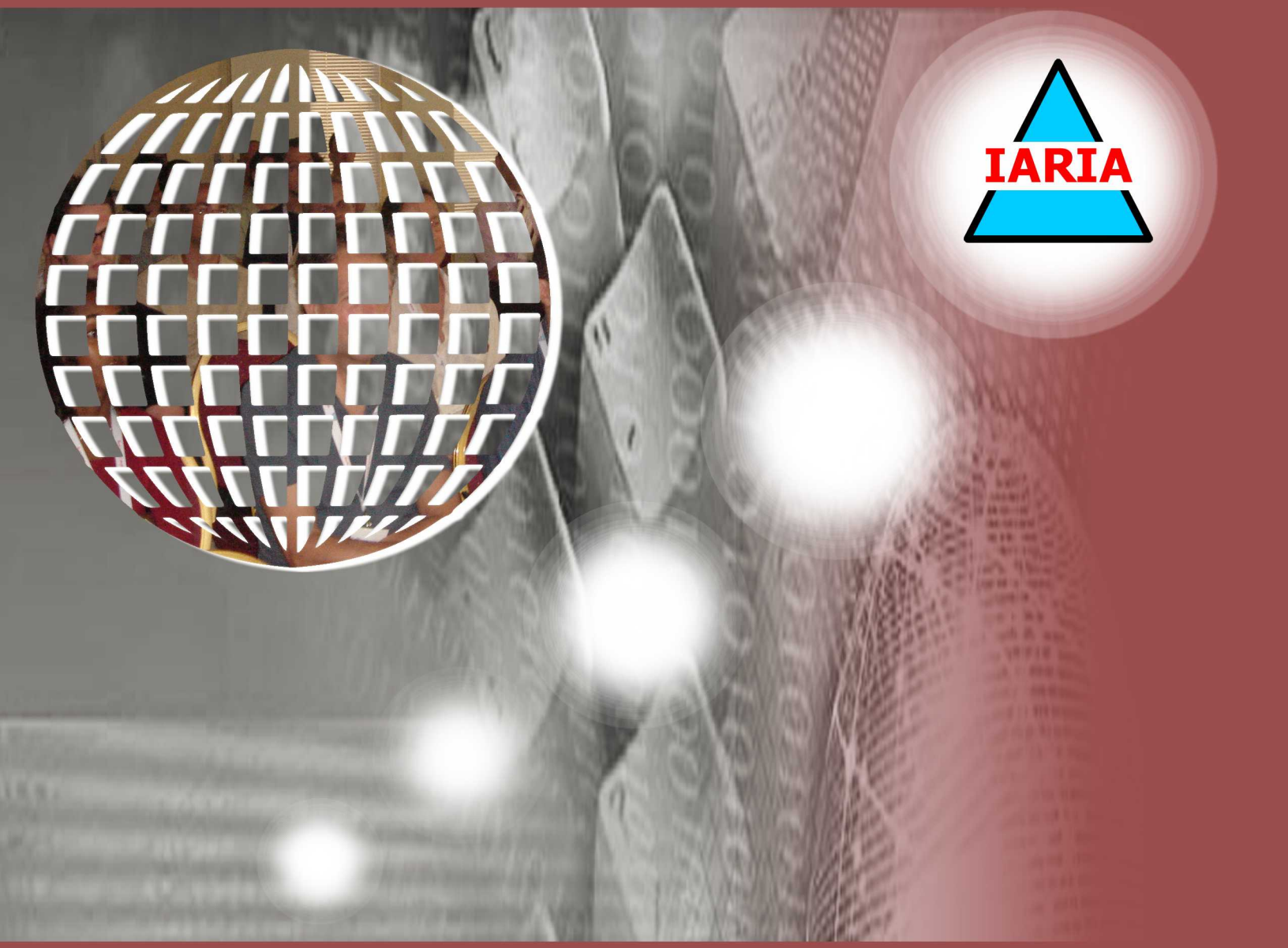
International Journal on Advances in Telecommunications



2014 vol. 7 nr. 1&2

The *International Journal on Advances in Telecommunications* is published by IARIA.
ISSN: 1942-2601
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Telecommunications
Volume 7, Number 1 & 2, 2014

Marco Baldi, Università Politecnica delle Marche, Italy
Ilija Basicevic, University of Novi Sad, Serbia
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Mark Bentum, University of Twente, The Netherlands
David Bernstein, Huawei Technologies, Ltd., USA
Eugen Borgoci, University "Politehnica"of Bucharest (UPB), Romania
Fernando Boronat Seguí, Universidad Politecnica de Valencia, Spain
Christos Bouras, University of Patras, Greece
David Boyle, Tyndall National Institute, University College Cork, Ireland
Martin Brandl, Danube University Krems, Austria
Julien Broisin, IRIT, France
Dumitru Burdescu, University of Craiova, Romania
Andi Buzo, University "Politehnica" of Bucharest (UPB), Romania
Shkelzen Cakaj, Telecom of Kosovo / Prishtina University, Kosovo
Enzo Alberto Candreva, DEIS-University of Bologna, Italy
Rodrigo Capobianco Guido, University of Sao Paulo, Brazil
Hakima Chaouchi, Telecom SudParis, France
Silviu Ciochina, Universitatea Politehnica din Bucuresti, Romania
José Coimbra, Universidade do Algarve, Portugal
Hugo Coll Ferri, Polytechnic University of Valencia, Spain
Denis Collange, Orange Labs, France
Noel Crespi, Institut TELECOM SudParis-Evry, France
Leonardo Dagui de Oliveira, Escola Politécnica da Universidade de São Paulo, Brazil
Gerard Damm, Alcatel-Lucent, USA
Danco Davcev, University for Information Science & Technology "St. Paul the Apostle", Ohrid, Macedonia
Francescantonio Della Rosa, Tampere University of Technology, Finland
Chérif Diallo, Consultant Sécurité des Systèmes d'Information, France
Klaus Drechsler, Fraunhofer Institute for Computer Graphics Research IGD, Germany
Jawad Drissi, Cameron University , USA
António Manuel Duarte Nogueira, University of Aveiro / Institute of Telecommunications, Portugal
Alban Duverdier, CNES (French Space Agency) Paris, France
Nicholas Evans, EURECOM, France
Fabrizio Falchi, ISTI - CNR, Italy
Mário F. S. Ferreira, University of Aveiro, Portugal
Bruno Filipe Marques, Polytechnic Institute of Viseu, Portugal
Robert Forster, Edgemount Solutions, USA
John-Austen Francisco, Rutgers, the State University of New Jersey, USA
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan
Shauneen Furlong , University of Ottawa, Canada / Liverpool John Moores University, UK
Ana-Belén García-Hernando, Universidad Politécnica de Madrid, Spain
Bezalel Gavish, Southern Methodist University, USA
Christos K. Georgiadis, University of Macedonia, Greece
Mariusz Glabowski, Poznan University of Technology, Poland
Katie Goeman, Hogeschool-Universiteit Brussel, Belgium
Hock Guan Goh, Universiti Tunku Abdul Rahman, Malaysia
Pedro Gonçalves, ESTGA - Universidade de Aveiro, Portugal

Valerie Gouet-Brunet, Conservatoire National des Arts et Métiers (CNAM), Paris
Christos Grecos, University of West of Scotland, UK
Stefanos Gritzalis, University of the Aegean, Greece
William I. Grosky, University of Michigan-Dearborn, USA
Vic Grout, Glyndwr University, UK
Xiang Gui, Massey University, New Zealand
Huaqun Guo, Institute for Infocomm Research, A*STAR, Singapore
Song Guo, University of Aizu, Japan
Ching-Hsien (Robert) Hsu, Chung Hua University, Taiwan
Javier Ibanez-Guzman, Renault S.A., France
Lamiaa Fattouh Ibrahim, King Abdul Aziz University, Saudi Arabia
Theodoros Iliou, University of the Aegean, Greece
Mohsen Jahanshahi, Islamic Azad University, Iran
Antonio Jara, University of Murcia, Spain
Carlos Juiz, Universitat de les Illes Balears, Spain
Adrian Kacso, Universität Siegen, Germany
György Kálmán, ABB AS, Norway
Eleni Kaplani, Technological Educational Institute of Patras, Greece
Behrouz Khoshnevis, University of Toronto, Canada
Ki Hong Kim, ETRI: Electronics and Telecommunications Research Institute, Korea
Atsushi Koike, Seikei University, Japan
Ousmane Kone, UPPA - University of Bordeaux, France
Dragana Krstic, University of Nis, Serbia
Archana Kumar, Delhi Institute of Technology & Management, Haryana, India
Romain Laborde, University Paul Sabatier (Toulouse III), France
Massimiliano Laddomada, Texas A&M University-Texarkana, USA
Thomas D. Lagkas, University of Western Macedonia, Greece
Wen-Hsing Lai, National Kaohsiung First University of Science and Technology, Taiwan
Zhihua Lai, Ranplan Wireless Network Design Ltd., UK
Jong-Hyouk Lee, INRIA, France
Wolfgang Leister, Norsk Regnesentral, Norway
Elizabeth I. Leonard, Naval Research Laboratory - Washington DC, USA
Jia-Chin Lin, National Central University, Taiwan
Chi (Harold) Liu, IBM Research - China, China
Haibin Liu, China Aerospace Science and Technology Corporation, China
Diogo Lobato Acatauassu Nunes, Federal University of Pará, Brazil
Andreas Loeffler, Friedrich-Alexander-University of Erlangen-Nuremberg, Germany
Michael D. Logothetis, University of Patras, Greece
Renata Lopes Rosa, University of São Paulo, Brazil
Hongli Luo, Indiana University Purdue University Fort Wayne, USA
Christian Maciocco, Intel Corporation, USA
Dario Maggiorini, University of Milano, Italy
Maryam Tayefeh Mahmoudi, Research Institute for ICT, Iran
Krešimir Malarić, University of Zagreb, Croatia
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Herwig Mannaert, University of Antwerp, Belgium

Adrian Matei, Orange Romania S.A, part of France Telecom Group, Romania
Natarajan Meghanathan, Jackson State University, USA
Emmanouel T. Michailidis, University of Piraeus, Greece
Ioannis D. Moscholios, University of Peloponnese, Greece
Djafar Mynbaev, City University of New York, USA
Pubudu N. Pathirana, Deakin University, Australia
Christopher Nguyen, Intel Corp., USA
Lim Nguyen, University of Nebraska-Lincoln, USA
Brian Niehöfer, TU Dortmund University, Germany
Serban Georgica Obreja, University Politehnica Bucharest, Romania
Peter Orosz, University of Debrecen, Hungary
Patrik Österberg, Mid Sweden University, Sweden
Harald Øverby, ITEM/NTNU, Norway
Tudor Palade, Technical University of Cluj-Napoca, Romania
Constantin Paleologu, University Politehnica of Bucharest, Romania
Stelios Papaharalabos, National Observatory of Athens, Greece
Gerard Parr, University of Ulster Coleraine, UK
Ling Pei, Finnish Geodetic Institute, Finland
Jun Peng, University of Texas - Pan American, USA
Cathryn Peoples, University of Ulster, UK
Dionysia Petraki, National Technical University of Athens, Greece
Dennis Pfisterer, University of Luebeck, Germany
Timothy Pham, Jet Propulsion Laboratory, California Institute of Technology, USA
Roger Pierre Fabris Hoefel, Federal University of Rio Grande do Sul (UFRGS), Brazil
Przemyslaw Pochec, University of New Brunswick, Canada
Anastasios Politis, Technological & Educational Institute of Serres, Greece
Adrian Popescu, Blekinge Institute of Technology, Sweden
Neeli R. Prasad, Aalborg University, Denmark
Dušan Radović, TES Electronic Solutions, Stuttgart, Germany
Victor Ramos, UAM Iztapalapa, Mexico
Gianluca Reali, Università degli Studi di Perugia, Italy
Eric Renault, Telecom SudParis, France
Leon Reznik, Rochester Institute of Technology, USA
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
David Sánchez Rodríguez, University of Las Palmas de Gran Canaria (ULPGC), Spain
Panagiotis Sarigiannidis, University of Western Macedonia, Greece
Michael Sauer, Corning Incorporated, USA
Reijo Savola, VTT Technical Research Centre of Finland, Finland
Marialisa Scatà, University of Catania, Italy
Zary Segall, Chair Professor, Royal Institute of Technology, Sweden
Sergei Semenov, Broadcom, Finland
Sandra Sendra Compte, Polytechnic University of Valencia, Spain
Dimitrios Serpanos, University of Patras and ISI/RC Athena, Greece
Adão Silva, University of Aveiro / Institute of Telecommunications, Portugal
Pushpendra Bahadur Singh, MindTree Ltd, India
Danai Skournetou, Tampere University of Technology, Finland

Mariusz Skrocki, Orange Labs Poland / Telekomunikacja Polska S.A., Poland
Leonel Sousa, INESC-ID/IST, TU-Lisbon, Portugal
Liana Stanescu, University of Craiova, Romania
Cosmin Stoica Spahiu, University of Craiova, Romania
Radu Stoleru, Texas A&M University, USA
Young-Joo Suh, POSTECH (Pohang University of Science and Technology), Korea
Hailong Sun, Beihang University, China
Jani Suomalainen, VTT Technical Research Centre of Finland, Finland
Fatma Tansu, Eastern Mediterranean University, Cyprus
Ioan Toma, STI Innsbruck/University Innsbruck, Austria
Božo Tomas, HT Mostar, Bosnia and Herzegovina
Piotr Tyczka, Poznan University of Technology, Poland
John Vardakas, University of Patras, Greece
Andreas Veglis, Aristotle University of Thessaloniki, Greece
Luís Veiga, Instituto Superior Técnico / INESC-ID Lisboa, Portugal
Calin Vladeanu, "Politehnica" University of Bucharest, Romania
Natalija Vlajic, York University - Toronto, Canada
Benno Volk, ETH Zurich, Switzerland
Krzysztof Walczak, Poznan University of Economics, Poland
Krzysztof Walkowiak, Wroclaw University of Technology, Poland
Yang Wang, Georgia State University, USA
Yean-Fu Wen, National Taipei University, Taiwan, R.O.C.
Bernd E. Wolfinger, University of Hamburg, Germany
Riaan Wolhuter, Universiteit Stellenbosch University, South Africa
Yulei Wu, Chinese Academy of Sciences, China
Mudasser F. Wyne, National University, USA
Gaoxi Xiao, Nanyang Technological University, Singapore
Bashir Yahya, University of Versailles, France
Abdulrahman Yarali, Murray State University, USA
Mehmet Erkan Yüksel, Istanbul University, Turkey
Pooneh Bagheri Zadeh, Staffordshire University, UK
Giannis Zaoudis, University of Patras, Greece
Liaoyuan Zeng, University of Electronic Science and Technology of China, China
Rong Zhao , Detecon International GmbH, Germany
Zhiwen Zhu, Communications Research Centre, Canada
Martin Zimmermann, University of Applied Sciences Offenburg, Germany
Piotr Zwierzykowski, Poznan University of Technology, Poland

International Journal on Advances in Telecommunications
Volume 7, Numbers 1 & 2, 2014

CONTENTS

Second Order Statistics of SSC/SC Combiner Operating Over Rician Fading Channel

Dragana Krstić
Faculty of Electronic Engineering, University of Niš
Niš, Serbia
dragana.krstic@elfak.ni.ac.rs

Petar Nikolić
Tigartyres, Pirot, Serbia
nikpetar@gmail.com

***Abstract*—The second order statistics, such as average level crossing rate (LCR) and average fade duration (AFD), for dual branch SSC/SC combiner output signals in the presence of Rician fading, are determined in this paper by using earlier obtained closed-form expressions for probability density functions (PDFs) of derivatives at two time instants. The results are graphically presented in some figures versus some parameters values. The analysis of the parameters influence and different types of combiners is given.**

Keywords- Average Fade Duration, Level Crossing Rate, Probability Density Function, Rician Fading, Selection Combining, Switch and Stay Combining, Time Derivative.

I. INTRODUCTION

The fading is one of the most important causes of system performance degradation in wireless communication systems. The communication systems are subjected to fading which is caused by multipath propagation due to reflection, refraction and scattering by buildings, trees and other large structures. After that received signal represents a sum of many signals that arrive via different propagation paths.

Some statistical models are used in the literature to describe the fading envelope of the received signal. These distributions are: Rayleigh, Rician, Nakagami, Weibull, Hoyt, and others. They are used to characterize the envelope of faded signals over small geographical area, or short term fading. The log-normal and gamma distribution are used for describing long term fading, when much wider geographical area is involved.

The performance analysis of complex SSC/MRC combiner in Rice fading channel is given in [1]. This paper presents an extended analysis of the second order statistics, such as average level crossing rate (LCR) and average fade duration (AFD), for dual branch SSC/SC combiner output signals in the presence of Rician fading, based on closed-form expressions for probability density functions (PDFs) of derivatives at two time instants obtained in [2].

Ricean fading is a stochastic model for radio propagation irregularity caused by partial cancellation of a radio signal, i.e., the signal arrives at the receiver by several different paths and at least one of them is lengthened or shorted [3]. Rician fading occurs when one of the paths, typically a line of sight signal, is much stronger than the others. In Rician fading, the amplitude gain is characterized by a Rician distribution [4], [5]. Rayleigh fading is the specialized model for stochastic fading when there is no line of sight signal, and it is considered as a special case of the more generalized concept of Rician fading. In Rayleigh fading, the amplitude is described by a Rayleigh distribution.

There are several ways to reduce fading influence on system performances without increasing the signal power and channel bandwidth. The diversity reception techniques are used extensively in fading radio channels to reduce the fading influence on system performances [5]. Various diversity combining techniques are used.

Selection combining (SC), where the strongest signal is selected between the N received signals [6]. When the N signals are independent and Rayleigh distributed, the expected diversity gain has been shown to be inversely proportional to the number of antennas [7, 8].

Switched combining: In the case of dual branch SSC, the first branch stay selected as long as its current value of signal-to-noise ratio (SNR) is greater than predetermined switching threshold, even if the SNR value in the second branch maybe is largerer at that time [3]. The receiver switches to another signal when the currently selected signal drops below a predefined threshold [9]. This is a less efficient technique than selection combining.

The consideration of SSC systems in the literature has been restricted to low-complexity mobile units where the number of diversity antennas is typically limited to two. Performance analysis of SSC diversity receiver over correlated Ricean fading channels in the presence of co-channel interference is carried out in [10].

Equal-gain combining (EGC): All the received signals are summed coherently [11].

Maximal-ratio combining (MRC) is often used in large phased-array systems. The received signals are weighted with respect to their SNR and then summed [12]. This is the best and most complicated combining scheme.

In this paper, the probability density functions (PDFs) of derivatives at two time instants for dual branch Switch and Stay (SSC) combiner output signals in the presence of Rician fading in closed-form expressions are presented and used for calculating the second order statistics of SSC/SC combiner. The results are shown in some graphs versus different parameters values. An analysis of the results is provided also. To the best author knowledge the

performance of SSC/SC combiner is not reported in open technical literature by other authors.

