCCH is transmitted in a frequency region on the edges of the system bandwidth. In order to minimize the resource needed for transmission of control signaling in one subframe is a single slot at or near one edge of the system bandwidth, followed by a second RB at or near the opposite edge of the system bandwidth, as shown in Fig. 3. Similarly to PUCCH structure, Figures 3(a) and 3(b) represent the conventional PUCCH formats 1 and 2 and their narrowband transmissions version for adaptation to satellite uplink. Fig. 3(b) shows the narrowband PUCCH structure with the RB size of 90 kHz, which has 6 subcarriers and 4 slots. Other narrowband PUCCH structures can be applied in a same way to Fig. 2(c) and 2(d) for narrowband PUSCH structures.

III. SIMULATION RESULTS

For system-level simulation, we exploit the evaluation configuration parameters in the Table 1. Evaluation is performed in open environment defined in ITU-R Report M.2176, which identifies visions and requirements for the satellite component of IMT-Advanced [8]. We assumed that UEs are randomly distributed over whole coverage and are located outdoor with the mobility of 3km/h. For assessment

of beam spectral efficiency, beam spectral efficiency is defined as the aggregate throughput of all users (the number of correctly received bits, i.e. the number of bits contained in the Service Data Units (SDUs) delivered to Layer 3) divided by the channel bandwidth by the number of beams. Also, full buffer best effort service profile is considered. VoIP capacity is derived assuming a 12.2 kbps codec with a 50% activity factor such that percentage of users in outage is less than 2%, where a user is defined to have experienced a voice outage if less than 98% of the VoIP packets have been delivered successfully to the user within a one-way radio access delay bout of 400 ms, considering maximum transfer delay of one way for the real-time services in the satellite component.

TABLE I.   EVALUATION CONFIGURAITON PARAMETERS

Parameters	Values used for evaluation
Deployment scenario	Open environment, GEO satellite
Duplex method and bandwidths	FDD: 5(Up) + 5(Down) MHz, 2.1 GHz carrier frequency
Frequency reuse plan	Reuse factor 6
Number of beams	20 (3dB of beam edge loss)
Transmission scheme	SISO
Scheduler	Channel dependent
Power control	None (allocate full power)
Link adaptation	Non-ideal based on delayed SRS-based measurements: MCS based on LTE transport formats and SRS period and bandwidths according to LTE Rel-8
HARQ scheme	Incremental redundancy or Chase combining None for VoIP traffic
Receiver type	MMSE
Satellite antenna	ITU-R Recommendation S.672, 50dBi gain
UE antenna	Omnidirectional, 0dBi gain
UE transmit power	250 mW
Channel estimation	Non-ideal
Feedback and control channel errors	None
HARQ/ARQ interaction	HARQ/ARQ interaction scheme for full buffer traffic.
MAC/RLC header overhead	Assume minimum size of specification
Layout	Hexagonal grid
Inter-site distance	180 km
Satellite system noise temperature	450 K
G/T	23.47 dB/K
Target packet error rate	1 %
Path loss	189.5 (LoS) + 2.5 (fading margin) dB

Based on these assumptions, we compare performance of a narrowband PUSCH (format 2 in Fig. 1) with the conventional PUSCH (format 1 in Fig. 1). Figure 4 shows system level simulation results with respect to average beam spectral efficiency and VoIP capacity. As seen in the figure, the use of narrowband PUSCH format 2 can get the increase of total system throughput more than three times of that in the use of the conventional PUSCH channel. Furthermore, an UE using the narrowband PUSCH format 2 can support slightly higher data rate than an UE using the conventional PUSCH even if the narrowband PUSCH format 2 has the half of the conventional PUSCH bandwidth.

From the performance analysis of a narrowband PUSCH format 2, we expect that the use of narrowband PUSCH format 3 and 4 could make total throughput increase more than that of the narrowband PUSCH format 2 as well as conventional PUSCH format.

IV.   CONCLUSIONS

The proposed narrowband transmission scheme is more granular than in terrestrial LTE and allows allocation of reduced bandwidth resources to an UE, consistent with the power constrained nature of the satellite channel. In the proposed scheme, higher power allocation on each subcarrier in RB can make higher modulation and channel coding rate be used, thus we can achieve the increase of total system throughput. On the other hand, modification of physical layer and no change of the transport block size from terrestrial LTE in the proposed scheme can make us design fully compatible satellite LTE radio interface over MAC layer.

REFERENCES

[1] Soo Yun Hwang, et al., "Design and implementation of a latency efficient encoder for LTE systems," ETRI Journal, vol. 32, no.4, Aug. 2010, pp.493-502.

[2] Ziad A. Shamsan, et al., "Coexistence of OFDM-based IMT-Advanced and FM brodcasting systems," ETRI Journal, vol.33, no.2, Apr. 2011, pp. 279-282.

[3] ETSI EN 302 583, "DVB; Framing structure, channel coding and modulation for DVB-SH below 3 GHz," March 2008.

[4] ETSI TR 102 443, "SES; Satellite component of UMTS/IMT-2000; Evaluation of the OFDM as a satellite radio interface," Aug. 2008.