The remainder of the document is organized as follows: in Section II related works are mentioned; Section III introduces the model of the SSC combiner; the probability density functions of derivatives are presented and graphically shown in Section IV. In Section V, the model of complex SSC/SC combiner is given and statistics for this combiner is calculated. In the last section, some conclusions are presented.

II. Related Works

The probability density functions (PDFs) of derivatives for Switch and Stay Combiner (SSC) output signals at two time instants in the presence of Rician fading are determined in [2]. Now, in this paper, the second-order characteristics will be determined using these PDFs.

The probability density functions and joint probability density functions for SSC combiner output signals at two time instants in the presence of different fading distributions are determined by these authors. Then, they are used for calculating first order system performances, such as the bit error rate and the outage probability, for complex systems sampling at two time instants. Performance analysis of SSC/SC combiner in the presence of Rayleigh fading is given in [13]. The outage probability analysis of the SSC/SC combiner at two time instants in the presence of lognormal fading is done in [14]. The PDF of the combiner output signal is derived. Then, the outage probability is numerically calculated using this PDF. The results are shown graphically in order to compare performances of the SSC/SC combiner with regard to classical SSC and SC combiners sampling at one time instant.

This work is motivated by the desire to obtain better system performance, in the presence of Rician fading, with complex combiner consisting of two cheaper and simpler diversity systems (SSC and SC) against MRC diversity combining scheme witch is much more expensive.

III. System Model of SSC combiner

The SSC combiner with two branches at two time instants is discussed in this section. The model is shown in Fig. 1.

The input signals, at the first time moment, are r_{11} and r_{21} and they are r_{12} and r_{22} at the second time instant. The derivatives are $\dot{r}_{11}$ and $\dot{r}_{21}$ at the first time instant and $\dot{r}_{12}$ and $\dot{r}_{22}$ at the second one.

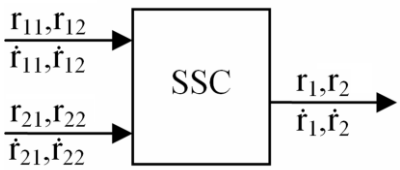


Figure 1. Model of the SSC combiner with two inputs at two time instants

The signals at the output are r_1 and r_2. The derivatives of the SSC combiner output signals are $\dot{r}_1$ and $\dot{r}_2$.

The indices for input signals and their derivatives are as follows: the first index represents the branch ordinal number and the other one signs the time instant observed. The indices for the output signal correspond to the time instants observed.

The probability that combiner examines first the signal from the first branch is P_1 and P_2 for the second one. The values of P_1 and P_2 for SSC combiner are obtained in [3].

The four different cases are discussed here, according to the working algorithm of SSC combiner:

1) $r_1<r_T, r_2<r_T$

In this case all signals are less then threshold r_T, i.e.: $r_{11}<r_T$, $r_{12}<r_T$, $r_{21}<r_T$, and $r_{22}<r_T$. Let combiner considers first the signal r_{11}. Because $r_{11}<r_T$, then $\dot{r}_1 = \dot{r}_{21}$, and because of $r_{22}<r_T$, then $\dot{r}_2 = \dot{r}_{12}$. The probability of this event is P_1. If combiner examines first the signal r_{21}, then $r_{21}<r_T$, $\dot{r}_1 = \dot{r}_{11}$, as $r_{21}<r_T$, $\dot{r}_2 = \dot{r}_{22}$. The probability of this event is P_2.

2) $r_1\geq r_T, r_2<r_T$

In this case first of the signals is greater than the threshold r_T, but the other is less. The possible combinations related to the probability of the first examination of first or second input are:

- $r_{11}\geq r_T$, $r_{12}<r_T$, $r_{22}<r_T$, $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{22}$ P_1
- $r_{11}<r_T$, $r_{21}\geq r_T$ $r_{22}<r_T$, $r_{12}<r_T$, $\dot{r}_1 = \dot{r}_{21}$ $\dot{r}_2 = \dot{r}_{12}$ P_1
- $r_{21}\geq r_T$, $r_{22}<r_T$, $r_{12}<r_T$, $\dot{r}_1 = \dot{r}_{21}$ $\dot{r}_2 = \dot{r}_{12}$ P_2
- $r_{21}<r_T$, $r_{11}\geq r_T$, $r_{12}<r_T$, $r_{22}<r_T$, $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{22}$ P_2

3) $r_1<r_T, r_2\geq r_T$

In this case first of the signals is less then threshold r_T, but the other is bigger. The possible combinations for this case tied with probabilities of order of inputs' consideration are:

- $r_{11}<r_T$, $r_{21}<r_T$, $r_{22}\geq r_T$, $\dot{r}_1 = \dot{r}_{21}$ $\dot{r}_2 = \dot{r}_{22}$ P_1
- $r_{11}<r_T$, $r_{21}<r_T$, $r_{22}<r_T$, $r_{12}\geq r_T$, $\dot{r}_1 = \dot{r}_{21}$ $\dot{r}_2 = \dot{r}_{12}$ P_1
- $r_{21}<r_T$, $r_{11}<r_T$, $r_{12}\geq r_T$, $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{12}$ P_2
- $r_{21}<r_T$, $r_{11}<r_T$, $r_{12}<r_T$, $r_{22}\geq r_T$, $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{22}$ P_2

4) $r_1\geq r_T, r_2\geq r_T$

In the last case all signals are greater then threshold r_T, i.e.: $r_{11}\geq r_T$, $r_{12}\geq r_T$, $r_{21}\geq r_T$, and $r_{22}\geq r_T$. Now, the possible combinations of probabilities of inputs' examinations are:

- $r_{11}\geq r_T$, $r_{12}\geq r_T$, $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{12}$ P_1
- $r_{11}\geq r_T$, $r_{12}<r_T$, $r_{22}\geq r_T$ $\dot{r}_1 = \dot{r}_{11}$ $\dot{r}_2 = \dot{r}_{22}$ P_1

- $r_{11}<r_T$, $r_{21}\geq r_T$, $r_{22}\geq r_T$, $\dot{r}_1=\dot{r}_{21}$ $\dot{r}_2=\dot{r}_{22}$ P_1
- $r_{11}<r_T$, $r_{21}\geq r_T$, $r_{22}<r_T$, $r_{12}<r_T$ $\dot{r}_1=\dot{r}_{21}$ $\dot{r}_2=\dot{r}_{12}$ P_1
- $r_{21}\geq r_T$, $r_{22}\geq r_T$, $\dot{r}_1=\dot{r}_{21}$ $\dot{r}_2=\dot{r}_{22}$ P_2
- $r_{21}\geq r_T$, $r_{22}<r_T$, $r_{12}\geq r_T$, $\dot{r}_1=\dot{r}_{21}$ $\dot{r}_2=\dot{r}_{12}$ P_2
- $r_{21}<r_T$, $r_{11}\geq r_T$, $r_{12}\geq r_T$, $\dot{r}_1=\dot{r}_{11}$ $\dot{r}_2=\dot{r}_{12}$ P_2
- $r_{21}<r_T$, $r_{11}\geq r_T$, $r_{12}<r_T$, $r_{22}\geq r_T$, $\dot{r}_1=\dot{r}_{11}$ $\dot{r}_2=\dot{r}_{22}$ P_2

IV. Probability Density Functions of Derivatives

The joint probability density functions of signal derivatives are:

$r_1<r_T$, $r_2<r_T$

$$p_{r_1r_2\dot{r}_1\dot{r}_2}(r_1,r_2,\dot{r}_1,\dot{r}_2)=P_1\int_0^{r_T}dr_{11}\int_0^{r_T}dr_{22}\,p_{r_{11}r_{22}r_{21}\dot{r}_{12}\dot{r}_{21}\dot{r}_{12}}(r_{11},r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\int_0^{r_T}dr_{12}\,p_{r_{21}r_{12}\dot{r}_{11}\dot{r}_{22}\dot{r}_{11}\dot{r}_{22}}(r_{21},r_{12},r_1,r_2,\dot{r}_1,\dot{r}_2) \quad (1)$$

$r_1\geq r_T$, $r_2<r_T$

$$p_{r_1r_2\dot{r}_1\dot{r}_2}(r_1,r_2,\dot{r}_1,\dot{r}_2)=P_1\int_0^{r_T}dr_{12}\,p_{r_1r_2\dot{r}_1\dot{r}_2\dot{r}_1\dot{r}_{22}}(r_{12},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{11}\int_0^{r_T}dr_{22}\,p_{r_{11}r_{22}r_2\dot{r}_{11}\dot{r}_2\dot{r}_{12}}(r_{11},r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{22}\,p_{r_{22}r_2\dot{r}_{12}\dot{r}_2\dot{r}_{12}}(r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\int_0^{r_T}dr_{12}\,p_{r_{21}r_{12}\dot{r}_{11}\dot{r}_{22}\dot{r}_{11}\dot{r}_{22}}(r_{21},r_{12},r_1,r,\dot{r}_1,\dot{r}_2) \quad (2)$$

$r_1<r_T$, $r_2\geq r_T$

$$p_{r_1r_2\dot{r}_1\dot{r}_2}(r_1,r_2,\dot{r}_1,\dot{r}_2)=P_1\int_0^{r_T}dr_{11}\,p_{r_1r_2r_{22}\dot{r}_2\dot{r}_{22}}(r_{11},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{11}\int_0^{r_T}dr_{22}\,p_{r_{11}r_{22}r_2\dot{r}_{11}\dot{r}_2\dot{r}_{12}}(r_{11},r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\,p_{r_{21}r_1\dot{r}_{11}\dot{r}_{12}\dot{r}_{11}\dot{r}_{12}}(r_{21},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\int_0^{r_T}dr_{12}\,p_{r_{21}r_{12}\dot{r}_{11}\dot{r}_{22}\dot{r}_{11}\dot{r}_{22}}(r_{21},r_{12},r_1,r,\dot{r}_1,\dot{r}_2) \quad (3)$$

$r_1\geq r_T$, $r_2\geq r_T$

$$p_{r_1r_2\dot{r}_1\dot{r}_2}(r_1,r_2,\dot{r}_1,\dot{r}_2)=P_1\,p_{r_{11}r_{12}\dot{r}_{11}\dot{r}_{12}}(r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{12}\,p_{r_{12}r_{11}r_{22}\dot{r}_{11}\dot{r}_{22}}(r_{12},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{11}\,p_{r_{11}r_{12}r_{22}\dot{r}_{12}\dot{r}_{22}\dot{r}_{12}}(r_{11},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{11}\int_0^{r_T}dr_{22}\,p_{r_{11}r_{22}r_1r_{12}\dot{r}_{21}\dot{r}_{12}}(r_{11},r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\,p_{r_{21}r_{22}\dot{r}_{21}\dot{r}_{22}}(r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{22}\,p_{r_{22}r_{21}r_{12}\dot{r}_{21}\dot{r}_{12}}(r_{22},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\,p_{r_{21}r_{11}r_{12}\dot{r}_{11}\dot{r}_{12}}(r_{21},r_1,r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\int_0^{r_T}dr_{12}\,p_{r_{21}r_{12}r_{11}r_{22}\dot{r}_{11}\dot{r}_{22}}(r_{21},r_{12},r_1,r,\dot{r}_1,\dot{r}_2) \quad (4)$$

For the case that signal and its derivative are not correlated, after integrating of the whole range of signal values and some mathematical manipulations, the joint PDF of derivative can be expressed as:

$$p_{\dot{r}_1\dot{r}_2}(\dot{r}_1,\dot{r}_2)=P_1\int_0^{r_T}dr_{11}\int_0^{r_T}dr_{22}\,p_{r_{11}r_{22}\dot{r}_{21}\dot{r}_{12}}(r_{11},r_{22},\dot{r}_1,\dot{r}_2)+$$
$$+P_2\int_0^{r_T}dr_{21}\int_0^{r_T}dr_{12}\,p_{r_{21}r_{12}\dot{r}_1\dot{r}_{22}}(r_{21},r_{12},\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{12}\int_{r_T}^{\infty}dr_1\,p_{r_{12}r_1\dot{r}_1\dot{r}_{22}}(r_{12},r_1,\dot{r}_1,\dot{r}_2)+P_2\int_0^{r_T}dr_{22}\int_{r_T}^{\infty}dr_1\,p_{r_{22}r_{21}\dot{r}_{21}\dot{r}_{12}}(r_{22},r_1,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_0^{r_T}dr_{11}\int_{r_T}^{\infty}dr_2\,p_{r_1r_{22}\dot{r}_2\dot{r}_{22}}(r_{11},r_2,\dot{r}_1,\dot{r}_2)+P_2\int_0^{r_T}dr_{21}\int_{r_T}^{\infty}dr_2\,p_{r_{21}r_1\dot{r}_1\dot{r}_{12}}(r_{21},r_2,\dot{r}_1,\dot{r}_2)+$$
$$+P_1\int_{r_T}^{\infty}dr_1\int_{r_T}^{\infty}dr_2\,p_{r_1r_2\dot{r}_1\dot{r}_{12}}(r_1,r_2,\dot{r}_1,\dot{r}_2)+P_2\int_{r_T}^{\infty}dr_1\int_{r_T}^{\infty}dr_2\,p_{r_{21}r_{22}\dot{r}_{21}\dot{r}_{22}}(r_1,r_2,\dot{r}_1,\dot{r}_2) \quad (5)$$

The signal derivatives PDFs can be found from joint PDF based on:

$$p_{\dot{r}_1}(\dot{r}_1)=\int_{-\infty}^{\infty}p_{\dot{r}_1\dot{r}_2}(\dot{r}_1,\dot{r}_2)d\dot{r}_2 \quad (6)$$

$$p_{\dot{r}_2}(\dot{r}_2)=\int_{-\infty}^{\infty}p_{\dot{r}_1\dot{r}_2}(\dot{r}_1,\dot{r}_2)d\dot{r}_1 \quad (7)$$

By replacing (5) in (6) and (7), it is obtained:

$$p_{\dot{r}_1}(\dot{r}_1)=P_1p_{\dot{r}_{11}}(\dot{r}_1)+P_2p_{\dot{r}_{21}}(\dot{r}_1)+$$
$$+\left(P_2F_{r_{21}}(r_T)-P_1F_{r_{11}}(r_T)\right)p_{\dot{r}_{11}}(\dot{r}_1)+$$
$$+\left(P_1F_{r_{11}}(r_T)-P_2F_{r_{21}}(r_T)\right)p_{\dot{r}_{21}}(\dot{r}_1) \quad (8)$$

$$p_{\dot{r}_2}(\dot{r}_2) = P_1 F_{r_{11}}(r_T) F_{r_{22}}(r_T) p_{\dot{r}_{12}}(\dot{r}_2) + P_2 F_{r_{21}}(r_T) F_{r_{12}}(r_T) p_{\dot{r}_{22}}(\dot{r}_2) +$$
$$+ P_1 B_1(r_T) p_{\dot{r}_{22}}(\dot{r}_2) + P_2 B_2(r_T) p_{\dot{r}_{12}}(\dot{r}_2) +$$
$$+ P_1 F_{r_{11}}(r_T)\left(1 - F_{r_{22}}(r_T)\right) p_{\dot{r}_{22}}(\dot{r}_2) + P_2 F_{r_{21}}(r_T)\left(1 - F_{r_{12}}(r_T)\right) p_{\dot{r}_{12}}(\dot{r}_2) +$$
$$+ P_1 C_1(r_T) p_{\dot{r}_{12}}(\dot{r}_2) + P_2 C_2(r_T) p_{\dot{r}_{22}}(\dot{r}_2) \quad (9)$$

where $F_{r_{ij}}(r_T)$ are signals' CDFs and $F_{r_{i1}}(r_T) = F_{r_{i2}}(r_T)$, while $B_i(r_T)$ and $C_i(r_T)$ are obtained based on [(1), 15]

$$B_i(r_T) = \left(1 - \rho_i^2\right) e^{-\frac{A_i^2}{\sigma_i^2(1+\rho_i)}} \cdot$$
$$\cdot \sum_{k,l_1,l_2,l_3=0}^{\infty} \varepsilon_k \cdot \frac{1}{l_1! l_2! l_3! (k+l_1)! (k+l_2)! (k+l_3)!} \cdot$$
$$\cdot \rho^{k+2l_1} \left[\left(\frac{1-\rho_i}{1+\rho_i}\right)\left(\frac{A_i}{2\sigma_i^2}\right)\right]^{k+l_2+l_3} \cdot$$
$$\cdot \gamma\left(k + l_1 + l_2 + 1, \frac{r_T^2}{2\sigma_i^2(1-\rho_i^2)}\right) \cdot$$
$$\cdot \left[1 - \gamma\left(k + l_1 + l_3 + 1, \frac{r_T^2}{2\sigma_i^2(1-\rho_i^2)}\right)\right] \quad (10)$$

$$C_i(r_T) = \left(1 - \rho_i^2\right) e^{-\frac{A_i^2}{\sigma_i^2(1+\rho_i)}} \cdot$$
$$\cdot \sum_{k,l_1,l_2,l_3=0}^{\infty} \varepsilon_k \cdot \frac{1}{l_1! l_2! l_3! (k+l_1)! (k+l_2)! (k+l_3)!} \cdot$$
$$\cdot \rho^{k+2l_1} \left[\left(\frac{1-\rho_i}{1+\rho_i}\right)\left(\frac{A_i}{2\sigma_i^2}\right)\right]^{k+l_2+l_3} \cdot$$
$$\cdot \left[1 - \gamma\left(k + l_1 + l_2 + 1, \frac{r_T^2}{2\sigma_i^2(1-\rho_i^2)}\right)\right] \cdot$$
$$\cdot \left[1 - \gamma\left(k + l_1 + l_3 + 1, \frac{r_T^2}{2\sigma_i^2(1-\rho_i^2)}\right)\right] \quad (11)$$

$\gamma(.)$ is incomplete gamma function and ε_k is Neumman factor defined with

$$\varepsilon_k = \begin{cases} 1, & k = 0 \\ 2, & k > 0 \end{cases}$$

The probability density functions of signal derivatives in the presence of Rician fading at the combiner input has normal distribution with zero mean value [16, 17]:

$$p_{\dot{r}_i}(\dot{r}_{i,j}) = \frac{1}{\sqrt{2\pi}\dot{\sigma}_i} e^{-\frac{\dot{r}_{i,j}^2}{2\dot{\sigma}_i^2}}, \quad -\infty < \dot{r}_{i,j} < \infty \quad (12)$$

where i=1,2; j=1,2 and $\dot{\sigma}_i^2 = 2\sigma_i^2\pi^2 f_m^2$ is the variance and f_m is maximal Doppler frequency.