[5] Arif Ansari, et al., "S-WiMAX: Adaptation of IEEE 802.16e for mobile satellite services," IEEE Communication Magazine, vol. 47, no.6, June 2009, pp.150-155.

[6] Hee Wook Ki, et al., "Applicability of orthogonal frequency division multiple access in satellite communication," IEEE VTC2011Spring Conference, May 2011, pp.1-6.

[7] 3GPP TS 36.211, "3GPP; Technical specificaiton group radio access network; E-UTRA; Physical channel and modulation (Release 8)," May 2008.

[8] ITU-R Report M.2176, "Visions and requirements for the satellite component of IMT-Advanced," Jan. 2011.

Data-Bits Asynchronous Tracking Loop Scheme for High Performance Real-Time GNSS Receivers

Pedro A. Roncagliolo, Javier G. García and Carlos H. Muravchik

Laboratorio de Electrónica Industrial, Control e Instrumentación (LEICI),

Dto. Electrotecnia, Facultad Ingeniería, UNLP, La Plata, Argentina.

Emails: {agustinr, jgarcia, carlosm}@ing.unlp.edu.ar

Abstract—Tracking loops are very often chosen for code and carrier phase estimation in real-time Global Navigation Satellite System receivers due to their low computational complexity. The inputs to these loops are obtained from correlations of the received signal with locally generated replicas. Usually, correlation intervals are chosen synchronously with the data-bits sent with each satellite signal. As a consequence, each loop operates at its own time and the navigation task must extrapolate loop measurements to a common instant. We propose to change this philosophy using a common correlation interval for every satellite signal. In this way, the tracking loops work in synchronism with the navigation process, rather than with the data bits. We show how to account for the occurrence of bit transitions inside a correlation interval and how to derive a suitable discriminator for phase and code errors. The performance of this discriminator is very close to that obtained with the usual bit-synchronous correlations. The proposed scheme was applied to a scalar phase lock loop structure intended for high dynamics Global Navigation Satellite System receivers. The loop is shown to have almost the same tracking threshold and phase estimation quality than those working bit-synchronously. However, control of the measurement instant can produce significant improvements in phase estimation. Moreover, the main contribution of this scheme is for the implementation of real-time vector tracking loops, since it naturally generates a vector of simultaneous measurements in real-time.

Keywords—*GNSS; Real-Time Receivers; Phase Locked Loops; Vector Tracking Loops.*

I. INTRODUCTION

Measuring the propagation delay of the broadcasted signals is the key of the position calculations made in every modern Global Navigation Satellite System (GNSS) receiver. For this purpose, the receiver has to be synchronized with the visible satellite signals. Direct Sequence Spread Spectrum (DS-SS) signals are utilized due to their desired properties of high time-resolution and Code Division Multiple Access (CDMA) and therefore code and carrier synchronization are required [1]. A correlation stage is also needed at the receiver to de-spread the incoming signals so that the synchronization and navigation algorithms can operate with reasonable signal to noise ratios. The required economy of operations in real-time receivers makes impractical the use of complex estimation schemes and usually tracking loop schemes are adopted for synchronization purposes. Phase measurements are considerably less noisy than code delay and so, code loops are usually aided by carrier loops [2]. However, the signal phase is affected by the wavelength ambiguity and hence the basic measurement used for standard position determination is code delay. On the contrary, the techniques used in high precision positioning applications usually take advantage of the phase measurements. In general, code delay and carrier phase or frequency measurements used by the GNSS receiver for position and velocity determination are referred as navigation measurements or raw track data.

Typically, the GNSS signal has also a data structure to send useful information to the receivers, such as orbit parameters needed for satellite position calculations, clock corrections, ionospheric corrections, signal quality indexes, etc. The bits carrying this information are modulated usually in phase, and of course the receiver has to be able to demodulate them. The presence of these data-bits imposes restrictions to the receiver operation from the point of view of navigation measurements generation. Indeed, the correlation time is, in principle, limited to the bit duration time and the corresponding signal to noise ratio increase due to despreading gain is limited too. In some applications this is not a limitation at all, but in others, such us indoor positioning, the use of some long-correlation techniques is unavoidable [3]. Moreover, since different satellite signals experience different propagation delays, the edges of these bits are in general asynchronous. As a consequence, the correlation intervals used for each signal satellite are also asynchronous. In standard real-time receivers, this causes that the tracking loops for each satellite operate synchronously with the bit edges, but asynchronously among them. For the navigation process, this implies that the measurements do not correspond to the same time instant and the receiver has to extrapolate them [1], [2]. These lag differences make it difficult to take advantage of the correlation between the received signals, since each signal is tracked independently.

The convenience of joint tracking the signals by means of the so-called vector tracking loops, has been envisioned since the conception of the GNSS systems [1]. Nowadays, due to their potential advantages together with the growing computation capacity available in a GNSS receiver, many researchers and developers, are considering vector tracking loop schemes. These loops can obtain up to 6 dB of improvement in tracking threshold, in addition to high dynamic capacity, multipath immunity and robustness [4]. Vector tracking loops have been mainly applied in Software-based receivers [5], [6]. Recently, a real-time implementation using Field-Programmable-Gate-Arrays (FPGA) with a fast microprocessor has been reported in [7]. This implementation operates with asynchronous

correlations of the different signals, either extrapolating the navigation measurements or asynchronously incorporating the measurements to the main processing algorithm. Other off-line implementations use data bit removal in order to get simultaneous navigation measurements [8]. In this work, we make a different and novel approach, which is based on the use of synchronous correlations for the received satellite signals so that the navigation measurements are naturally simultaneous. As a consequence, the tracking loops operate asynchronously with respect to the bit edges of the signals and their inputs, i.e., the code and carrier phase errors, have to be calculated for signal intervals with a possible bit-transition inside. Our approach is simple: compute partial correlations before and after the bit edge and calculate a discriminated error based on them. By means of simulations made with scalar tracking loops, we show that this scheme offers some improvements in the measurement quality in high-dynamics conditions, and also that the degradation in the tracking threshold is less than 0.5dB compared with a bit-synchronous loop. This value is completely insignificant compared to the potential gain of using a vector loop in real-time.

The rest of paper is organized as follows. A digital model for the received GNSS signal is presented in Section 2. Since the emphasis on this work is on phase loops, our UFA-PLL scheme will be briefly explained. The proposed phase error discriminator for the correlation periods with possible bit-transition is presented in section 3. The extension of the idea to the code loops, which is straightforward, is also shortly discussed. The bit-asynchronous scheme is applied to a high-dynamic scalar carrier tracking loop and the effects in its performance are analyzed in terms of phase measurements quality and tracking threshold in section 4. Finally, the conclusions and future work lines are given in section 5.

II. Digital Measurements Model

As stated above, the received signal must be correlated with the locally generated replicas for each visible satellite in a GNSS receiver. The complex correlations of the signal from a given satellite with carrier power to noise power spectral density C/N_0 and for the i-th correlation interval of duration T can be expressed as [1]

$$C_i = D_i\sqrt{T\frac{C}{N_0}}\text{sinc}(\Delta f_i)R(\Delta\tau_i)e^{j(\pi\Delta f_i T + \Delta\theta_i)} + n_i \quad (1)$$

where $\Delta\tau_i = \tau_i - \hat{\tau}_i$ is the code delay estimation error, $\Delta f_i = f_i - \hat{f}_i$ the frequency estimation error, both assumed constant during the integration time, and $\Delta\theta_i = \theta_i - \hat{\theta}_i$ the initial phase estimation error. The term n_i is a complex white Gaussian noise sequence with unit variance, $R(\cdot)$ is the code correlation function and $\text{sinc}(x) = \sin(\pi x)/(\pi x)$. This expression assumes that there are binary data bits $D_i = \pm 1$ and that correlations are computed within the same bit period. This type of modulation, i.e., Binary Phase Shift Keying (BPSK), is used in the GPS civil signal and in the data components of composite modernized GNSS signals.

After the acquisition process has been completed, i.e., in tracking conditions [1], code and frequency estimation errors are sufficiently small so that the functions $\text{sinc}(\cdot)$ and $R(\cdot)$ can be approximated by 1. Hence, (1) becomes

$$C_i = I_i + jQ_i = D_i\sqrt{T\frac{C}{N_0}}e^{j\Delta\phi_i} + n_i \quad (2)$$

where we have defined $\Delta\phi_i = \phi_i - \hat{\phi}_i$, with $\phi_i = \pi f_i T + \theta_i$ and $\hat{\phi}_i = \pi\hat{f}_i T + \hat{\theta}_i$. With the help of these sequences the carrier tracking loop can be modeled as a digital single-input single-output (SISO) system. It is important to note that Δf_i and $\Delta\phi_i$ can be interpreted as the average frequency error and average phase error during the correlation interval respectively.

The phase estimation error is obtained from the angle of the complex correlation. In the case of BPSK modulation the phase error must be insensitive to the bit changes and a two quadrant discriminator should be utilized. Then,

$$e_i = \tan^{-1}\left(\frac{Q_i}{I_i}\right) = [\Delta\phi_i + n_{\phi_i}]_\pi \quad (3)$$

where the notation $[\cdot]_\pi$ indicates that its argument is kept within the interval $(-\frac{\pi}{2}, \frac{\pi}{2}]$ by adding or subtracting π as many times as needed. The noise term $n_{\phi i}$ has zero mean and a complicated probability distribution in general. However, in high C/N_0 conditions it can be approximated by a Gaussian distribution with zero mean and variance $1/(2TC/N_0)$.