Probability density function of signal derivatives $\dot{r}_1$ and $\dot{r}_2$ at the SSC combiner output at two time moments in the presence of Rician fading is obtained when (12) putting in previously obtained general expressions for PDFs of signal derivatives and replacing of CDF with [18]:

$$F_{r_i}(r_{i,j}) = 1 - Q_1\left(A_i/\sigma_i, r_{i,j}/\sigma_i\right), \quad r_{i,j} \geq 0 \quad (13)$$

where $Q_1(.)$ is Marcum Q-function of first order. It is obtained as:

$$p_{\dot{r}_1}(\dot{r}_1) = P_1 \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} + P_2 \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} +$$
$$+ \left(P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] - P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right]\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} +$$
$$+ \left(P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] - P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right]\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} \quad (14)$$

$$p_{\dot{r}_2}(\dot{r}_2) = P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right]\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} +$$
$$+ P_2\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right]\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} +$$
$$+ P_1 B_1(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} + P_2 B_2(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} +$$
$$+ P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} +$$
$$+ P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} +$$
$$+ P_1 C_1(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} + P_2 C_2(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} \quad (15)$$

The curves for probability density functions of signal derivatives at the SSC combiner output at two time instants, versus signal derivatives $\dot{r}$, are presented in Fig. 2 for different values of parameter $\dot{\sigma}_i$.

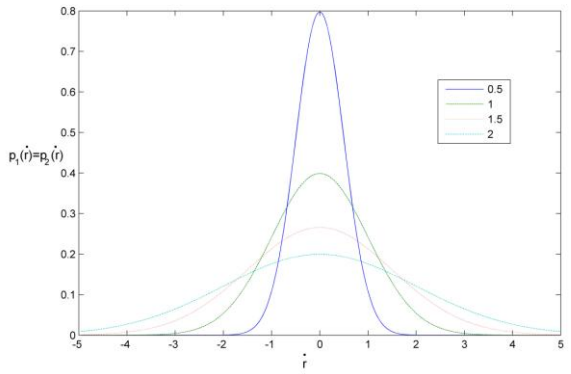


Figure 2. The probability density functions of derivatives at the SSC combiner output at two time instants with parameter $\dot{\sigma}_i$ = 0.5; 1; 1.5; 2

The parameters of PDFs curves are different values of parameter $\dot{\sigma}_i$ (0.5; 1; 1.5; 2). The case of channels with identical distribution is observed.

V. System Model and Statistics of SSC/SC Combiner

The model of the SSC/SC combiner with two inputs at two time instants, considering in this paper, is shown in Fig. 3. The SSC combiner input signals, r_{11} and r_{21} at first time instant, and r_{12} and r_{22} at the second time instant, are overall system inputs. The output signals from SSC part of combiner are r_1 and r_2 and they are inputs for second part of combiner. The overall output signal from complex system is r.

The joint probability density functions at the SSC combiner outputs at two time instants in Rician fading channels for four different cases are:

$r_1<r_T$, $r_2<r_T$

$$p_{r_1r_2}(r_1,r_2)=P_1D_2(r_1)D_1(r_2)+P_2D_1(r_1)D_2(r_2) \qquad (16)$$

$r_1\geq r_T$, $r_2<r_T$

$$p_{r_1r_2}(r_1,r_2)=P_1\frac{r_2}{\sigma_2^{\,2}}e^{-\frac{r_2^{\,2}+A_2^{\,2}}{2\sigma_2^{\,2}}}I_0\left(\frac{r_2A_2}{\sigma_2^{\,2}}\right)D_1(r_1)+P_1D_2(r_1)D_1(r_2)+$$

$$+P_2\frac{r_2}{\sigma_1^{\,2}}e^{-\frac{r_2^{\,2}+A_1^{\,2}}{2\sigma_1^{\,2}}}I_0\left(\frac{r_2A_1}{\sigma_1^{\,2}}\right)D_2(r_1)+P_2D_1(r_1)D_2(r_2) \qquad (17)$$

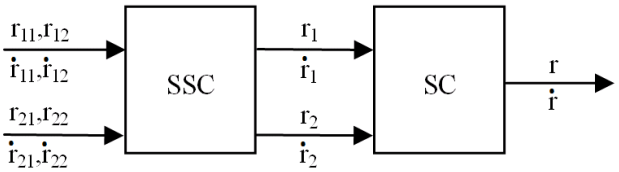


Figure 3. Model of the SSC/SC combiner with two inputs at two time instants

$r_1<r_T$, $r_2\geq r_T$

$$p_{r_1r_2}(r_1,r_2)=P_1\left(1-Q_1\left(A/\sigma_1,r_t/\sigma_1\right)\right)\cdot$$

$$\cdot\frac{r_1r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_1D_2(r_1)D_1(r_2)+$$

$$*+P_2\left(1-Q_1\left(A/\sigma_2,r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_2D_1(r_1)D_2(r_2) \qquad (18)$$

$r_1\geq r_T$, $r_2\geq r_T$

$$p_{r_1r_2}(r_1,r_2)=P_1\frac{r_1r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_1\left(1-Q_1\left(A/\sigma_1,r_t/\sigma_1\right)\right)\cdot$$

$$\cdot\frac{r_1r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_1\frac{r_2}{\sigma_2^{\,2}}e^{-\frac{r_2^{\,2}+A_2^{\,2}}{2\sigma_2^{\,2}}}I_0\left(\frac{r_2A_2}{\sigma_2^{\,2}}\right)D_1(r_1)+P_1D_2(r_1)D_1(r_2)+$$

$$+P_2\frac{r_1r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_2\left(1-Q_1\left(A/\sigma_2,r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+ P_2 \frac{r_2}{\sigma_1^2} e^{-\frac{r_2^2 + A_1^2}{2\sigma_1^2}} I_0\left(\frac{r_2 A_1}{\sigma_1^2}\right) D_2(r_1) + P_2 D_1(r_1) D_2(r_2) \quad (19)$$

where:

$$D_i(r) = \frac{1}{\sigma_i^2} e^{-\frac{A_2^2}{\sigma_i^2(1+\rho)}} \cdot$$

$$\cdot \sum_{k,l_1,l_2,l_3=0}^{\infty} \varepsilon_i \frac{1}{l_1! l_2! l_3! (k+l_1)! (k+l_2)! (k+l_3)!} \frac{\rho^{k+2l_1}}{\left(2\sigma_i^2(1-\rho^2)\right)^{2k+l_1+l_2+2l_3}} \cdot$$

$$\cdot A_i^{2k+2l_2+2l_3} (1-\rho)^{2k+2l_2+2l_3} \cdot r^{2k+2l_1+2l_3+1} e^{-\frac{r^2}{2\sigma_i^2(1-\rho^2)}} \cdot$$

$$\cdot \gamma\left(k+l_1+l_2+1, \frac{r_T^2}{2\sigma_i^2(1-\rho^2)}\right) \quad (20)$$

For dual SC receiver with correlated diversity branches the joint PDF $p_{r\dot{r}}(r,\dot{r})$ is given by [19, eq. (8.42)]

$$p_{r\dot{r}}(r,\dot{r}) = p_{\dot{r}_1}(\dot{r}) \int_0^r p_{r_1 r_2}(r, r_2) dr_2 + p_{\dot{r}_2}(\dot{r}) \int_0^r p_{r_1 r_2}(r_1, r) dr_1 \quad (21)$$

For SSC/SC combiner, the joint PDF $p_{r\dot{r}}(r,\dot{r})$ is obtained for different values of r:

$r<r_T$ $(r_1<r_T, r_2<r_T)$

$$p_{r\dot{r}}(r,\dot{r}) = p_{\dot{r}_1}(\dot{r}) \int_0^r p_{r_1 r_2}(r, r_2) dr_2 + p_{\dot{r}_2}(\dot{r}) \int_0^r p_{r_1 r_2}(r_1, r) dr_1 \quad (22)$$

$r\geq r_T$ $(r_1\geq r_T, r_2<r_T)$

$$p_{r\dot{r}}{}^1(r,\dot{r}) = p_{\dot{r}_1}(\dot{r}) \int_0^{r_T} p_{r_1 r_2}(r, r_2) dr_2 + p_{\dot{r}_2}(\dot{r}) \int_{r_T}^r p_{r_1 r_2}(r_1, r) dr_1 \quad (23)$$

$r\geq r_T$ $(r_1<r_T, r_2\geq r_T)$

$$p_{r\dot{r}}{}^2(r,\dot{r}) = p_{\dot{r}_1}(\dot{r}) \int_{r_T}^r p_{r_1 r_2}(r, r_2) dr_2 + p_{\dot{r}_2}(\dot{r}) \int_0^{r_T} p_{r_1 r_2}(r_1, r) dr_1 \quad (24)$$

$r\geq r_T$ $(r_1\geq r_T, r_2\geq r_T)$

$$p_{r\dot{r}}{}^3(r,\dot{r}) = p_{\dot{r}_1}(\dot{r}) \int_{r_T}^r p_{r_1 r_2}(r, r_2) dr_2 + p_{\dot{r}_2}(\dot{r}) \int_{r_T}^r p_{r_1 r_2}(r_1, r) dr_1 \quad (25)$$

$r\geq r_T$ $(r_1\geq r_T, r_2<r_T;\ r_1<r_T, r_2\geq r_T;\ r_1\geq r_T, r_2\geq r_T)$

$$p_{r\dot{r}}(r,\dot{r}) = p_{r\dot{r}}{}^1(r,\dot{r}) + p_{r\dot{r}}{}^2(r,\dot{r}) + p_{r\dot{r}}{}^3(r,\dot{r}) \quad (26)$$

Using (14) and (15) for PDFs of derivatives, and (16) – (19) for PDFs of signals and putting them into (22) – (25), the joint PDF $p_{r\dot{r}}(r,\dot{r})$ is obtained as:

For $r<r_T$:

$$p_{r\dot{r}}(r,\dot{r}) = \left[P_1 \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} + P_2 \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} + \right.$$

$$+ \left(P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] - P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] \right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} +$$

$$\left. + \left(P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] - P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} \right] \cdot$$

$$\cdot \int_0^r \left[P_1 D_2(r) D_1(r_2) + P_2 D_1(r) D_2(r_2) \right] dr_2 +$$

$$+ \left[P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right]\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} + \right.$$

$$\left. + P_2\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right]\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} + \right.$$

$$+ P_1 B_1(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} + P_2 B_2(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} +$$

$$+ P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} +$$

$$+ P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} +$$

$$\left. + P_1 C_1(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_1^2}} + P_2 C_2(r_T) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_2^2}{2\dot{\sigma}_2^2}} \right] \cdot$$

$$\cdot \int_0^r \left[P_1 D_2(r_1) D_1(r) + P_2 D_1(r_1) D_2(r) \right] dr_1 \quad (27)$$

For $r\geq r_T$

$$p_{r\dot{r}}(r,\dot{r}) = \left[P_1 \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} + P_2 \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} + \right.$$

$$+ \left(P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] - P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] \right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_1} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_1^2}} +$$

$$\left. + \left(P_1\left[1 - Q_1\left(\frac{A_1}{\sigma_1}, \frac{r_T}{\sigma_1}\right)\right] - P_2\left[1 - Q_1\left(\frac{A_2}{\sigma_2}, \frac{r_T}{\sigma_2}\right)\right] \right) \frac{1}{\sqrt{2\pi}\dot{\sigma}_2} e^{-\frac{\dot{r}_1^2}{2\dot{\sigma}_2^2}} \right] \cdot$$

$$\cdot\left[\int_0^{r_T}\left[P_1\frac{r_2}{{\sigma_2}^2}e^{-\frac{{r_2}^2+{A_2}^2}{2{\sigma_2}^2}}I_0\left(\frac{r_2A_2}{{\sigma_2}^2}\right)D_1(r)+P_1D_2(r)D_1(r_2)+\right.\right.$$

$$\left.\left.+P_2\frac{r_2}{{\sigma_1}^2}e^{-\frac{{r_2}^2+{A_1}^2}{2{\sigma_1}^2}}I_0\left(\frac{r_2A_1}{{\sigma_1}^2}\right)D_2(r)+P_2D_1(r)D_2(r_2)\right]dr_2+\right.$$

$$+\int_{r_T}^{r}\left[P_1\left(1-Q_1\left(A/\sigma_1,r_t/\sigma_1\right)\right)\cdot\right.$$

$$\left.\cdot\frac{rr_2}{{\sigma_2}^4(1-\rho^2)}e^{-\frac{r^2+{r_2}^2+2{A_2}^2(1-\rho)}{2{\sigma_2}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho rr_2}{{\sigma_2}^2(1-\rho^2)}\right)I_i\left(\frac{A_2r}{{\sigma_2}^2(1+\rho)}\right)I_i\left(\frac{A_2r_2}{{\sigma_2}^2(1+\rho)}\right)+\right.$$

$$\left.+P_1D_2(r_1)D_1(r_2)+\right.$$

$$\left.+P_2\left(1-Q_1\left(A/\sigma_2,r_t/\sigma_2\right)\right)\cdot\right.$$

$$\left.\cdot\frac{r_1r_2}{{\sigma_1}^4(1-\rho^2)}e^{-\frac{{r_1}^2+{r_2}^2+2{A_1}^2(1-\rho)}{2{\sigma_1}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{{\sigma_1}^2(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{{\sigma_1}^2(1+\rho)}\right)I_i\left(\frac{A_1r_2}{{\sigma_1}^2(1+\rho)}\right)+\right.$$

$$\left.+P_2D_1(r)D_2(r_2)\right]dr_2+$$

$$+\int_{r_T}^{r}\left[P_1\frac{r_1r_2}{{\sigma_1}^4(1-\rho^2)}e^{-\frac{{r_1}^2+{r_2}^2+2{A_1}^2(1-\rho)}{2{\sigma_1}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{{\sigma_1}^2(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{{\sigma_1}^2(1+\rho)}\right)I_i\left(\frac{A_1r_2}{{\sigma_1}^2(1+\rho)}\right)+\right.$$

$$\left.+P_1\left(1-Q_1\left(A/\sigma_1,r_t/\sigma_1\right)\right)\cdot\right.$$

$$\left.\cdot\frac{r_1r_2}{{\sigma_2}^4(1-\rho^2)}e^{-\frac{{r_1}^2+{r_2}^2+2{A_2}^2(1-\rho)}{2{\sigma_2}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{{\sigma_2}^2(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{{\sigma_2}^2(1+\rho)}\right)I_i\left(\frac{A_2r_2}{{\sigma_2}^2(1+\rho)}\right)+\right.$$

$$\left.+P_1\frac{r_2}{{\sigma_2}^2}e^{-\frac{{r_2}^2+{A_2}^2}{2{\sigma_2}^2}}I_0\left(\frac{r_2A_2}{{\sigma_2}^2}\right)D_1(r_1)+P_1D_2(r_1)D_1(r_2)+\right.$$

$$\left.+P_2\frac{r_1r_2}{{\sigma_2}^4(1-\rho^2)}e^{-\frac{{r_1}^2+{r_2}^2+2{A_2}^2(1-\rho)}{2{\sigma_2}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{{\sigma_2}^2(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{{\sigma_2}^2(1+\rho)}\right)I_i\left(\frac{A_2r_2}{{\sigma_2}^2(1+\rho)}\right)+\right.$$

$$\left.+P_2\left(1-Q_1\left(A/\sigma_2,r_t/\sigma_2\right)\right)\cdot\right.$$

$$\left.\cdot\frac{r_1r_2}{{\sigma_1}^4(1-\rho^2)}e^{-\frac{{r_1}^2+{r_2}^2+2{A_1}^2(1-\rho)}{2{\sigma_1}^2(1-\rho^2)}}\cdot\right.$$

$$\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r_2}{{\sigma_1}^2(1-\rho^2)}\right)I_i\left(\frac{A_1r_1}{{\sigma_1}^2(1+\rho)}\right)I_i\left(\frac{A_1r_2}{{\sigma_1}^2(1+\rho)}\right)+\right.$$

$$\left.+P_2\frac{r_2}{{\sigma_1}^2}e^{-\frac{{r_2}^2+{A_1}^2}{2{\sigma_1}^2}}I_0\left(\frac{r_2A_1}{{\sigma_1}^2}\right)D_2(r_1)+P_2D_1(r_1)D_2(r_2)\right]dr_2+\right.$$

$$\left.+\left[P_1\left[1-Q_1\left(\frac{A_1}{\sigma_1},\frac{r_T}{\sigma_1}\right)\right]\left[1-Q_1\left(\frac{A_2}{\sigma_2},\frac{r_T}{\sigma_2}\right)\right]\frac{1}{\sqrt{2\pi}\dot\sigma_1}e^{-\frac{\dot r_2^2}{2\dot\sigma_1^2}}+\right.\right.$$

$$\left.\left.+P_2\left[1-Q_1\left(\frac{A_1}{\sigma_1},\frac{r_T}{\sigma_1}\right)\right]\left[1-Q_1\left(\frac{A_2}{\sigma_2},\frac{r_T}{\sigma_2}\right)\right]\frac{1}{\sqrt{2\pi}\dot\sigma_2}e^{-\frac{\dot r_2^2}{2\dot\sigma_2^2}}+\right.\right.$$

$$\left.\left.+P_1B_1(r_T)\frac{1}{\sqrt{2\pi}\dot\sigma_2}e^{-\frac{\dot r_2^2}{2\dot\sigma_2^2}}+P_2B_2(r_T)\frac{1}{\sqrt{2\pi}\dot\sigma_1}e^{-\frac{\dot r_2^2}{2\dot\sigma_1^2}}+\right.\right.$$

$$\left.\left.+P_1\left[1-Q_1\left(\frac{A_1}{\sigma_1},\frac{r_T}{\sigma_1}\right)\right]Q_1\left(\frac{A_2}{\sigma_2},\frac{r_T}{\sigma_2}\right)\frac{1}{\sqrt{2\pi}\dot\sigma_2}e^{-\frac{\dot r_2^2}{2\dot\sigma_2^2}}+\right.\right.$$

$$\left.\left.+P_2\left[1-Q_1\left(\frac{A_2}{\sigma_2},\frac{r_T}{\sigma_2}\right)\right]Q_1\left(\frac{A_1}{\sigma_1},\frac{r_T}{\sigma_1}\right)\frac{1}{\sqrt{2\pi}\dot\sigma_1}e^{-\frac{\dot r_2^2}{2\dot\sigma_1^2}}+\right.\right.$$

$$\left.\left.+P_1C_1(r_T)\frac{1}{\sqrt{2\pi}\dot\sigma_1}e^{-\frac{\dot r_2^2}{2\dot\sigma_1^2}}+P_2C_2(r_T)\frac{1}{\sqrt{2\pi}\dot\sigma_2}e^{-\frac{\dot r_2^2}{2\dot\sigma_2^2}}\right]\cdot\right.$$

$$\left.\cdot\left[\int_{r_T}^{r}\left[P_1\frac{r}{{\sigma_2}^2}e^{-\frac{r^2+{A_2}^2}{2{\sigma_2}^2}}I_0\left(\frac{rA_2}{{\sigma_2}^2}\right)D_1(r_1)+P_1D_2(r_1)D_1(r)+\right.\right.\right.$$

$$\left.\left.\left.+P_2\frac{r}{{\sigma_1}^2}e^{-\frac{{r_2}^2+{A_1}^2}{2{\sigma_1}^2}}I_0\left(\frac{rA_1}{{\sigma_1}^2}\right)D_2(r_1)+P_2D_1(r_1)D_2(r)\right]dr_1+\right.\right.$$

$$\left.\left.+\int_0^{r_T}\left[P_1\left(1-Q_1\left(A/\sigma_1,r_t/\sigma_1\right)\right)\cdot\right.\right.\right.$$

$$\left.\left.\left.\cdot\frac{r_1r}{{\sigma_2}^4(1-\rho^2)}e^{-\frac{{r_1}^2+r^2+2{A_2}^2(1-\rho)}{2{\sigma_2}^2(1-\rho^2)}}\cdot\right.\right.\right.$$

$$\left.\left.\left.\cdot\sum_{i=0}^{\infty}\varepsilon_iI_i\left(\frac{\rho r_1r}{{\sigma_2}^2(1-\rho^2)}\right)I_i\left(\frac{A_2r_1}{{\sigma_2}^2(1+\rho)}\right)I_i\left(\frac{A_2r}{{\sigma_2}^2(1+\rho)}\right)+\right.\right.\right.$$

$$\left.\left.\left.+P_1D_2(r_1)D_1(r)+\right.\right.\right.$$

$$+P_2\left(1-Q_1\left(A/\sigma_2, r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1 r}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r^2+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_2 D_1(r_1)D_2(r)\Big]dr_1+$$

$$+\int_{r_T}^{r}\Bigg[P_1\frac{r_1 r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_1\left(1-Q_1\left(A/\sigma_1, r_t/\sigma_1\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2 r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2 r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_1\frac{r_2}{\sigma_2^{\,2}}e^{-\frac{r_2^{\,2}+A_2^{\,2}}{2\sigma_2^{\,2}}}I_0\left(\frac{r_2 A_2}{\sigma_2^{\,2}}\right)D_1(r_1)+P_1 D_2(r_1)D_1(r_2)+$$

$$+P_2\frac{r_1 r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2 r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2 r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_2\left(1-Q_1\left(A/\sigma_2, r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_2\frac{r_2}{\sigma_1^{\,2}}e^{-\frac{r_2^{\,2}+A_1^{\,2}}{2\sigma_1^{\,2}}}I_0\left(\frac{r_2 A_1}{\sigma_1^{\,2}}\right)D_2(r_1)+P_2 D_1(r_1)D_2(r_2)\Bigg]dr_1\Bigg] \quad (28)$$

The average LCR at envelope level, r, is defined as the rate at which a fading signal envelope crosses level r in a positive or negative going direction. Denoting the signal envelope and its time derivative by r and $\dot{r}$, respectively, the average LCR is given by [16],[17]

$$N_r(r)=\int_0^{\infty}\dot{r}p_{r\dot{r}}(r,\dot{r})d\dot{r} \quad (29)$$

The average fade duration is defined as the average time that the fading envelope remains below the specified signal level, after crossing the level in a downward direction and is given by [20]

$$T_r(r)=\frac{P_{out}(r)}{N_r(r)} \quad (30)$$

where $P_{out}(r)$ is outage probability at the output of SSC/SC combiner. It can be obtained by using (16) - (19) similarly like for obtaining LCR:

$r<r_T$

$$P_{out}(r)=\int_0^{r}\int_0^{r}P_1 D_2(r_1)D_1(r_2)+P_2 D_1(r_1)D_2(r_2)dr_1 dr_2 \quad (31)$$

$r\geq r_T$

$$P_{out}(r)=\int_0^{r_T}\int_0^{r_T}P_1 D_2(r_1)D_1(r_2)+P_2 D_1(r_1)D_2(r_2)dr_1 dr_2+$$

$$+\int_{r_T}^{r}\int_0^{r_T}\Bigg[P_1\frac{r_2}{\sigma_2^{\,2}}e^{-\frac{r_2^{\,2}+A_2^{\,2}}{2\sigma_2^{\,2}}}I_0\left(\frac{r_2 A_2}{\sigma_2^{\,2}}\right)D_1(r_1)+P_1 D_2(r_1)D_1(r_2)+$$

$$+P_2\frac{r_2}{\sigma_1^{\,2}}e^{-\frac{r_2^{\,2}+A_1^{\,2}}{2\sigma_1^{\,2}}}I_0\left(\frac{r_2 A_1}{\sigma_1^{\,2}}\right)D_2(r_1)+P_2 D_1(r_1)D_2(r_2)\Bigg]dr_1 dr_2+$$

$$+\int_0^{r_T}\int_{r_T}^{r}\Big[P_1\left(1-Q_1\left(A/\sigma_1, r_t/\sigma_1\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2 r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2 r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_1 D_2(r_1)D_1(r_2)+$$

$$+P_2\left(1-Q_1\left(A/\sigma_2, r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_2 D_1(r_1)D_2(r_2)\Big]dr_1 dr_2+$$

$$+\int_{r_T}^{r}\int_{r_T}^{r}\Bigg[P_1\frac{r_1 r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_1\left(1-Q_1\left(A/\sigma_1, r_t/\sigma_1\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2 r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2 r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_1\frac{r_2}{\sigma_2^{\,2}}e^{-\frac{r_2^{\,2}+A_2^{\,2}}{2\sigma_2^{\,2}}}I_0\left(\frac{r_2 A_2}{\sigma_2^{\,2}}\right)D_1(r_1)+P_1 D_2(r_1)D_1(r_2)+$$

$$+P_2\frac{r_1 r_2}{\sigma_2^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_2^{\,2}(1-\rho)}{2\sigma_2^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_2^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_2 r_1}{\sigma_2^{\,2}(1+\rho)}\right)I_i\left(\frac{A_2 r_2}{\sigma_2^{\,2}(1+\rho)}\right)+$$

$$+P_2\left(1-Q_1\left(A/\sigma_2, r_t/\sigma_2\right)\right)\cdot$$

$$\cdot\frac{r_1 r_2}{\sigma_1^{\,4}(1-\rho^2)}e^{-\frac{r_1^{\,2}+r_2^{\,2}+2A_1^{\,2}(1-\rho)}{2\sigma_1^{\,2}(1-\rho^2)}}\cdot$$

$$\cdot\sum_{i=0}^{\infty}\varepsilon_i I_i\left(\frac{\rho r_1 r_2}{\sigma_1^{\,2}(1-\rho^2)}\right)I_i\left(\frac{A_1 r_1}{\sigma_1^{\,2}(1+\rho)}\right)I_i\left(\frac{A_1 r_2}{\sigma_1^{\,2}(1+\rho)}\right)+$$

$$+P_2\frac{r_2}{\sigma_1^{\,2}}e^{-\frac{r_2^{\,2}+A_1^{\,2}}{2\sigma_1^{\,2}}}I_0\left(\frac{r_2 A_1}{\sigma_1^{\,2}}\right)D_2(r_1)+P_2 D_1(r_1)D_2(r_2)\Bigg]dr_1 dr_2 \quad (32)$$

The level crossing rate (LRC) and average fade duration (AFD) curves for complex SSC/SC combiner at two time instants in the presence of Rician fading, depending on different values of the signal and derivative distributions' parameters, are shown in Figs. 4 to 7.

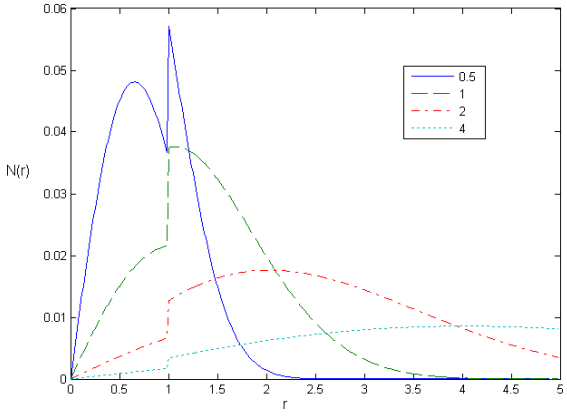


Figure 4. Level crossing rate $N(r)$ of SSC/SC combiner, at two time instants, versus signal amplitude for r_t=1, σ =0.5; 1; 2; 4, A =0.5 and $\dot{\sigma}$ =0.2

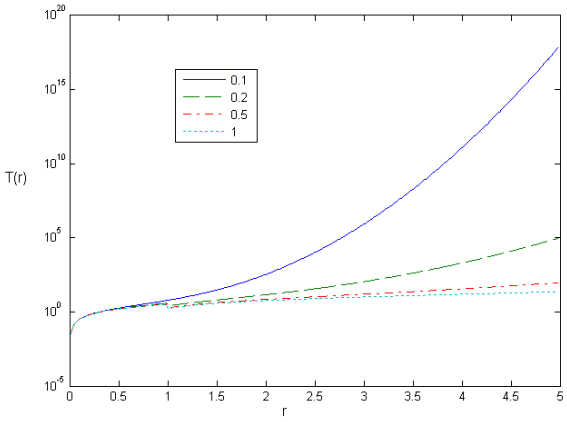


Figure 5. Average fade duration $T(r)$ of SSC/SC combiner, at two time instants, versus signal amplitude r, for r_t=1, σ =0.1; 0.2; 0.5; 1, A =0.5 and $\dot{\sigma}$ =0.2

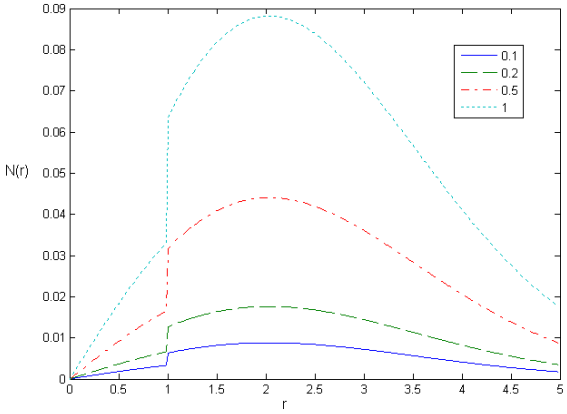


Figure 6. Level crossing rate $N(r)$ of SSC/SC combiner, at two time instants, versus signal amplitude r, for r_t=1, σ =2, A =0.5 and $\dot{\sigma}$ =0.1; 0.2; 0.5; 1

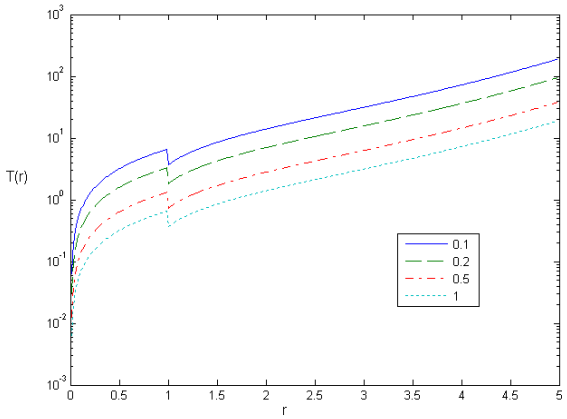


Figure 7. Average fade duration $T(r)$ of SSC/SC combiner, at two time instants, versus signal amplitude r, for r_t=1, σ =2, A =0.5 and $\dot{\sigma}$ =0.1; 0.2; 0.5; 1

Observing the first order system characteristics of SSC/SC combiner at two time instants presented in Fig. 2. [13, 14], it is obvious that the benefit of using complex SSC/SC combiner exists and it increases with decreasing of correlation between input signals.

Because of that, the expressions for second order system performance are derived in this paper and the influence of system parameters at second order system characteristics is shown in all these figures. These results are useful for designing and analyzing of wireless communication systems.

The level crossing rate of SSC/SC combiner at two time instants versus signal amplitude, $N(r)$, for r_t=1, A =0.5 and $\dot{\sigma}$=0 and different values of σ is presented in Fig. 4. It is visible from this figure that system performances are better for bigger values of parameter σ, because the system performances are better for lower values of average level crossing rate.

The level crossing rate of SSC/SC combiner at two time instants $N(r)$, for r_t=1, σ =2, A =0.5 and different values of $\dot{\sigma}$ is plotted in Fig. 6. One can see from this figure that system performances are better for lower values of parameter $\dot{\sigma}$.

The curves for average fade duration of SSC/SC combiner at two time instants, $T(r)$, are drawn in Figs. 5 and 7. The parameters' values are r_t=1, A =0.5 and $\dot{\sigma}$=0.2 in Fig. 5. Variable is parameter σ. It is evident that performances are better for greater values of parameter σ (lower values for AFD). The parameters' values are r_t=1, σ=2, A=0.5 in Fig. 7, with changeable $\dot{\sigma}$. One can conclude that results are more favorable for greater values of parameter $\dot{\sigma}$.

From all this figures it can be noticed that values of the LCR and AFD are growing, have discontinuity in the value of the threshold, where there is a drop in the values, and then are continuing to grow. In the case of LCR, the increasing of the value is evident till it reaches a maximum and then begins to fall. The system performance are better for lower values of average level crossing rate. For smaller LCR (i.e., the shallower fades), correspondingly AFD is larger. AFD increases with the value of the signal amplitude in whole range, but the curves also have drops at the threshold values.

VI. Conclusion

In this paper, the expressions for probability density functions (PDFs) of the time derivatives at two time instants for output signals from dual branch SSC combiner in the presence of Rician fading are given. The second order characteristics: the average level crossing rate and the average fade duration for complex combiner who makes the decision based on sampling at two time instants, are calculated by using derived closed-form expressions for the case of SSC/SC combiner. To point out the improvement of using complex SSC/SC combiner compared to classical SSC and SC combiners at one time instant, some graphs are presented earlier for several fading distributions.

Here, the second order system performances are calculated and presented graphically. These results are valuable in analyzing and projecting of wireless communication systems in the presence of Rician fading.

Acknowledgment

This paper has been funded by the Serbian Ministry of Education, Science and Technological Development of Republic of Serbia under projects III-44006 and TR-33035.

References

[1] D. Krstić, P. Nikolić, A. Stevanović and G. Stamenović, "The performance analysis of complex SSC/MRC combiner in Rice fading channel," The Ninth International Conference on Wireless and Mobile Communications, ICWMC 2013, July 21 - 26, 2013 - Nice, France, pp. 195-199, ISSN: 2308-4219, ISBN: 978-1-61208-284-4

[2] D. Krstić, P. Nikolić, and G. Stamenović, "Probability density functions of derivatives in two time instants for SSC combiner in Rician fading channel," The Eighth International Conference on Wireless and Mobile Communications, ICWMC 2012, June 24-29, 2012 - Venice, Italy, pp. 329-334, ISBN: 978-1-61208-203-5

[3] M. K. Simon and M. S. Alouni, Digital Communication over Fading Channels, Second Edition, Wiley-Interscience, A John Wiley&Sons, Inc., Publications, New Jersey, 2005.

[4] A. Abdi, C. Tepedelenlioglu, M. Kaveh, and G. Giannakis, "On the estimation of the K parameter for the Rice fading distribution," IEEE Communications Letters, pp. 92 -94, Mar. 2001.

[5] M. A. Richards, Rice Distribution for RCS, Georgia Institute of Technology, Sept. 2006.

[6] D.G. Brennan, "Linear diversity combining techniques," Proc. IRE, vol.47, no.1, pp.1075–1102, June 1959.

[7] D. Milovic, M. Stefanovic, and D. Pokrajac, "Stochastic approach for output SINR computation at SC diversity systems with correlated Nakagami-m fading," European Transactions on Telecommunications, vol. 20, no. 5, pp. 482-486, 2009.

[8] A. Cvetković, M. Stefanović, N. Sekulović, D. Milić, D. Stefanović, and Z. Popović, "Second-order statistics of dual SC macrodiversity system over channels affected by Nakagami-*m* fading and correlated gamma shadowing," Electrical Review (Przeglad Elektrotechniczny), vol. 87, no. 6, pp. 284-288, June 2011.

[9] P. Spalević, S. Panić, Ć. Dolićanin M. Stefanović, and A.Mosić, "SSC diversiity receiver over correlated α-μ fading channels in the presence of co-channel interference," EURASIP Journal on Wireless Communications and Networking, vol. 2010, doi:10.1155/2010/142392.

[10] Đ. V. Banđur, M. Stefanović, and M. V. Banđur, "Performance analysis of SSC diversity receiver over correlated Rician fading channels in the presence of co-channel interference," Electronics Letters, vol. 44, no. 9, pp. 587-588, 2008.

[11] G. T. Djordjevic, D. N. Milic, A. M. Cvetkovic, and M. C. Stefanovic, "Influence of imperfect cophasing on performance of EGC receiver of BPSK and QPSK signals transmitted over Weibull fading channel," European Transactions on Telecommunications, vol. 22, Issue 6, pp. 268–275, Oct. 2011, DOI: 10.1002/ett.1475.

[12] Z. Popovic, S. Panic, J. Anastasov, M. Stefanovic, and P.Spalevic, "Cooperative MRC diversity over Hoyt fading channels," Przeglad Elektrotechniczny (Electrical Review), vol. 87, no. 12, pp. 150-152, Dec. 2011.

[13] P. Nikolić, D. Krstić, M. Milić, and M. Stefanović, "Performance analysis of SSC/SC combiner at two time instants in the presence of Rayleigh fading," Frequenz, vol. 65, Issue 11-12, pp. 319–325, Nov. 2011, ISSN (Online)

2191-6349, ISSN (Print) 0016-1136, DOI 10.1515/FREQ.2011.048

[14] M. Stefanović, P. Nikolić, D. Krstić, and V. Doljak, "Outage probability of the SSC/SC combiner at two time instants in the presence of lognormal fading," Przeglad Elektrotechniczny (Electrical Review), R. 88 NR 3a/2012, pp. 237-240, March 2012, ISSN 0033-2097.

[15] M. K. Simon, "Comments on infinite-series representations associated with the bivariate Rician distribution and their applications," IEEE Trans. Commun., vol. 54, no 8, pp. 2149 – 2153, 25-28 Sept. 2005.

[16] T. S. Rappaport, Wireless Communications: Principles and Practice. Upper Saddle River, NJ: PTR Prentice-Hall, 1996.

[17] L. Yang and M.-S. Alouini, "Average level crossing rate and average outage duration of generalized selection combining," IEEE Transactions on Communications, vol. 51, no. 12, pp. 1063-1067, Dec. 2003.

[18] S. O. Rice, "Statistical properties of a sine wave plus random noise," Bell Syst. Tech. J., vol. 27, pp. 109–157, Jan. 1948.

[19] L. Yang and M. S. Alouini, "Average outage duration of wireless communication systems," ch. 8, Wireless Communications Systems and Networks. Springer, 2004.

[20] W. C. Jakes, Microwave Mobile Communications. New York: Wiley, 1974.

A No-reference Voice Quality Estimation Method for Opus-based VoIP Services

Péter Orosz, Tamás Skopkó, Zoltán Nagy, and Tamás Lukovics
Faculty of Informatics
University of Debrecen
Debrecen, Hungary
e-mail: oroszp@unideb.hu

Abstract— **By the emergence of real time media applications, correlating users' satisfaction with measured service quality is a constant challenge. Accordingly, this area was under intensive research in the last decade. Finding the correlation among Quality of Experience (QoE) for voice, measured Quality of Service (QoS) parameters in the network, and objective voice performance metrics is a key task. Most of the voice metrics use the reference content to compare the quality of the received stream. In contrast, this paper introduces a mathematical low-complexity, no-reference method that performs real-time estimation of QoE for Opus-based voice services. To determine the estimator function, we performed combined (subjective and objective) assessments to build a reference data set of 3-tuples of MOS, jitter and loss values. Applying polynomial regression, we used the reference data set to search a low-degree two-variable polynomial to hash objective QoS metrics (jitter and loss) to the subjective MOS score of the service quality. In the final phase of our investigation, we were evaluated the performance of the polynomials with a set of four audio clips.**

Keywords- Opus codec; real-time voice; QoE; QoE estimation; QoS-QoE correlation.