A. UFA Phase Discriminator

The Unambiguous Frequency Aided (UFA) algorithm uses the frequency error information to correct the non-linearity of a Phase Locked Loop (PLL), instead of adding a Frequency Locked Loop (FLL) to cope with high dynamics. Thus, the advantages of a frequency loop are added to the PLL obtaining the same dynamic tolerance of an FLL but also avoiding cycle slips during tracking [9]. The UFA phase discriminator works correcting the ambiguous values of e_i by adding or subtracting an integer number of π. The correction is such that the difference between successive values of the corrected phase error u_i is less than a quarter of a cycle in magnitude. Then, the corrected phase error estimate, with starting value $u_0 = e_0$, is

$$u_i = e_i - I_\pi(e_i - u_{i-1}) \quad (4)$$

where $I_\pi(x) = x - [x]_\pi$ acts similarly to the integer part function, but with steps at the multiples of π. Created in this way, the sequence phase errors u_i has unambiguous values as long as the loop frequency error is lower than $1/(4T)$ in magnitude, i.e., half of the Nyquist rate from uniform sampling theory. Under this condition, the sequence u_i allows to measure the loop frequency error with a simple difference of successive phase errors, giving to the UFA-PLL the same extra-information that usually has an FLL but not a PLL. In previous works we have also shown that the UFA-PLL has the same noise resistance, and so the same tracking threshold, that an equivalent FLL [10].
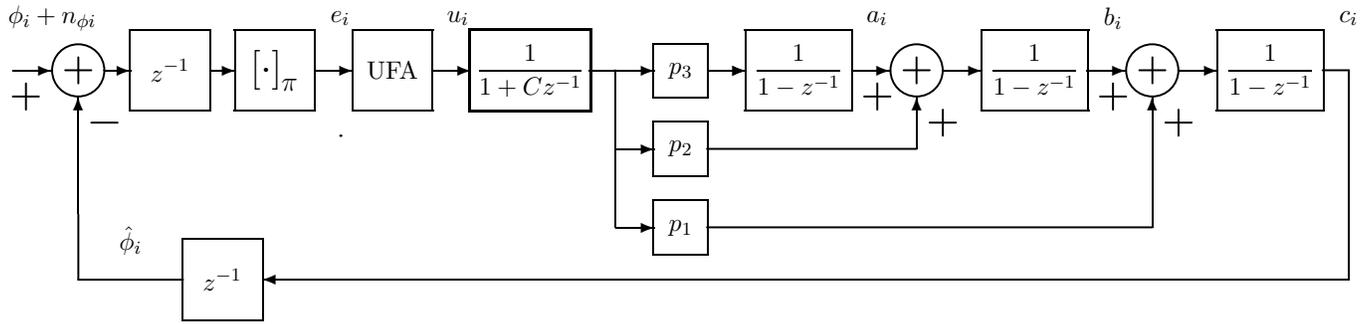
Figure 1. Block diagram of the UFA-PLL model.

III. BIT ASYNCHRONOUS PHASE DISCRIMINATION

Assume the receiver is tracking a given satellite and it knows when a data bit edge will occur during a correlation interval. This requires that a bit synchronization stage has been completed previously. This is not a limitation since the required signal strength for tracking at the high dynamics considered in this work must be high enough to detect bit transitions. For the same reason, multiple data-bits long correlation intervals will not be considered. However, notice that the receiver will not use bit transitions to synchronize the correlation intervals. In our scheme, the receiver uses them to compute the code and phase errors as described in the following and the correlation intervals are dictated by the navigation task. Specifically, assume for the i-th correlation interval of duration T the bit edge will occur T_1 seconds after the beginning and T_2 seconds before its end. Clearly, $T_1 + T_2 = T$. In that case, a coherent correlation of T seconds will not be effective since the possible change of phase will produce a signal cancelation. The worst case when there is a bit reversal is $T_1 = T_2 = T/2$, where a complete signal cancelation occurs. Therefore, the receiver should compute two partial correlations, namely C_1 and C_2. The corresponding phase errors, obtained as in (3), are

$$e_1 = \tan^{-1}\left(\frac{Q_1}{I_1}\right) = [\Delta\phi_1 + n_{\phi_1}]_\pi \quad (5)$$

$$e_2 = \tan^{-1}\left(\frac{Q_2}{I_2}\right) = [\Delta\phi_2 + n_{\phi_2}]_\pi \quad (6)$$

where $\Delta\phi_1 = \Delta\theta_i + \pi\Delta f_i T_1$ and $\Delta\phi_2 = \Delta\theta_i + 2\pi\Delta f_i T_1 + \pi\Delta f_i T_2$ according to the assumed linear evolution of the phase error. Leaving aside for a moment the nonlinearity of the $\tan^{-1}(\cdot)$ function, we can think that these two phase errors are partial averages and therefore they should be averaged to obtain the desired phase error for the i-th correlation interval. The weighted average of them according to the duration of each correlation should be

$$e_i = \frac{T_1}{T}e_1 + \frac{T_2}{T}e_2 \approx \Delta\theta_i + \pi\Delta f_i T + n_{eq} = \Delta\phi_i + n_{eq} \quad (7)$$

with n_{eq} equal to the weighted average of n_{ϕ_1} and n_{ϕ_2}. Under the Gaussian approximation for both noise terms, n_{eq} has a Gaussian distribution with zero mean and variance $1/(2TC/N_0)$. That is, the same variance as if the bit edge was not present. Of course, if T_1 or T_2 are not long enough

the approximation is not valid, and we still have to deal with the nonlinearity of the $\tan^{-1}(\cdot)$ function.