I. INTRODUCTION

Real-time audio services are sensitive to the timing and transmission performance of the network infrastructure. Since voice communication is interactive, low latency (≤150 ms) is a crucial requirement for an acceptable level of user experience. While monitoring of these performance metrics is a common solution, especially in dedicated infrastructures (like mobile networks), none of these parameters alone shows direct correlation with the perceptual quality (QoE). Therefore, service providers also apply subjective evaluation methods occasionally. Mean Opinion Score (MOS) and similar simplified scale, such as Absolute Category Rating (ACR) are generally used. Since subjective evaluation is time consuming and circumstantial, service providers mainly use these tools for a short but intensive period. A further drawback of an evaluation based on user feedback is the static evaluation of a whole user session (i.e., at the end of the conversation). Though the method can be extended by localizing the time moments of errors that affect QoE dramatically (e.g., by pressing a specific key straight after a disturbing glitch or noise), still the most efficient way for a service provider to assess perceptual quality of a service would be a dynamic, real-time analysis method, which estimates user experience. To accomplish this aim, the relationship between objective metrics (QoS) and perceptual quality should be investigated. Since the measurements should be run on network nodes aggregating intensive traffic (e.g., routers and switches), the method has to enable a hardware-accelerated implementation. Therefore, the complexity of the algorithm should be kept to a relatively low level.

We already investigated the performance of the evolved Opus audio codec in real network conditions [1]. In the related research we subjectively evaluated the voice content using the original, source audio content as reference. During the transmission we used pre-determined QoS parameters in a full-controlled laboratory network. We pointed out that the investigated Opus codec provides very good voice reproduction with a wide range of network parameters. We also found that there is a close-to-linear relation between MOS and packet loss rate. The experiments motivated us to unfold deeper relations between QoS and QoE. We set up a subsequent investigation consisting of three phases. The first phase is built upon the subjective evaluation technique discussed in our previous paper. In the second phase we continue to focus on VoIP services provided on managed networks instead of Over-the-Top (OTT) and we use statistical analysis methods, i.e., polynomial regression and correlation analysis on the reference data set of the first phase to describe the relation between the QoS parameters and QoE. In the third phase, we introduce the optimal coefficients for the polynomials that can be used to estimate the QoE. We are also evaluating the generality of the method by performing subjective as well as objective assessments with four independent voice clips.

In Section II, we summarize researches and developments relevant to the quality of the Opus audio codec. Section III shows some important details of voice transmission that have to be considered for the quality assessment of a voice codec. In Section IV, we present our evaluation method and its associated measurement configuration among some important details about VoIP transmission. In Section V, the investigation of the suitable statistical description is introduced as well as its mathematical background. We also validate our evaluation method by further subjective assessments detailed in Section IV. In the closing Section VII, we summarize our

observations, the introduced evaluation method and sketch further research aspects.

II. Related Works

A. Raake summarized the elements of VoIP speech quality evaluation and assessments [2].

A. Ramö and H. Toukomaa evaluated Opus' MDCT and LP modes by subjective listening tests and compared them with 3GPP AMR, AMR-WB and ITU-T G.718B, and they stated Opus to be a good alternative for the aforementioned codecs [3]. The paper of C. Hoene et al. includes different listening tests and compares the codec to Speex (both NB and WB), iLBC, G.722.1, G.722.1C, AMR-NB and AMR-WB [4]. They conclude that Opus performs better, though at lower rates, however AMRNB and AMR-WB still outperform the new Opus codec. Valin *et al.* present further improvements in the Opus encoder that help to minimize the impact of coding artifacts [5].

The International Telecommunication Union (ITU) has its own recommendation for standardized evaluation of speech quality: after many years of development (superseding PEAQ, PSQM, PESQ and PESQ-WB algorithms), POLQA (ITU-T P.863) is able to evaluate speech sampled up to 14 kHz using the MOS metrics [6].

Neves *et al* proposed a No Reference model for monitoring VoIP QoE based on the E-Model [7]. The method was proven to be a class C2 conformance in terms of subjective MOS scoring.

S. Cardeal *et al* introduced ArQoS, a probing system that integrates network performance monitoring methods as well as QoE assessment methods in a telecommunication infrastructure [8].

W. Cherif *et al* presented a non-intrusive QoE prediction method called A_PSQA based on a Random Neural Network (RNN) approach [9].

L. Fei *et al* discuss packet delay and bandwidth as main factors affecting QoE and introduce a carrier scheduling scheme for LTE based on their research [10]. They have significant QoE improvements in their simulation results.

Jelassi S. *et al* made an extensive survey on objective and subjective QoE assessment methods [11].

V. Aggarwal *et al* employs machine learning technique to passive QoS measurement information to predict VoIP QoE with at least 80% accuracy [12].

In our previous paper, we summarized objective voice quality evaluation methods, as well as subjective approaches. We examined the correlation of QoS metrics packet loss and jitter with subjective evaluations for VoIP audio transmission [1]. We used pre-defined QoS: loss and jitter values were iterated through a selected range. We streamed real human voice in emulated WAN environment based on the specific QoS parameters. The results were evaluated by a number of volunteers. We ran the experiment with both Opus and its predecessor, the Speex codec. Opus performed more uniformly on a wide range of QoS parameters than the Speex codec. Opus also showed a close-to-linear correlation with packet loss rate. These results opened the way for the construction of an estimator function.

III. Background

Since interactive real-time audio streaming is very sensitive to timing parameters, it is very common to use a specific protocol for media transmission. UDP-based Real-Time Transport Protocol provides the necessary parameters for time-sensitive data exchange [13]. Its most important metadata are the sequence number and the timestamp. The former provides a way of determining packet losses, reorders and duplications. The protocol assigns a per-application play-out buffer, where the packets are sorted using the timestamps and sequence numbers and accordingly duplicated packets can also be filtered out. To handle timing, the time-related information is also used. Since the goal is to ensure a continuously decoded media stream, this layer avoids both buffer over and underruns. A moving time window will specify what packets are received on time. Too early and too late packets will be dropped.

For optimal operation, it is necessary to have a constantly available guaranteed bandwidth. This can be assured most easily by using a constant bitrate (CBR), when packets of roughly the same size are transmitted with fixed rate. Most real-time media services still use CBR operation mode to effectively manage allocation of resources, and also, scheduling mechanisms can better handle a constant packet rate.

During an RTP/UDP real-time audio transmission using the Opus audio codec, the audio frames are all the same type, in contrast to the video frames of the H.264 codec. There are no key-frames that store data for a full video frame as a reference for several subsequent B-type and P-type frames. An audio frame stores audio samples for a fixed period of time (10-40 ms, typically). Thus, the effect of a lost packet always produces a gap in the voice stream. In contrast, jitter itself does not necessarily lead to data loss with appropriate buffering at the receiver side. We are not focusing on packet reordering and duplication since RTP handles these anomalies transparently. As far as sequence numbering and playout buffering allows these metrics do not cause degradation of quality for the end user.

Research of quality evaluation methods is not a novel field. However, creating objective methods that correlate well with subjective assessments is still a challenge. Most objective methods are full reference (FR), as the original media content is required for them to work. Since the source material is rarely available in the real world, there is a demand for solutions operating without the original content or only with some trace of it. They are no-reference (NR) or reduced reference (RR) methods. If we want to develop such a method for real-time QoE prediction, low calculation overhead is a further requirement, since embedded systems and mobile devices have limited computing power and resources.

IV. First phase of the investigation: reference assessments

As a basis of our research, we planned an environment providing laboratory conditions for the evaluation that can be repeated many times and reproduced by other groups. Since

network anomalies such as packet loss and jitter occur in real networks accidentally, we decided to use a network emulation tool that can introduce these types of error in a statistical way.

A. Source media content

We selected one high quality, pre-recorded voice clip containing an easy to understand male voice in the native language of the volunteers who participated in the evaluation. The clip was taken from an audio book, stored in standard CDDA resolution (44.1 kHz, 16 bit, stereo) and was resampled to 48 kHz, 16 bit, mono at the source of the stream. Its length was cut to 60 seconds. This length allows the listener to pick up the pace, and also enables a statistics-based network emulator to reach the steady state.

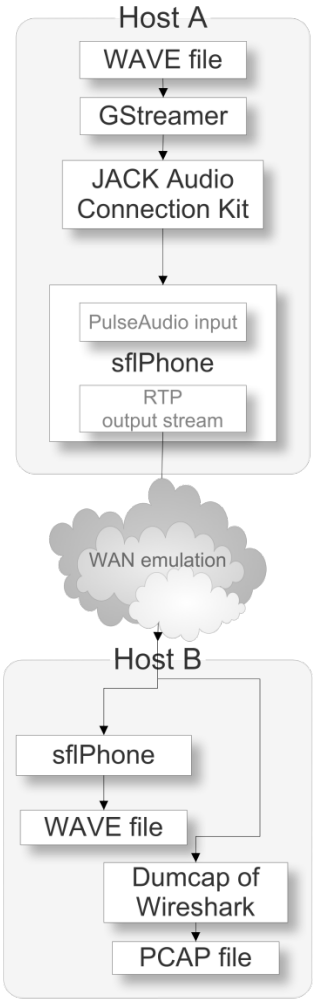


Figure 1. The measurement setup: audio is fed into the VoIP client on Host A and is transported through RTP to the other client on Host B

B. Measurement setup

An emulated WAN connection including two communication endpoints was constructed for the assessments (see Fig. 1).

Endpoints feature generic multi-core x64-based architectures and they were equipped with Intel PRO/1000 NICs. Fedora Core 18 was installed to both hosts with unmodified Linux 3.8.1-x kernel (with a jiffy setting of 1000 Hz). We have chosen version 1.2.2 of the *sflPhone* VoIP application, since it natively supports the Opus codec and our initial traffic measurements confirmed the expected QoS performance (i.e., packet rate, uniform distribution of inter-arrival times and packet sizes) [14].

The source voice clip was injected into the input of the softphone on Host A. *JACK Audio Connection Kit* is a general audio tool and is able to connect audio inputs and outputs of different applications and audio devices [15]. The current version of sflPhone can accept *ALSA* and *PulseAudio* datastream at its input. *PulseAudio* was selected since it can be connected with JACK. Since it has a native output plugin (sink) for *JACK*, the audio clip was fed into *JACK* from an uncompressed PCM WAVE file with the *GStreamer* application [16]. We carefully configured the applications not to perform unnecessary audio sample rate conversion throughout the digital audio path. The sflPhone application on Host B was configured to save the audio data into an uncompressed PCM WAVE file for further QoE assessments. Noise reduction and echo cancellation features were turned off. During the measurement, we used the *Netem* Linux kernel module, which was configured symmetrically on both directly connected interfaces to emulate a WAN connection and produced various network anomalies that affect QoS (i.e., packet loss and variation of network delay) [17].

Since we don't have to distinguish between different types of media packets, like in the case of a video stream, where packets contain different type of video frames (and key-frames can be transmitted in multiple consecutive packets), Netem's Layer-2 emulation technique is suitable for our measurement goals. During the iterated measurements, we stored both the WAVE file from the receiving softphone and the network traffic trace containing the received RTP stream [18]. ITU recommends delay to be kept under 150 ms for an acceptable interactive service and we wanted to emulate a generic WAN connection therefore we have chosen 100 ms of delay in each direction. The codec was measured with a series of parameters (Table 1). Netem network parameters were iterated using the following scheme:

TABLE I. PRE-DEFINED QOS VALUES FOR NETWORK EMULATION

Netem parameters	Set of values
jitter	0-20 ms in 1 ms steps
packet loss	0-40% in 1% steps

The softphone client generated 100 RTP packets per second. Packet size varied between 40 to 159 bytes (see Fig. 2).The codec generated an audio frame in each 8 ms. Its operation mode was constrained VBR (CVBR), which forces the encoder to operate at an average nominal bitrate. In our case this was 64 kbps and the variation of the inter-arrival time was in the range of ±500 µs.

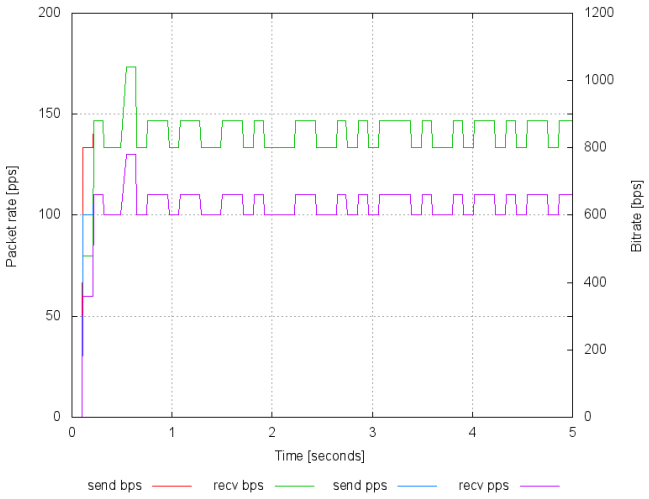


Figure 2. Packet rate and bandwidth during the voice transfer

C. *Evaluating the content*

Since we worked with a wide range of parameters, the iterated measurements resulted in more than 100 audio clips. Choosing finer resolution of these parameters would increase the number of clips even further, that could make the the evaluation significantly complicated by the monotony of the huge set of samples. As a reference for the evaluation, an initial measurement with zero jitter and no packet loss were run. The one-minute audio files got subjective QoE values by 15 volunteers who were in a calm and peaceful home environment. The evaluators were not VoIP professionals, had no deep knowledge of VoIP services and audio codecs.. Everyone did the tests at her own pace, with a comfortable timing, to minimize monotony. Each assessment session was started by listening to the reference voice clip. In the QoE analysis, the average of the MOS ratings was assigned to all audio files.

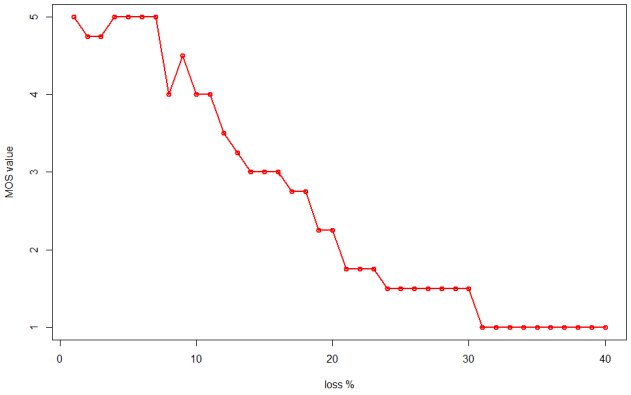


Figure 3. Correlation between packet loss ratio and subjective quality of experience expressed in MOS

We experienced near-linear relationship between quality degradation and the amount of lost packets (see Fig. 3).

Opus over RTP tolerates packet losses well even up to 30%. It can be a significant advantage when used with wireless access. Lower subjective quality right after the reference clip was a side-effect of the subjective assessment: evaluators got used to the good quality and even a small increment in the packet loss rate caused lower MOS ratings.

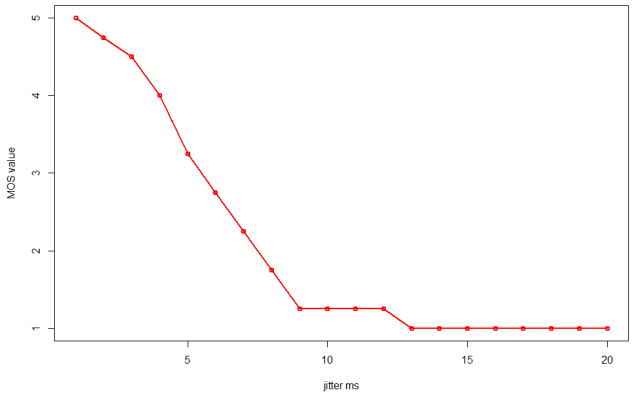


Figure 4. Correlation between jitter and subjective quality of experience expressed in MOS

The jitter-sensitivity of the transmission was also consistent but was further from the linear relation (see Fig. 4). The reason is that jitter can cause packet loss in the RTP layer since too early (generally caused by bursty packet forwarding) and too late packets were discarded by the transport protocol. Opus tolerates well a small amount of jitter, typically below 4 ms. However, the result of jitter elimination depends on the size of the receiver buffer. Greater jitter may cause significant degradation of perceived quality. In our investigation, this quality drop was experienced up to 9 ms. In contrast, in the range of 10 to 12 ms of jitter, an acceptable steady state was perceptible.

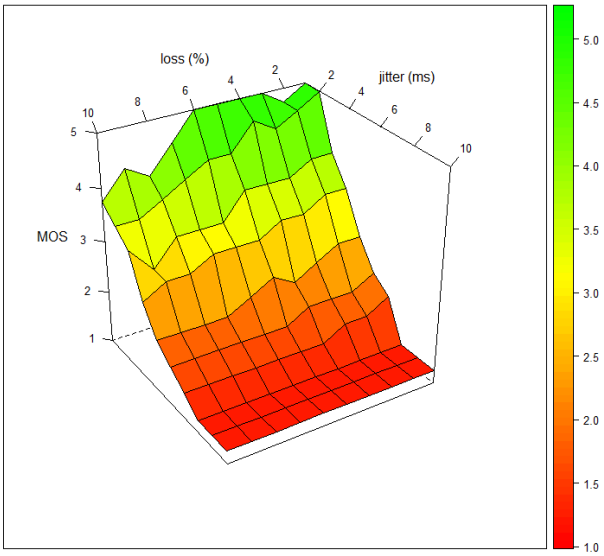


Figure 5. MOS values under mixed network conditions

Sensitivity to combined effect of the aforementioned network anomalies confirmed the experiences of the simple

scenarios. Opus was performing uniformly in term of packet loss. Contrarily, the effect of jitter made the relationship more complex as seen on Fig 5. However, the surface mapped to the MOS values of the combined measurements did not result in a very complex pattern and therefore it made us motivated to undertake a deeper investigation of the correlation between QoS and QoE.