The issue about the nonlinearity is caused by the ambiguity of the phase, indicated by the function $[\cdot]_\pi$. If this operation acts after the weighted average we would obtain a result equivalent to (3) for the correlation interval with a bit transition. However, in (7) the $[\cdot]_\pi$ function actually acted before the average, and then (7) is not correct. Fortunately, the same idea used to build the UFA algorithm can be applied here to test the result and correct it when needed. The hypothesis is that the frequency error is kept under $1/(4T)$ in magnitude. Hence, the signal part of a difference between the partial phase errors in (5) must be bounded. Indeed, (5) can be written as

$$e_1 = \tan^{-1}\left(\frac{Q_1}{I_1}\right) = \Delta\phi_1 + n_{\phi_1} + k_1\pi \quad (8)$$

$$e_2 = \tan^{-1}\left(\frac{Q_2}{I_2}\right) = \Delta\phi_2 + n_{\phi_2} + k_2\pi \quad (9)$$

with $k_1, k_2 \in \mathbb{Z}$. Then, different values of k_1 and k_2 will produce a wrong result at the average (7). This situation has to be detected, and a simple hypothesis test can be built. The decision variable is

$$e_1 - e_2 = \Delta\theta_1 - \Delta\theta_2 + n_d + k_d\pi = \pi\Delta f_i T + n_d + k_d\pi \quad (10)$$

where $n_d = n_{\phi_1} - n_{\phi_2}$ and $k_d = k_1 - k_2$. Since $|\Delta f_i T| < 1/4$ and n_d is a zero mean symmetrically distributed noise term, the optimum decision for the k_d value is $\hat{k}_d = I_\pi(e_1 - e_2)/\pi$. Notice that the possible values for k_d are only three: -1, 0 and 1. Then, if $\hat{k}_d = 0$ no correction is needed and (7) can be applied directly. If $\hat{k}_d \neq 0$ either e_1 or e_2 have to be corrected. Which one is not important since the π ambiguity of the e_i value will be solved later by the UFA algorithm. For simplicity, assume e_2 is corrected when $\hat{k}_d \neq 0$. Then, the final expression for the phase error is

$$e_i = \frac{T_1}{T}e_1 + \frac{T_2}{T}\{e_2 + I_\pi(e_1 - e_2)\}. \quad (11)$$

The computational cost of the new scheme is only one more phase error calculation each time a bit transition could be present, plus the weighted average. The logic needed for the last π ambiguity correction can be neglected compared with the cost of angle calculations and multiplications. Naturally, the idea of the weighted average for combining the discriminated errors from partial correlations in the presence of a
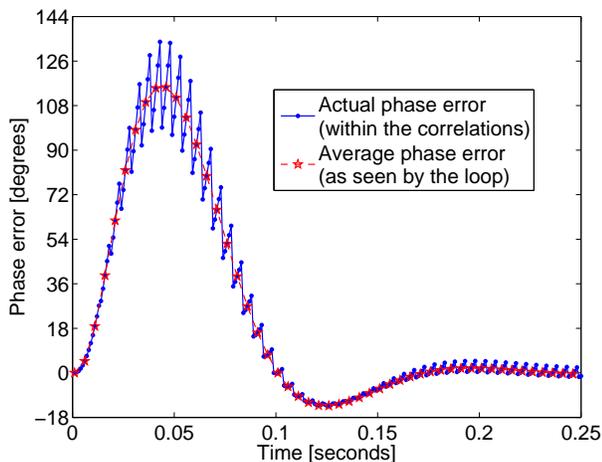
Figure 2.   Phase estimation error during a step of 20g.
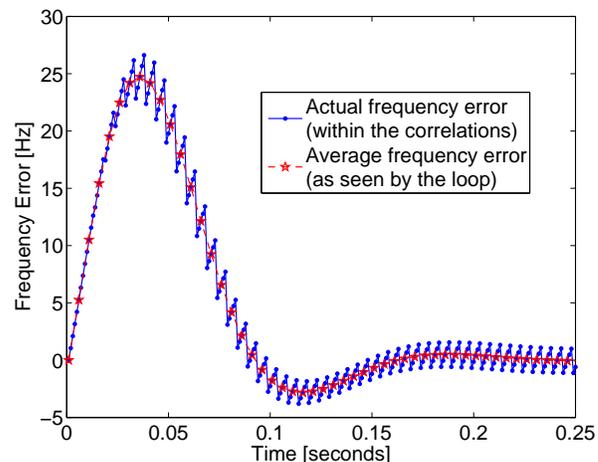
Figure 3.   Frequency estimation error during a step of 20g.

bit transition can be also applied for the case of code delay. Significant changes of the code delay are not expected during a correlation interval, and in the case of high dynamics the code loop is aided by the carrier loop and then only the residual and low code dynamics is tracked. Thus, the code delay error will not evolve significantly during the two partial correlations, and the significant effect will be only to reduce the noise to the same level obtained with a correlation time T. In addition, the delay discriminators are not ambiguous and then no further corrections are needed to the average.