V. SECOND PHASE OF THE INVESTIGATION: STATISTICAL ANALYSIS

A. Polynomial regression

The result of the first phase, i.e., assessed packet loss, jitter and MOS values, are presented in a 3-axis graph. We used polynomial regression to determine a two-variable polynomial of a matched surface. This is calculated using the class of functions (1).

$$F = \{p_n(x) = a_0 + a_1 x + \cdots + a_n x^n\} \quad (1)$$

A matched polynomial regression is the result class of polynomials of (2).

$$M(\eta - f^*(\xi))^2 = \min_{\forall f \in F} M(\eta - f(\xi))^2 \quad (2)$$

Coefficients are determined by solving the linear simultaneous equations in (3).

$$\begin{pmatrix} 1 & M\xi & \cdots & M\xi^n \\ M\xi & M\xi^2 & \cdots & M\xi^{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ M\xi^i & M\xi^{i+1} & \cdots & M\xi^{i+n} \\ \vdots & \vdots & \ddots & \vdots \\ M\xi^n & M\xi^{n+1} & \cdots & M\xi^{2n} \end{pmatrix} * \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} M\eta \\ M\eta\xi \\ \vdots \\ M\eta\xi^i \\ \vdots \\ M\eta\xi^n \end{pmatrix} \quad (3)$$

The polynomial regression was calculated using the *Matlab* function *poly*{ij}. The surface model can be tuned by the degree of input parameters. In a polynomial regression using second degree for the first variable, and first degree for the second one (poly21) can be described as (4).

$$Z = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy \quad (4)$$

where *x* denotes the amount of packet loss and *y* specifies the amount of jitter.

B. Surface inspection using SSE

Sum of Squares due to Error (SSE) determines the standard deviation between a set of points and the matched surface and is calculated using (5).

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (5)$$

The surface is less accurate match for the specified set of points if the SSE value falls further from zero.

C. Matched surfaces – poly11

Figure 6 plots the matched surface for this purely linear model, where the surface is determined by (6) and its coefficients are summarized in Table II.

$$f(x, y) = p_{00} + p_{10}x + p_{01}y \quad (6)$$

TABLE II. COEFFICIENTS FOR (6)

p_{00}	5.151
p_{10}	-0.07513
p_{01}	-0.4111

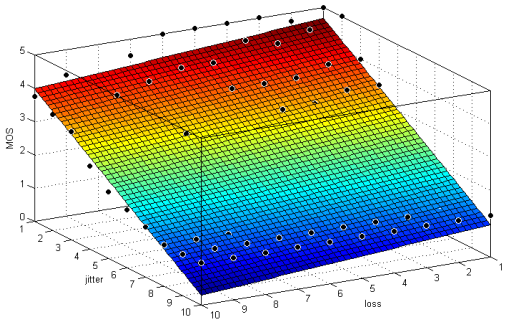


Figure 6. The surface determined by (6) and the MOS values from subjective evaluation in Phase 1

Goodness of fit with SSE is 16.24, and thus the MOS scores fall far from the surface displayed on Fig. 6.

D. Matched surfaces – poly12

If we use a second degree polynomial for the jitter (y) variable, the describing function will be (7). The coefficients are specified in Table III.

$$f(x, y) = p_{00} + p_{10}x + p_{01}y + p_{02}y^2 + p_{11}xy \quad (7)$$

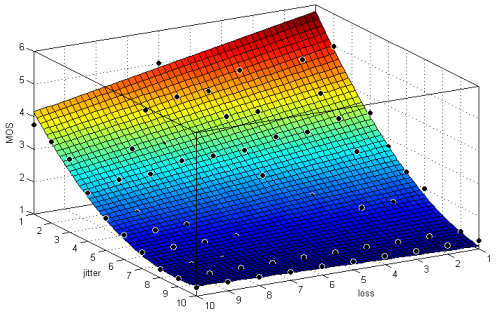


Figure 7. The surface determined by (7) and the MOS values from subjective evaluation in Phase 1

TABLE III. COEFFICIENTS FOR (7)

p_{00}	6.985
p_{10}	-0.2052
p_{01}	-1.063
p_{11}	0.02292
p_{02}	0.04696

Goodness of fit using SSE is 2.336. Introducing a second degree jitter variable resulted in a significant improvement as shown on Fig. 7.

E. *Matched surfaces – poly22*

In this case we are using a second degree function for the loss parameter also. The corresponding polynomial is (8).

$$f(x,y) = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2 \quad (8)$$

TABLE IV. COEFFICIENTS FOR (8)

p_{00}	7.067
p_{10}	-0.2382
p_{01}	-1.07
p_{20}	0.003474
p_{11}	0.02217
p_{02}	0.04787

The SSE goodness of fit is 2.289 and means only a slightly closer match was achieved by introducing a more complex calculation. However, increasing the degree of the loss variable, which shows close-to-linear relationship to the MOS values, did not imply a significantly lower SSE value.

F. *Matched surfaces – poly13*

We also inspected the effect of a higher degree polynomial for jitter. In this case the describing function is (9) and its coefficients are specified in Table V.

$$f(x,y) = p_{00} + p_{10}x + p_{01}y + p_{11}xy + p_{02}y^2 + p_{12}xy^2 + p_{13}y^3 \quad (9)$$

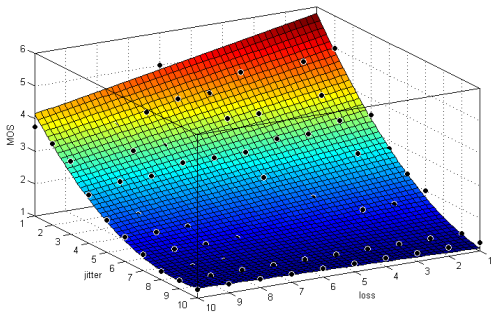


Figure 8. The surface determined by (9) and the MOS values from subjective evaluation in Phase 1

TABLE V. COEFFICIENTS FOR (9)

p_{00}	6.728
p_{10}	-0.241
p_{01}	-0.7999
p_{11}	0.04121
p_{02}	-0.01849
p_{12}	-0.001633
p_{03}	0.004354

For the surface demonstrated on Fig. 8, the SSE goodness of fit value is 1.69. It means a better matching surface but also adds a significant amount of complexity to the calculation.

G. *More matched surfaces*

We calculated a number of matched surfaces of higher degree functions. Fig. 9 shows their matching efficiency using SSE goodness of fit versus the degree of functions. The list of calculated polynomials, their number of coefficients and their error values are listed in Table VI and Table VII. Some questions are raised at this point raised: Do higher degree polynomials lead to better correlation results? What is the optimal complexity-error trade-off for a real-time NR method?

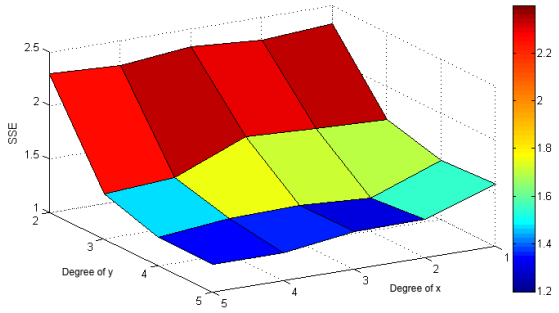


Figure 9. SSE goodness of fit for surface matching using polynomial regression and second and higher degree functions for packet loss and jitter parameters

TABLE VI. NUMBER OF COEFFICIENTS FOR DIFFERENT DEGREES OF LOSS AND JITTER VARIABLES

Degree of x	**Degree of y**				
	1	**2**	**3**	**4**	**5**
1	3	5	7	9	11
2	5	6	9	12	15
3	7	9	10	14	18
4	9	12	14	14	20
5	11	15	18	20	21

TABLE VII. SSE GOODNESS OF FIT VALUES FOR DIFFERENT DEGREES OF LOSS AND JITTER VARIABLES

Degree of x	Degree of y				
	1	2	3	4	5
1	16.2389	2.3363	1.6905	1.5559	1.5829
2	14.1052	2.2895	1.7138	1.3088	1.362
3	13.8641	2.3346	1.7445	1.353	1.3542
4	13.9947	2.2665	1.4665	1.3375	1.2664
5	13.7948	2.3008	1.4275	1.2668	1.2674

VI. THIRD PHASE OF THE INVESTIGATION: VALIDATION OF THE METHOD

The primary aim of this last phase is to verify the generality of the correlation calculation based on the coefficients and polynomials determined in Phase 2. The measurement environment described in Phase 1 remained unchanged.

The final goal of investigation is to construct a low degree evaluation function that inputs measured objective values and outputs the estimated value of QoE on the MOS scale.

A. Source media contents

We selected four voice clips from different sources. A female voice was present in two of the clips (clips A and B), and a male voice in the other two (clips C and D). Only native language was used. All clips were taken from audio books, being stored in standard CDDA resolution (44.1 kHz, 16 bit, stereo), they were resampled to 48 kHz, 16 bit, mono at the source of the stream. All the audio clips were 60 seconds long.

B. Evaluating the goodness of QoS-QoE mapping

The evaluations were executed in the same way as presented in Phase 1. In this phase, 10 volunteers were involved to the subjective assessment. Clip A containing a female voice got lower scores on the MOS scale than expected. This may be caused by the intonation of the person's voice. We compared the uncompressed source material and the 64 kbps Opus-encoded audio without network errors. While no disturbing artifacts were experienced in the source material, the Opus codec made sometimes the woman's voice raised in volume and this may be the root cause of the lower MOS scores.

We made 3D surfaces of MOS scores in the inspected QoS parameter range for all of the audio clips. The surfaces are presented on Fig. 10.

There is a quasi-linear correlation between MOS value ranging from 2 to 4 for Clip A. Over 4 ms of jitter, the experienced quality started to degrade. However, increasing packet loss did not degrade the scores dramatically.

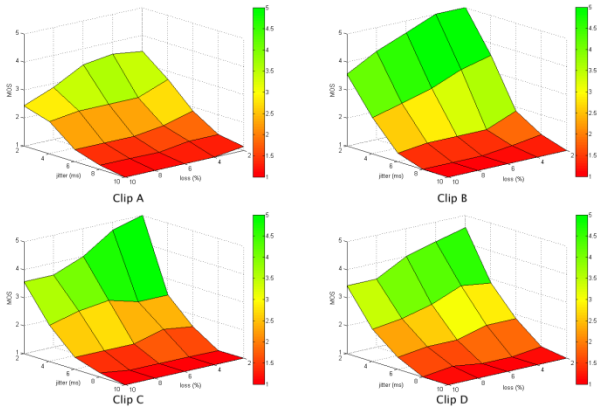


Figure 10. MOS values for different packet loss and jitter parameters

In Clip B also containing a female voice, the coding artifacts were not so significant. This resulted in higher MOS scores in general. The close-to-linear relation can be observed in the case of packet loss. Above the MOS value of 2, this correlation with the jitter is also present.

The correlation between MOS and jitter is still first-degree above MOS value 2 in Clip C with the male voice, while correlation with packet loss got closer to a second-degree function.

The surface of Clip D is similar to Clip C but got lower scores in general. This may also be caused by the source material.

Based on the surface analysis, we can conclude that there is a quasi-linear correlation between packet loss and MOS scores in the range of 2 to 4. This may enable good estimation of perceived quality. We assume that the most interesting range of the MOS scale is from 3 to 4. This range is where the customer will decide between continuing using the service and giving up. Around the MOS value of 4, there is a long-term acceptance-level of the service quality, while below 3, perceptive quality may become poor with annoying artifacts.

We also experienced the Opus codec being very sensitive to the voice intonation in some cases. All of the source materials were high quality, but the encoding process at this bitrate introduced some artifacts that affected the experienced quality very dramatically when packet losses or higher jitter occurs.

C. The results – Polynomial regression

Based on the conclusion of Phase 2, we assume that constructing an estimator function using first or second degree function is possible. It would be beneficial for a hardware implementation of such an estimator function to use low calculation complexity

The correlations of the objective and subjective MOS series, which is calculated from (1) are presented on Fig.11 to 14.

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{E[(X-\mu_x)(Y-\mu_Y)]}{\sigma_x \sigma_y} \qquad (10)$$

where μ is the expected value of the random variable and a σ is its standard deviation.

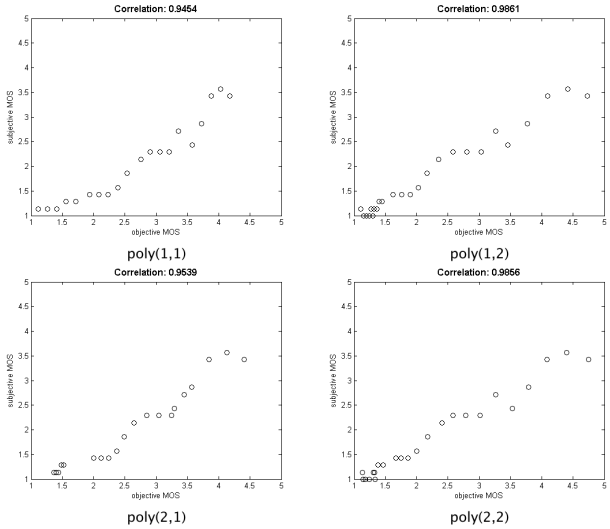


Figure 11. Correlation using different degree polynomials for Clip A

Poly(x,y) is the regression function, where packet loss is represented by *x*, and jitter is described by *y*.

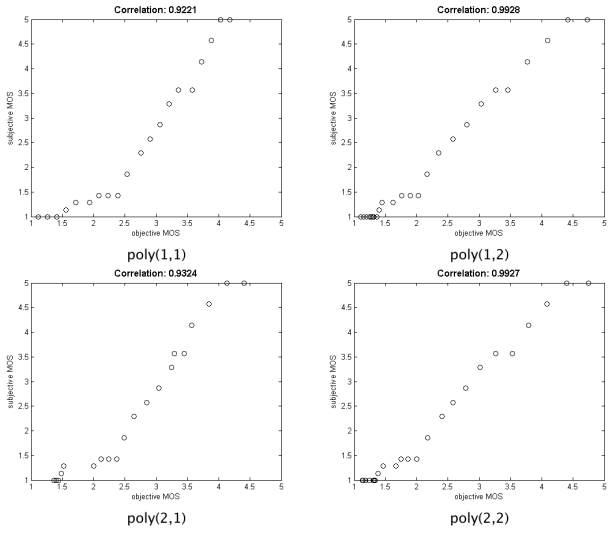


Figure 12. Correlation using different degree polynomials for Clip B

In the case of Clip A depicted on Fig. 11, the purely linear functions for both QoS parameters result in good correlation only within the MOS score range of 2 to 2.5. When a second degree function is used for packet loss, correlation only increased a small amount, with mainly the bottom range of the MOS scale showing better correlation. Using a second degree function for jitter, the correlation is better not only for lower MOS scores but in the whole range. If both functions are second degree, the level of correlation is not changed significantly.

Fig. 12 presents the analysis for Clip B. Linear functions only provide 92.21% correlation, and using a second degree packet loss function does not improve its accuracy perceptively. Second degree variable for jitter shows significant improvement in correlation, reaching 99%. In the MOS range from 2 to 4 it shows quasi-linear behavior in each case.

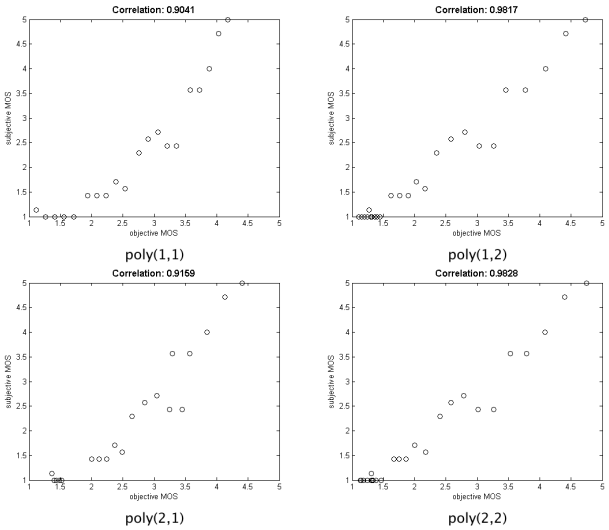


Figure 13. Correlation using different degree polynomials for Clip C

The male voice in Clip C led to similar results in general, but the uniformity in the aforementioned range is not as strong. The benefit of a second degree jitter variable in the function was also confirmed by this evaluation. This is demonstrated in Fig. 13.

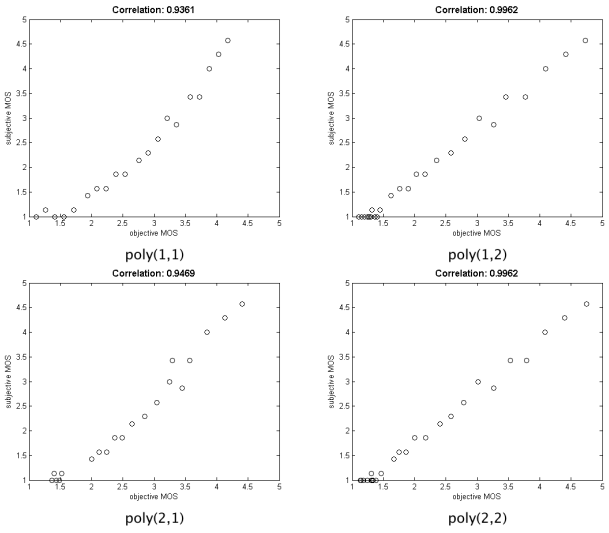


Figure 14. Correlation using different degree polynomials for Clip D

Assessment of Clip D plotted in Fig. 14 resulted in rather uniform results almost through the whole MOS range.

We calculated the overall correlation of first and second degree polynomials. The correlation values are summarized in Table VIII.

TABLE VIII. CORRELATION VALUES FOR FIRST AND SECOND DEGREE POLYNOMIAL REGRESSION ANALYSIS. FIRST PARAMETER: DEGREE FOR PACKET LOSS, SECOND PARAMETER: DEGREE FOR JITTER

	Clip A	Clip B	Clip C	Clip D
poly(1,1)	0.9454	0.9221	0.9041	0.9361
poly(1,2)	0.9861	0.9928	0.9817	0.9962
poly(2,1)	0.9539	0.9324	0.9159	0.9469
poly(2,2)	0.9856	0.9927	0.9828	0.9962

Based on these calculations we decided to use *poly*(1,2) as an optimal choice for the prediction, since increasing the degree of the function would not improve the performance significantly. Fig. 15 plots *poly*(1,2) points for all the audio clips.

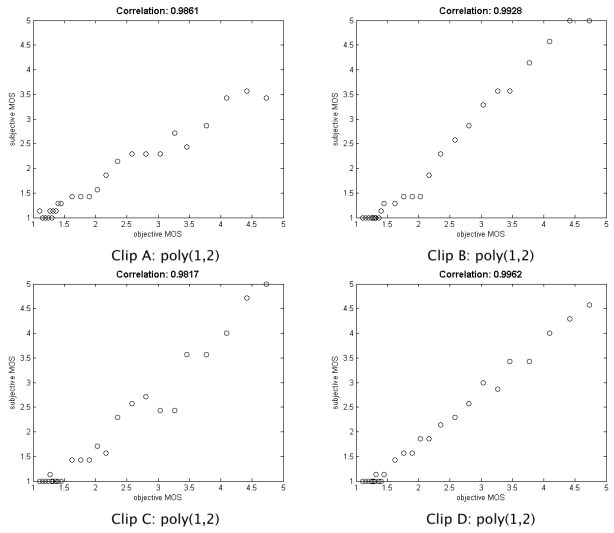


Figure 15. Correlation between subjective and estimated MOS values using first degree variable for packet loss, second degree variable for jitter in the estimator function

VII. CONCLUSION AND FUTURE WORK

We presented a three-phase investigation with the aim of constructing an NR-type objective method for estimating VoIP QoE based on the Opus audio codec. In the first phase of the investigation, we performed subjective QoE evaluations of voice content streamed through an emulated network environment with pre-defined QoS behavior. In the second phase, we determined the degree of relation between QoS parameters and QoE using polynomial regression. We also determined the coefficients for the specific degree of polynomials in the correlation analysis. In the third phase, we repeated the evaluation technique using four independent contents. We calculated the correlation for first and second degree functions in the polynomial regression. We concluded that using first degree functions a correlation of 90% is achievable in view of QoS parameters packet loss and jitter. If a second degree variable is used for the jitter, a correlation above 98% was observed.