IV. APPLICATION TO A SCALAR CARRIER LOOP

In this section, the proposed bit-asynchronous scheme is applied to a specific carrier tracking loop. We chose a digital UFA-PLL as shown in Figure 1 whose filter coefficients are $C = 0.5$, $p_1 = C = 0.5$, $p_2 = 0.105$, and $p_3 = 0.0123$. For the selected correlation time, $T = 5$ms, the resulting PLL has an equivalent noise bandwidth $B_N = 75.6$Hz. Notice that two delays are included in the loop model. One of them is due to the time spent in computation of the correlation. The other delay appears because the estimated values used to compute the correlations have to be known before the calculations begin. That is, the value $\hat{\phi}_i$ is obtained with the loop filter output of the $(i-1)$-th correlation interval, which in turn is calculated with the estimation errors of $\hat{\phi}_{(i-2)}$. The loop filter is optimized for the tracking of acceleration steps, which produces a quadratic ramp of phase at the loop input. These demanding high dynamics scenarios can be found for example in sounding rockets, at engine turn-on and turn-off. This loop design has been implemented in experimental GPS receivers [9]. According to the analysis made in [11] this design is almost optimal for tracking steps of 20g, in the sense that for a given C/N_0 it approximately produces the smallest pull-out probability.

In order to consider the effects of asynchronous bit transitions without increasing the simulation time excessively, a time step of 1ms was selected. This implies a quantization of the transition times to 5 possible values within a correlation interval of $T = 5$ms. During each correlation interval, partial 1ms correlations are computed according to (1). Then, these values are added to form the two partial correlations if there is a bit transition, or a single correlation if not. The phase estimate used for each middle instant of this period is $\hat{\phi}_i = c_{(i-1)}$, and the frequency estimate for this whole 5 ms period is $\hat{f}_i = b_{(i-1)} + a_{(i-1)}/2$. This frequency estimate was chosen because, as can be seen in the following simulations, it has zero stationary error for acceleration steps. In the following, the influence of the synchronization of the tracking loop with the data-bits in the quality of the navigation measurements is analyzed first. This analysis shows that even for a GNSS receiver with scalar tracking loops, the proposed bit-asynchronous scheme is beneficial. The more important results are given in the second subsection, where the pull-out probabilities of the same loop, operating in a bit-synchronous and in a bit-asynchronous mode are presented.

A. Effects on Phase Estimation Quality

The simulation of the phase evolution with a higher sampling rate than the loop iteration allows us to quantify the quality of phase measurements obtained from the tracking loops. If the loop operates synchronously with the data bits, it cannot be synchronous with the navigation process in general. Therefore, the measurement instants can occur at any instant during a correlation interval, not necessarily in the middle. As an example, consider the phase estimation error produced by the loop for an acceleration step of 20 g without noise, plotted in Figure 2. Clearly, the phase error is not constant during each correlation period. In fact, since the estimated carrier has constant frequency for each period the loop fits the incoming phase with a piecewise linear approximation. Hence, a residual quadratic ramp of phase appears as an estimation error. Notice that the level of error is lower if the relative location within the correlation interval is close to the middle. The same situation is found for the frequency estimation, plotted in Figure 3. Here, the residual error is a linear ramp inside each interval.

As long as pull-out or cycle slips do not occur, the loop response is linear and the noise effects can be analyzed separately. As an example, the loop response to noise only
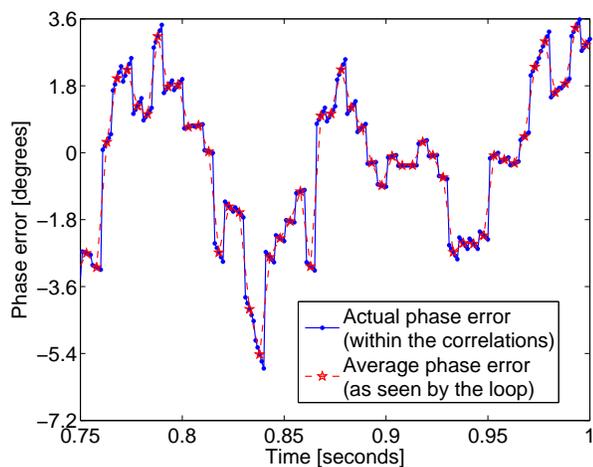
Figure 4. Phase estimation error with $C/N_0 =$ 48 dB/Hz.

Figure 5. Phase estimation error within the correlation interval.

with $C/N_0 = 48$dB/Hz is plotted in Figure 4. In this case, the resultant phase error seems to be lower if the relative location within the correlation interval is close to the beginning. This effect can be understood if we notice that in fact, the loop calculates a carrier prediction for the following correlation interval based on the available measurements. And as the prediction time grows, so does the noise variance of this prediction. Signal and noise results are summarized in Figure 5 where each contribution to the phase standard deviation is plotted for the different relative location within the correlation period. The signal parts correspond to the time average of a 1-second run like that shown in Figure 2. The noise variance was estimated with an average of 1000 runs of 1 second with only Gaussian noise of variance $\frac{1}{2TC/N_0}$ as input, and discarding the first 0.15 seconds to avoid the main part of the initial loop transient. It is possible to verify that the standard deviation for the loop output, i.e., at the middle point, is the same as that obtained with the equivalent noise bandwidth of the loop and the input noise variance. Indeed,

$$\sigma_{\hat{\phi}_i}^2 = \frac{1}{2TC/N_0} \times 2B_N T = \frac{B_N}{C/N_0} \qquad (12)$$