Using low degree polynomials, our new method can be the basis of a hardware-accelerated QoE estimation method for predicting perceptual quality in real-time by measuring QoS parameters (packet loss and jitter) and estimating the MOS score on-the-fly without using the original content. We have chosen the Opus audio codec as a reference but our model can also be applied to other audio codecs with a behavior similar to Opus.

ACKNOWLEDGMENT

The research of Tamás Skopkó, Zoltán Nagy and Tamás Lukovics was supported by the TÁMOP 4.2.2.C-11/1/KONV-2012-0001 project. The project was implemented through the New Széchenyi Plan, co-financed by the European Social Fund.

The research of Péter Orosz was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4.A/2-11-1-2012-0001 'National Excellence Program'.

REFERENCES

[1] P. Orosz, T. Skopkó, Z. Nagy, and T. Lukovics, "Performance analysis of the Opus codec in VoIP environment using QoE evaluation," ICSNC 2013, The Eighth International Conference on Systems and Networks Communications, IARIA, pp. 89-93, 2013.

[2] A. Raake, "Speech quality of VoIP: Assessment and prediction," John Wiley & Sons, 2007.

[3] A. Ramö, and H. Toukomaa, "Voice quality characterization of IETF Opus codec," INTERSPEECH, ISCA, pp. 2541-2544, 2011.

[4] C. Hoene, J. M. Valin, K. Vos, and J. Skoglund, "Summary of Opus listening test results draft-valin-codec-results-03," November 2013.

[5] J. M. Valin, G. Maxwell, TB. Terriberry, and K. Vos, "High-Quality, low-delay music coding in the Opus codec," 135th AES Convention, AES p. 8942, October 2013.

[6] ITU-T P.863: Perceptual objective listening quality assessment, January 2011.

[7] F. Neves, S. Cardeal, S. Soares, P. Assunção, F. Tavares, "Quality model for monitoring QoE in VoIP services," International Conference on Computer as a Tool (EUROCON), IEEE, pp.1-4, April 2011, doi:10.1109/EUROCON.2011.5929300

[8] S. Cardeal, F. Neves, S. Soares, F. Tavares, and P. Assunção, "ArQoS®: System to monitor QoS/QoE in VoIP," International Conference on Computer as a Tool (EUROCON), IEEE, pp. 1-2, April 2011, doi:10.1109/EUROCON.2011.5929310

[9] W. Cherif, A. Ksentini, D. Négru, and M. Sidibe, "A_PSQA: PESQ-like non-intrusive tool for QoE prediction in VoIP services," International Conference on Communications (ICC), IEEE, pp. 2124-2128, 2012, doi: 10.1109/ICC.2012.6364004

[10] F. Liu, W. Xiang, Y. Zhang, K. Zheng, and H. Zhao, "A novel QoE-based carrier scheduling scheme in LTE-Advanced

networks with multi-service," Vehicular Technology Conference (VTC Fall), IEEE, pp. 1-5, 2012, doi: 10.1109/VTCFall.2012.6398912

[11] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, and G. Pujolle, "Quality of experience of VoIP service: A survey of assessment approaches and open issues," Communications Surveys & Tutorials, IEEE, pp. 491-513, 2012, doi: 10.1109/SURV.2011.120811.00063

[12] V. Aggarwal, E. Halepovic, J. Pang, S. Venkataraman, and H. Yan, "Prometheus: Toward Quality-of-Experience estimation for mobile apps from passive network measurements," ACM HotMobile'14, ACM, p. 18, February 2014, doi: 10.1145/2565585.2565600

[13] V. Jacobson, R. Frederick, S. Casner, and H. Schulzrinne, "RTP: A transport protocol for real-time applications," RFC 3550, July 2003.

[14] sflPhone, [Online]. Available from: http://sflphone.org/, 20.06.2014

[15] JACK Audio Connaction Kit, [Online]. Available from: http://jackaudio.org/, 20.05.2014

[16] GStreamer open source multimedia framework, plugin list, [Online]. Available from: http://gstreamer.freedesktop.org/documentation/plugins.html, 20.06.2014

[17] netem: Linux Networking Emulator, [Online]. Available from: http://www.linuxfoundation.org/collaborate/workgroups/networking/netem, 20.06.2014

[18] J. Spittka, K. Vos, and J. M. Valin, "RTP Payload for Opus Speech and Audio Codec draft-ietf-payload-rtp-opus-01," August 2013.

A Novel Timing Synchronization Method for MIMO-OFDM Systems

Ali Rachini
IETR - INSA de Rennes, France
GET/UL - Lebanese University
Email: Ali.Rachini@ul.edu.lb

Fabienne Nouvel
IETR - INSA de Rennes
Rennes, France
fabienne.nouvel@insa-rennes.fr

Ali Beydoun and Bilal Beydoun
GET/UL - Lebanese University
Hadath, Lebanon
bilbey@ul.edu.lb
beydoual@yahoo.fr

***Abstract*—To increase the throughput of transmission systems, MIMO-OFDM technology enables better transmission rate and improves the reception. The synchronization between the transmitter and the receiver has become a big challenge. A bad timing synchronization causes the loss of a lot of information in a MIMO-OFDM system. In this paper, a novel timing synchronization method is proposed for a MIMO-OFDM systems with N_t transmit antennas and N_r receive antennas. The proposed method is based on transmit an orthogonal CAZAC sequences over different transmit antennas [1]. Simulations results show that the proposed solution presents a good performance at a low SNR (Signal to Noise Ratio) in AWGN and multipath fading Rayleigh channels, where Doppler Effect is not considered for current simulations. Furthermore, this method can be implemented for MIMO-OFDM system up to 8×8 as well. In the proposed method, the coarse and fine timing synchronization are done at the same time at each receive antenna due to the different training sequences transmitted over different transmit antennas.**

Keywords - MIMO-OFDM system, timing synchronization, CAZAC sequences, compact preamble

I. Introduction

The world of wireless communications and mobile communication is currently at a very important crossroads in its evolution. This crossroads introduces a variety of challenges such as multi-path signal reflections, and interference. These can reduce the performance of a receiver. To address these challenges, OFDM (Orthogonal Frequency Division Multiplexing) modulation combined with MIMO (Multiple Input-Multiple Output) is proposed in 802.11n [2].

The OFDM [3] divides a frequency bandwidth into several (N_{sc}) orthogonal sub-carrier. To simply achieve orthogonal frequency multiplex, Saltzberg and Weinstein [4] have shown that the use of the FFT algorithm (Fast Fourier Transform) can easily produce the OFDM symbol carried by N_{sc} orthogonal subcarriers. The IFFT (Inverse Fast Fourier Transform) algorithm is also used to demodulate the message at the receiver. To maintain the orthogonality, the sub-carrier spacing required is $\Delta f = \frac{1}{T_s}$, where T_s is the duration of OFDM symbol. Due to this orthogonality, OFDM is more resistant to frequency selective fading than FDM (Frequency Division Multiplexing) systems. The orthogonality of frequencies is primordial to eliminate the ICI (Inter Carrier Interference). On the other hand, it should be noted that the most important disadvantage of OFDM is probably the PAPR (Peak-to-Average Power Ratio), leading to replacing it on the up-link in LTE (Long-Term Evolution).

Otherwise, to reduce the ISI (Inter Symbol Interference) due to the multipath fading channel, OFDM technology uses a technique that consists to insert a GI (Guard Interval) [5] or CP (Cyclic Prefix) of duration T_g. The duration of CP should be greater than or equal to the maximum spread of the channel impulse response τ_{max}. The CP is usually a copy of the last portion of an OFDM symbol. Therefore, the useful part T_s of each OFDM symbol will no longer be affected by the ISI. The total duration of the OFDM symbol is $T_{tot} = T_s + T_g$.

Furthermore, MIMO is a wireless technology that uses multiple transmit antennas (N_t) and multiple receive antennas (N_r) to transmit data at the same time. Several applications, based on MIMO technology, have been proposed in various communication standards as WiFi (Wireless Fidelity), WiMax (Worldwide Interoperability for Microwave Access), HSPA+ (evolved High-Speed Packet Access), 3^{rd} and 4^{th} generation of mobile network and LTE. MIMO is based on two main techniques [6], [7]: Spatial Multiplexing techniques (SM) and Spatial Diversity techniques (SD).

1) Spatial Multiplexing technique (SM): The spatial multiplexing technique transmits independent data streams, over different antennas, in order to increase the transmission data rates in a given frequency bandwidth. MIMO systems are mainly used to increase the flow of wireless communications. Foshini et al. [8] and Telatar [9] have shown that the theoretical capacity of the MIMO channel, with N_t transmit antennas and N_r receive antennas, increases linearly with $min(N_t, N_r)$. The channel capacity of a MIMO system is defined by the equation (1) [8], [9]:

$$C = \log_2 \left[\det \left(I_{N_r} + \frac{\rho}{N_t} HH^{\dagger} \right) \right] bps/Hz. \quad (1)$$

with

- N_t : Number of transmit antennas.
- N_r : Number of receive antennas.
- I_{N_r} : Identity matrix $N_r \times N_r$.
- $(.)^{\dagger}$: Conjugate transpose.
- H : MIMO channel matrix $N_t \times N_r$.

- $\rho = \frac{P}{No.B}$: Signal to noise ratio (SNR).
- P : Total transmitted power.
- N_0 : Power Spectral Density (PSD).

2) Spatial Diversity technique (SD): Spatial diversity technique consists of sending the same data stream simultaneously on different transmit antennas. At the receiver, several versions of the signal are received and combined on each of the antennas. This combination reduces the attenuation of the signal and compensates the effect of fading channels. This diversity requires to make use of STC (Space-Time Codes) such as Alamouti codes [10] or trellis codes [11]. Spatial diversity has good efficiency when the MIMO sub-channels are decorrelated to each other. Furthermore, when the number of transmit antennas increases, the power of the received signal increases, thereby improving the detection of received signal.

In this paper, we will focus on spatial diversity technique using STBC (Space-Time Block Code) with Alamouti [10] encoder.

One of the main challenges for MIMO-OFDM system is the synchronization between transmitter and receiver. Two types of synchronization are necessary, the frequency and the timing synchronization. The frequency synchronization is based on the detection of the frequency offset between the transmitter and the local oscillator at the receiver [12]. The CFO (Carrier Frequency Offset) is presented due to the Doppler effect or the difference between the frequency between transmitter and receiver in MIMO-OFDM systems. In the other hand, timing synchronization is divided into two parts: the coarse timing synchronization detects the arrival of the OFDM frame and the fine timing synchronization is needed to detect the beginning of OFDM symbols on each frame. The timing synchronization is very important on each communication system and it is important to have a robust algorithm to estimate the timing synchronization in order to detect the start of each OFDM frame and symbol. In this paper, we focus on the timing synchronization in MIMO-OFDM systems.

In the literature, several synchronization approaches have been proposed for OFDM and MIMO-OFDM systems [13]–[17]. The main idea is the use of good synchronization preamble, at the transmitter, in order to detect the packet arrival, at the receiver.

In this paper, we propose a novel preamble structure for timing synchronization in MIMO-OFDM systems using CAZAC (Constant Amplitude Zero Auto-Correlation) sequences. The CAZAC sequences [18] are a class of complex-valued sequences with cyclic autocorrelation equal to zero. The main characteristics of CAZAC sequences are their zero auto-aorrelation; it means that a CAZAC code is always orthogonal with its cyclic shifted versions. Furthermore, they have a constant amplitude. The main benefits of CAZAC sequences are the reduction of ISI and they avoid interferences between multiple antennas. As a result, CAZAC sequences are regarded as preamble for timing synchronization in MIMO-OFDM systems.

This paper is organized as follows. Section II briefly describes the MIMO-OFDM system structure based on STBC code. Related work are presented in Section III. Section IV introduces the criteria to select a good synchronization sequences. The proposed method and preamble structure are presented in Section V. Simulation results and conclusion are given in Sections VI and VII, respectively.

II. System Model

Like any telecommunications system, MIMO-OFDM system consists of a transmitter, a channel, and a receiver. The transmitter generates OFDM symbols which are modulated using QAM (Quadrature Amplitude Modulation). The OFDM symbols are transmitted over multiple antennas using space time coding block (STBC) [10], [19]. The STBC is a technique to transmit multiple copies of a data stream across N_t transmit antennas in a MIMO system. It exploits the spatial diversity and increases the reliability of transmission. This type of code is divided into three main approaches [19]: OSTBC (Orthogonal Space-Time Block Codes), NOSTBC (Non-Orthogonal Space-Time Block Codes) and QSTBC (Quasi-Orthogonal Space-Time Block Codes). Our approach is based on the STBC using Alamouti code [10].

A. MIMO-OFDM System

Figure 1 presents a MIMO-OFDM system with N_t transmit antennas, N_r receive antennas and N_{sc} subcarriers per transmit antenna. The data block is used in order to generate the bitstream, then a serial/parallel (S-to-P) converter is used to distribute the datastream to the QAM modulator (Sub carrier mapping). This modulator maps the bit stream according to the 32-QAM modulation. The QAM symbols are then introduced into a STBC encoder. Then, we apply the IFFT to generate the OFDM symbols. A cyclic prefix block (Add CP) consist to insert a guard interval between successive OFDM symbols, which the last T_G seconds of each symbol is appended to the beginning of each OFDM symbol. The CP is used in order to minimize the ISI and it will be removed at the receiver before to the FFT. The synchronization block is used in order to insert the synchronization preamble at the beginning of each OFDM frame. Two different approaches are presented, the synchronization preamble is appended in frequency domain or in time domain. In this paper, we focus on the first approach. In [20], we proposed another method that generates the synchronization preamble in time domain.

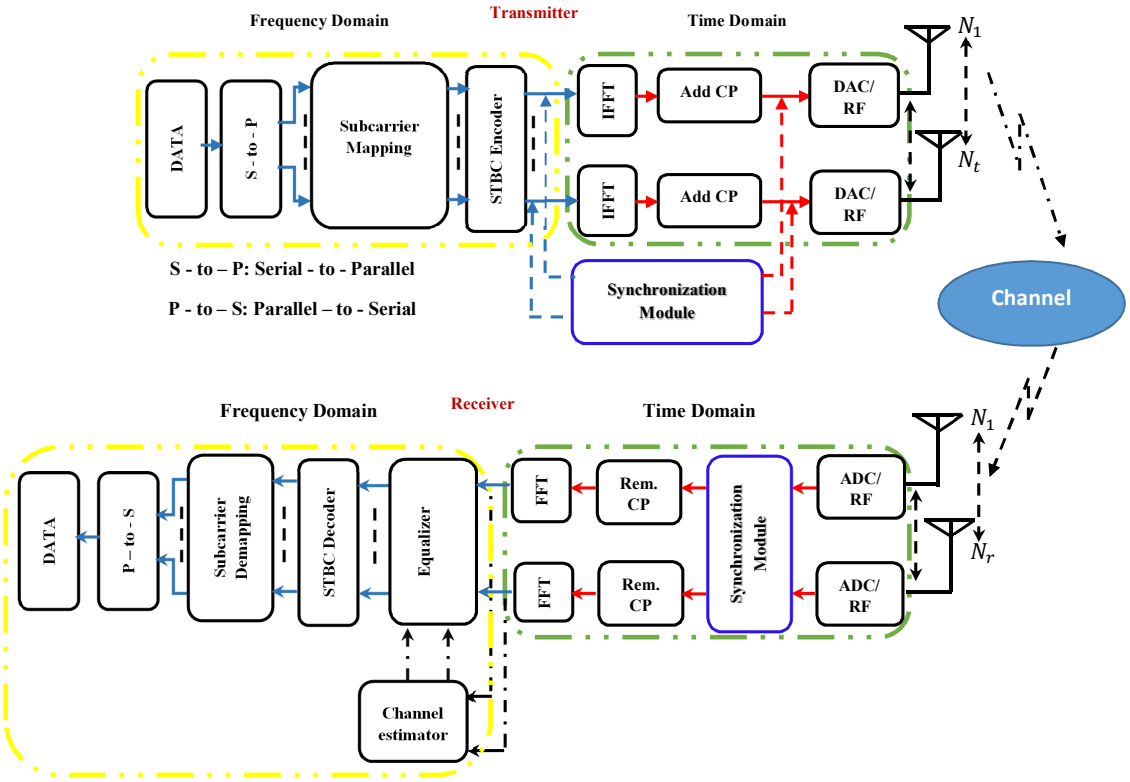


Fig. 1: Transmission system MIMO-OFDM-STBC

The transmitted signal s_i on each transmit antenna T_i is given by:

$$s_i(t) = \frac{1}{\sqrt{N_{sc}}} \sum_{k=0}^{N_{sc}-1} \mathfrak{Re}\left\{x_k e^{j.2\pi.f_k.t}\right\} \quad (2)$$

where x_k is the symbol on the frequency f_k.

The second part of MIMO-OFDM system is the receiver. The first block after the analog to digital converter (ADC) is the timing synchronization block. After a good timing synchronization, the cyclic prefix of each OFDM symbol is eliminated. The signal returns back into frequency domain thanks to FFT block. The equalization is used in order to fight the ISI problem, which is created by multipath fading channels, taking into account the channel estimation coefficients. The estimated symbols are decoded and combined by the STBC decoder. Then, a QAM demodulator is used to demodulate and recover the binary information. The parallel/serial (P-to-S) converter allows to reformatting the binary bit stream.

B. Channel Model

The transmitted signal reaches the receiver over different paths due to the multipath channel. The channel between the transmit antennas T_i and receive antenna R_j, $i \in \{1, N_t\}$ and $j \in \{1, N_r\}$, is given by:

$$H(t) = \sum_{l=1}^{L} H_l \delta(t - \tau_l) \quad (3)$$

where H_l are the matrix coefficients of the l^{th} path. This matrix is $N_t \times N_r$. δ is a pulse function and L is the maximum number of multipaths. H_l is given by:

$$H_l = \begin{bmatrix} h^l_{1,1} & h^l_{1,2} & \dots & h^l_{1,N_r} \\ h^l_{2,1} & h^l_{2,2} & \dots & h^l_{2,N_r} \\ \vdots & \vdots & \ddots & \vdots \\ h^l_{N_t,1} & h^l_{N_t,2} & \dots & h^l_{N_t,N_r} \end{bmatrix} \quad (4)$$

C. Received signal

The received signal r_j on each receive antenna R_j is given by:

$$r_j(t) = \sum_{i=1}^{N_t} \sum_{l=1}^{L} \left(h^l_{ij}(t).s_i(t) \right) + w_j(t) \quad (5)$$

where s_i is the transmitted signal on the transmit antenna T_i, w_j is the additive white Gaussian noise (AWGN) and h^l_{ij} represents the corresponding channel between T_i and R_j.