The expression gives $\sigma_{\hat{\phi}_i} = 1.98^\circ$ when $C/N_0 = 48$dB/Hz, as seen in ordinates of Figure 5. In summary, the phase estimation quality changes depending on the time when the navigation measurement is taken during the correlation interval. In the presented example better estimates are obtained if the measurements are taken at the middle, and this can only be done if the loop is synchronous with the navigation process and therefore it is bit-asynchronous. In a bit-synchronous loop, when the measurement instant is taken next to the end of the correlation interval the increase in noise variance is $(3.78^\circ/2.83^\circ)^2 \approx 1.78$, i.e., 78%. Another 1000 runs of the same loop operating according to the bit-asynchronous scheme, with the bit transition location chosen randomly at each run, reveals that the noise output variance is almost not affected, since its increase is less than 0.5% and when plotted looks the same as in Figure 5.
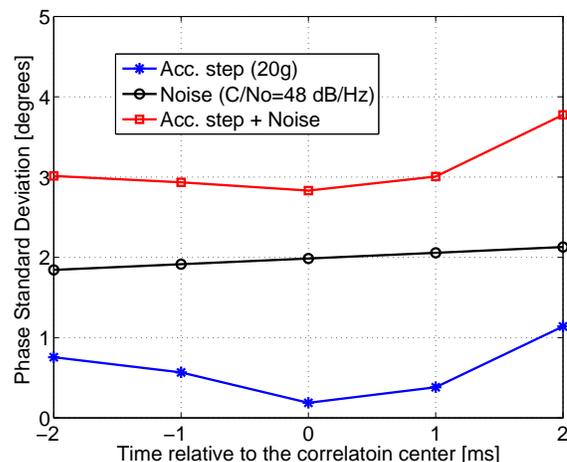
B. Effects on Tracking Threshold

In this section the non-linear performance of the proposed bit-asynchronous scheme for different acceleration and signal levels is determined by means of simulation. Main consequences of this non-linear behavior are cycle slips and pull-out events, i.e., to lose lock with the tracked signal. If a cycle slip occurs, it will produce a loop transient that could end with a pull-out event or not. This temporary loss of phase lock can degrade the data bit demodulation, but as long as the frequency error is low enough useful navigation measurements can be generated. Actually, since the expression (1) was used for the calculation of 1ms correlations used in the simulations, a frequency error also produces a signal power reduction due to the sinc($\cdot$) function factor. Therefore, the adopted criterion to declare a pull-out was that the frequency error exceeds $1/T = 200$ Hz. In this situation, the signal power is completely attenuated and then it can be considered as a practically irrecoverable state. An error of less than 200 Hz is a critical situation but it could still be recoverable. For each value of acceleration and C/N_0, 100,000 runs of 1 second (200 samples) were computed. Each run has an acceleration step of the selected value at the beginning. Two UFA-PLLs were simulated for comparison. One operates synchronously with the data-bits whereas the other implements the proposed scheme according to (11). Using the criteria mentioned above, runs that presented a pull-out event were detected and the pull-out probability (POP) estimated. Since the POP is computed for 1 second of tracking it can be also interpreted as the inverse of the mean-time to lose lock (MTLL) in seconds. The results are presented in Figs. 6 and 7 where POP level curves have been plotted for values of 0.1, 0.2 and so on. Defining the tracking threshold when the POP reaches a level of 0.1, as is usually done, it is interesting to note that the proposed scheme exhibits practically the same threshold than the synchronous loop. Actually, comparing Figs. 7 and 6 it can be clearly seen that the difference between them is always less than 0.5 dB.
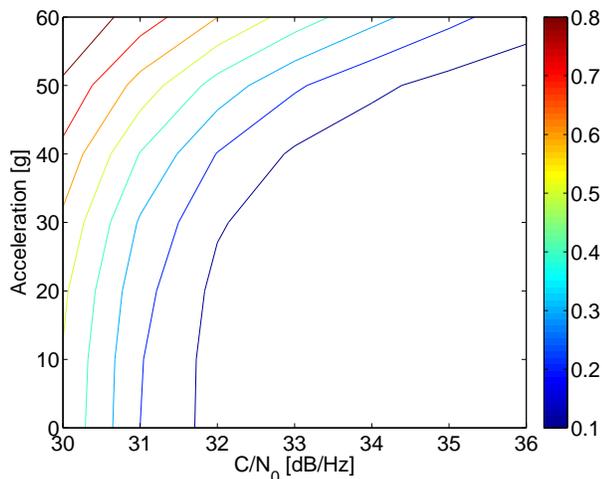
Figure 6.   POP of bit-asynchronous loop and its tracking threshold.
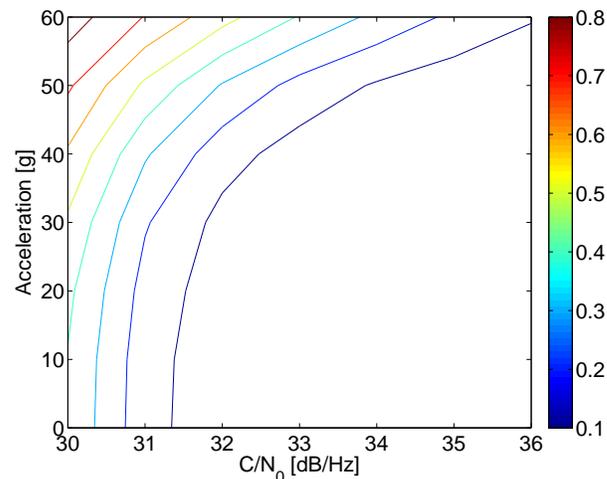
Figure 7.   POP of bit-synchronous loop and its tracking threshold.

V. CONCLUSION AND FUTURE WORK

An efficient bit-asynchronous tracking loop scheme for high performance real-time GNSS receivers has been presented. In the proposed scheme, the loop operation is synchronous with the navigation measurement generation process, rather than with the data-bits. The effect of a possible bit-transition inside the correlation interval is managed by the calculation of two partial correlations. We devised how to build a phase error discriminator based on these partial correlations and explained it in detail. The application of the same procedure for code delay error was briefly discussed. The proposed scheme was applied to a UFA-PLL intended for high dynamic GNSS receivers. We found that it produces an almost negligible impact on tracking threshold (< 0.5 dB) and estimation phase noise ($< 0.5\%$). However, the measurement instants of the tracked signals do not need to be extrapolated to a common instant and therefore a significant improvement can be obtained. An example was shown where a phase noise variance increase of up to 78% can be avoided controlling the measurement instant. Nevertheless, the new scheme main contribution is for the implementation of vector tracking loops in real time, since it will allow operating with a simultaneous vector of measurements from the received satellite signals obtained from correlations computed at a common time.

In terms of computational cost, there is some increase due to the calculation of the partial correlations and ensuing error discrimination. However, the operation of the different satellite tracking loops with the same timing can reduce the processor load depending on the adopted hardware/software architecture. In this case, a correlation stage capable of computing the two partial results for the same receiver estimates when the possible transition location is fed as an extra parameter could be very beneficial for a real-time implementation. The authors are now working on that correlator architecture to be implemented in FPGA, and in the vector tracking loop formulation in real-time with this philosophy.

VI. ACKNOWLEDGMENT

REFERENCES

[1] B. W. Parkinson and J. J. Spilker (eds.), *Global Positioning System: Theory and Applications*.   Washington: American Institute of Aeronautics and Astronautics (AIAA), 1996.

[2] E. D. Kaplan, *Understanding GPS: Principles and Applications*. Boston: Artech House, 1996.

[3] van Diggelen, F., "Indoor GPS theory & implementation," in *Proceedings of The Position Location and Navigation Symposium (PLANS), 2002 IEEE/ION*, Palm Springs, CA , USA, August 2002, pp. 240 – 247.

[4] M. Lashley, D. M. Bevly, and J. Y. Hung, "A valid comparison of vector and scalar tracking loops," in *Proceedings of The Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION*, Indian Wells, CA, USA, May 2010, pp. 464 – 474.

[5] M. Lashley, D. M. Bevly, and J. Y. Hung, "Performance analysis of vector tracking algorithms for weak gps signals in high dynamics," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 4, pp. 661 – 673, August 2009.

[6] S. Zhao, D. Akos, "An Open Source GPS/GNSS Vector Tracking Loop - Implementation, Filter Tuning, and Results," in *Proceedings of the 2011 International Technical Meeting of The Institute of Navigation, ION ITM 2011*, San Diego, CA, USA, January 2011, pp. 1293 – 1305.

[7] W. L. Edwards, M. Lashley, and D. M. Bevly, "FPGA Implementation of a Vector Tracking GPS Receiver using Model-Based Tools," in *Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation,ION GNSS 2009*, Savannah, GA, USA, September 2009, pp. 273 – 280.

[8] T. Lin, J. T. Curran, C. O´Driscoll, and G. Lachapelle, "Implementation of a Navigation Domain GNSS Signal Tracking Loop," in *Proceedings of the 24th International Technical Meeting of The Satellite Division of the Institute of Navigation, ION GNSS 2011*, Portland, Oregon, USA, September 2011, pp. 3644 – 3652.

[9] P. A. Roncagliolo and J. G. García, "High Dynamics and False Lock Resistant GNSS Carrier Tracking Loops," in *Proc. of The 20th Int. Technical Meeting of The Satellite Division of The Institute of Navigation, ION GNSS 2007*, Fort Worth, Texas, September 2007, pp. 2364 – 2375.

[10] P. A. Roncagliolo, J. García, and C. H. Muravchik, "Pull-out Probability and Tracking Threshold Analysis for High Dynamics GNSS Carrier Loops," in *Proceedings of The 21st International Technical Meeting of The Satellite Division of the Institute of Navigation, ION GNSS 2008*, Fort Worth, Texas, USA, September 2008, pp. 221 – 228.

[11] P. A. Roncagliolo, J. G. García, and C.H. Muravchik, "Pull-out Probability Considerations in High Dynamics GNSS Tracking Loops Design," in *Proceedings of The 10th International Symposium on Spread Spectrum Techniques and Applications, ISSSTA 2008*, Bologna, Italy, August 2008, pp. 53 – 57.