The matrix representation of the received signal is given by:

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{N_r} \end{bmatrix} = \begin{bmatrix} H_{1,1} & \dots & H_{1,N_r} \\ H_{2,1} & \dots & H_{2,N_r} \\ \vdots & \ddots & \vdots \\ H_{N_t,1} & \dots & H_{N_t,N_r} \end{bmatrix} \times \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{N_t} \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{N_r} \end{bmatrix} \quad (6)$$

III. Related work

In the literature, several synchronization approaches have been proposed for OFDM and MIMO-OFDM as shown in Section I. The most of the synchronization methods are preamble based.

Authors in [21] provides a method for timing and frequency synchronization for a MIMO-OFDM system using the LS (Loosely Synchronous) codes to detect the beginning of each received frame. LS codes have a good autocorrelation and cross-correlation functions within certain vicinity of the zero shifts. In this method, the synchronization process is divided into four steps. The first step is to estimate the coarse timing synchronization and, the second step is to estimate the coarse frequency synchronization. The third step detects the OFDM symbols and estimates the channel using the LS codes, the fourth step is for the fine frequency estimation. Furthermore, the structure of preambles used in the process is relatively complex.

Authors in [22] proposed a timing synchronization method based on OVSF (Orthogonal Variable Spreading Factor). Simulations results are done using MISO (Multiple Input-Single Output) systems 2×1. The length of each OFDM symbol is 256, and the length of the CP is 32. A synchronization preamble is appended at the beginning of each OFDM frame, this preamble has the same length as the CP. For MISO-OFDM systems 2×1, the timing acquisition probability is 1 for $SNR \geq -5\,dB$. The synchronization preamble for this method is appended in the time domain, this is the main drawback of this method. With such implementation, authors need an extra block to insert the preamble, while, in frequency domain we do not need any extra block thanks to IFFT.

Based on our proposed approach, a compact preamble design for synchronization in distributed MIMO-OFDM systems has been proposed in [23]. The authors have designed training symbols based on exclusive subband, where different adjacent subbands are spaced by guard bands to reduce their mutual interference. In [23], authors propose a compact preamble structure that has the same length as an OFDM symbol. This preamble, based on CAZAC sequences, is carried by a different subbands. The simulation results shown that for a MISO-OFDM (3×1), the acquisition probability for timing synchronization is 70% for an $SNR = 5\ dB$. In this work, we compare the simulation results of our proposed [1] with those of the method proposed by Chin-Liang et al [23].

The proposed method in [23] has several limitations such as the complexity of the structures to generate the synchronization preambles for a large number of transmit antennas. In [23], when the number of transmit antennas increases, the size of sub-bands should be reduced to take into account all transmit antennas. Therefore, at the receiver, the probability of synchronization decreases due to the length of synchronization sequence.

IV. Selection criteria for synchronization sequences

Most of timing synchronization methods are based on preamble approach. They consist in sending a synchronization preamble at the beginning of each OFDM frame. The main characteristic of synchronization preamble is to have good autocorrelation and cross-correlation properties in order to detect a correlation peak as closed as possible to a Dirac pulse. On the other hand, the synchronization preambles should be orthogonal to eliminate the interferences between preambles at simultaneous transmissions in MIMO systems. The main sequences used in the state of art are listed below.

A. Gold sequences

Gold sequences have been proposed by Robert Gold in 1967 [24]. These sequences are constructed by two PN (Pseudorandom Noise) sequences of the same length. The XOR ($\oplus$) is the operation to generate Gold sequences. Let x and y be two PN maximum length sequences of the length $L_C = 2^l - 1$. The set $S_{gold}(x, y)$ of Gold sequences is given by:

$$S_{gold}(x, y) = \{x, y, x \oplus y, x \oplus T^{-1}y, \dots, x \oplus T^{-(L_C-1)}y\} \quad (7)$$

where T^{-p} is an operator of cyclic shifts of p values to the left.

The cross-correlation peak of some periodic Gold sequences take one of the three possible values -1, $-t(l)$ or $t(l) - 2$, where $t(l)$ is given by:

$$t(l) = \begin{cases} 2^{(l+2)/2} + 1 & \text{if } l \text{ is even} \\ 2^{(l+1)/2} + 1 & \text{if } l \text{ is odd} \end{cases} \quad (8)$$

The disadvantage of Gold sequences is inter-code interference due to high cross-correlation. In this paper, the synchronization preamble is appended at the beginning of each OFDM frame and is transmitted aperiodic. Therefore, the autocorrelation and cross-correlation functions of aperiodic Gold sequences are shown in Figure 2 where index represents indices at which the correlation was estimated.

B. Walsh-Hadamard code

Walsh-Hadamard code are built recursively from a 2×2 Hadamard's matrix [25]. These codes are orthogonal and bipolar, where the Hadamard's matrix is given by:

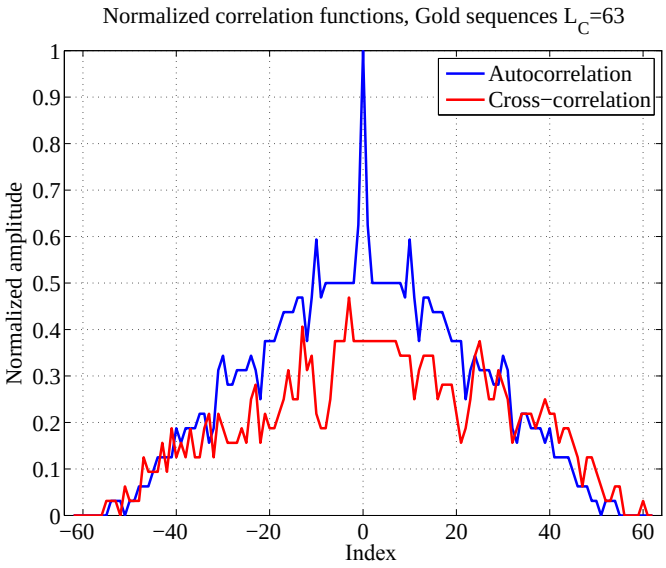


Fig. 2: Autocorrelation and cross-correlation of Gold sequence, $L_C = 63$

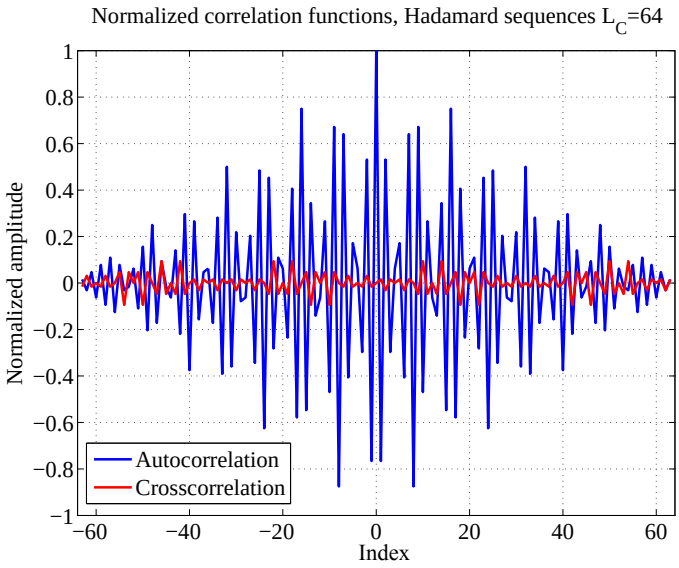


Fig. 3: Autocorrelation and cross-correlation of Hadamard sequence, $L_C = 64$

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \ldots, H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}$$

$$= H_2 \otimes H_{2^{k-1}} for 2 \leq k \in \mathbb{N} \quad (9)$$

where $\otimes$ denotes the Kronecker product.

An Hadamard matrix H_n satisfies the following property:

$$H_n . H_n^T = n I_n$$

where H_n^T is the conjugate transpose of H_n and I_n is a $n \times n$ identity matrix. The autocorrelation and cross-correlation functions of Hadamard sequences are shown in Figure 3.

C. CAZAC sequences

The CAZAC (Constant Amplitude Zero Auto-Correlation) sequences [26] are a class of complex-valued sequences with cyclic autocorrelation equal to zero. CAZAC sequences are characterized by their constant amplitude and their correlation function. The cross-correlation function between two different CAZAC sequences is near to zero and the autocorrelation function is as δ. Let $x_u(k)$ be a CAZAC sequence of length L_C, $x_u(k)$ is given in equation (10):

$$C(k) = \begin{cases} e^{j\left(\frac{\pi M k(k+1)}{L_C}\right)} & \text{if } L_C \text{ is odd} \\ e^{j\left(\frac{\pi M k^2}{L_C}\right)} & \text{if } L_C \text{ is even} \end{cases} \quad (10)$$

where $L_C = 2^n$ is the length of the CAZAC sequence, $n \in \mathbb{N}$, $M \in \mathbb{N}$ is a prime number with L_C and $k \in \{0, L_C - 1\}$ is the index of the sample.

Let $c(m)$ be the corresponding of $C(k)$ in the time domain after the IFFT. It should be noted that $c(m)$ is also a CAZAC sequence of length L_C [27]. This sequence $c(m)$ is given by the equation (11):

$$c(m) = \frac{1}{L_C} \sum_{k=0}^{L_C - 1} C(k) . e^{j\left(\frac{2\pi}{L_C}\right) mk}, m \in [0,\ L_C - 1] \quad (11)$$

The autocorrelation and cross-correlation functions of CAZAC sequences are shown in Figure 4.

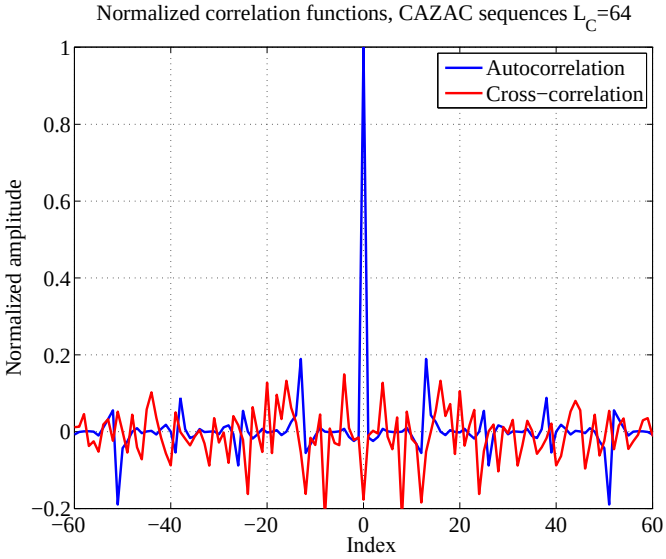


Fig. 4: Autocorrelation and cross-correlation of CAZAC sequence, $L_C = 64$

D. Selection criteria

Both Gold and Hadamard sequences have good correlation functions, but Gold sequences have a high value for their cross-correlation function, while for some Hadamard sequences, the autocorrelation function has secondary correlation peaks. In the other hand, and as shown in Figure 4, CAZAC sequences have a good autocorrelation and cross-correlation functions. In this paper, our work was focused on the use of CAZAC sequence as synchronization sequences in the synchronization preamble with different structures.

V. Proposed method for timing synchronization

In this section, we present our timing synchronization method based on CAZAC sequences. We define, in this section, C as a CAZAC sequence of length L_C, where:

$$L_C = L_{FFT}/2 \tag{12}$$

L_{FFT} is the size of the FFT and $-C^*$ is the minus conjugate of C of length L_C. This approach is based into two structures.

A. First preamble structure

The proposed timing synchronization structure, in this section, relies on sending a preamble structure performed in the frequency domain, as shown in Figure 5.

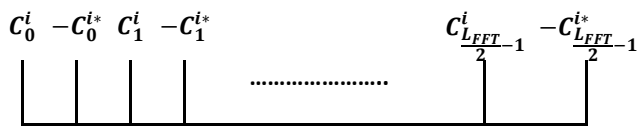


Fig. 5: General preamble structure in frequency domain for the first structure

In Figure 5, each preamble contains a CAZAC (C) sequence mapped on the odd subcarrier, in frequency domain, and the $-C^*$ is mapped on the even subcarrier. Let $X_u^i(k)$ be the preamble that appended at the beginning of each OFDM frame in frequency domain on the i^{th} transmit antenna. This preamble can be given by the following equation:

$$X_u^i(k) = \begin{cases} C^i\left(\frac{k}{2}\right) & if \quad k \mod 2 = 0 \\ -C^{i*}\left(\frac{k-1}{2}\right) & if \quad k \mod 2 \neq 0 \end{cases} \tag{13}$$

where $k \in \{0, L_{FFT}-1\}$ and $L_{FFT} = 2.L_C$.

Figure 6 presents the real and the imaginary parts of the synchronization preamble in time domain. The combination of a CAZAC sequence C with $-C^*$ gives a time-domain complex envelope form that have a good autocorrelation and cross-correlation functions. This combination does not destroy the orthogonality between subcarriers, and it retains the orthogonality between different preambles over different transmit antennas.

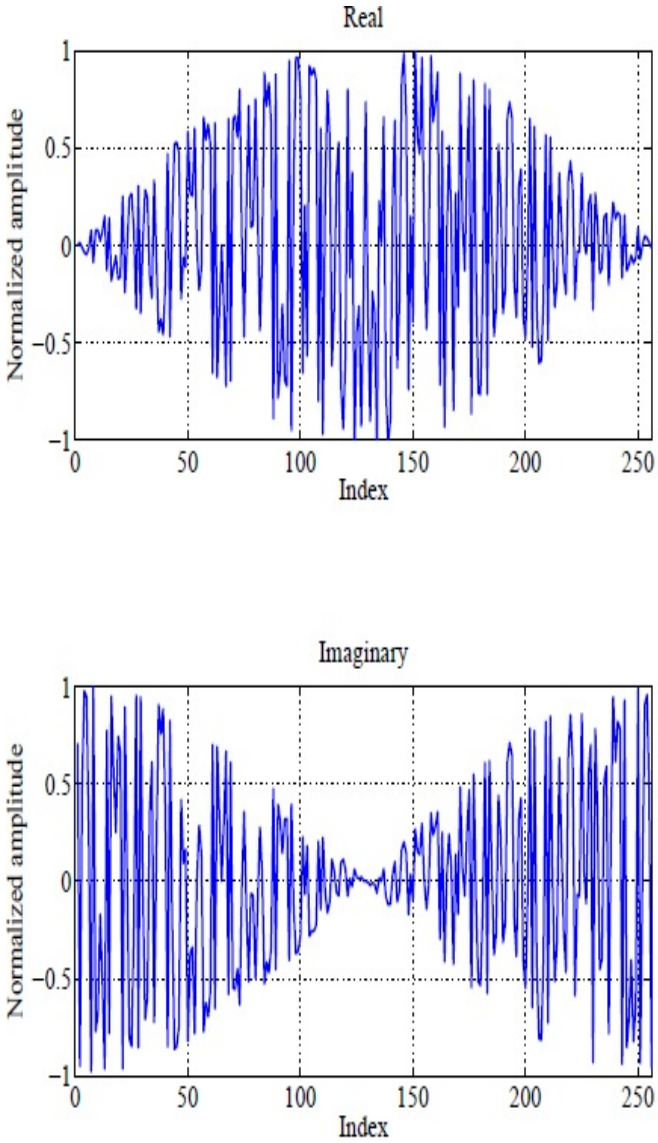


Fig. 6: Real and imaginary parts of the first preamble structure in time domain

Figure 7 presents the autocorrelation and the cross-correlation of the proposed preamble in time domain. This preamble shows a good correlation function in order to detect the timing synchronization peak.

Figure 8 represents different CAZAC sequences transmitted over different transmit antennas. The term C_k^i is the sample of the CAZAC sequence carried by the k^{th} subcarrier and transmitted by the transmitting antenna T_i. The proposed method can be applied regardless of the number of transmit or receive antennas.

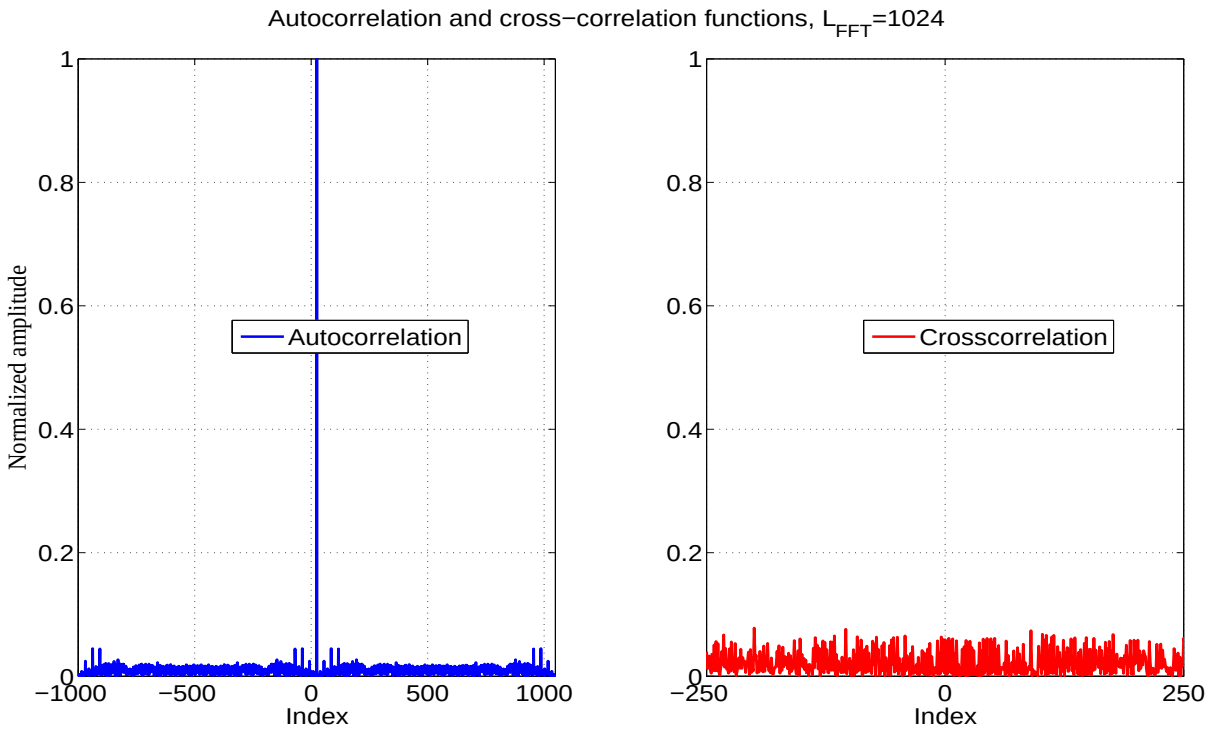


Fig. 7: Autocorrelation and cross-correlation functions of the first preamble structure ($L_{FFT} = 1024$)

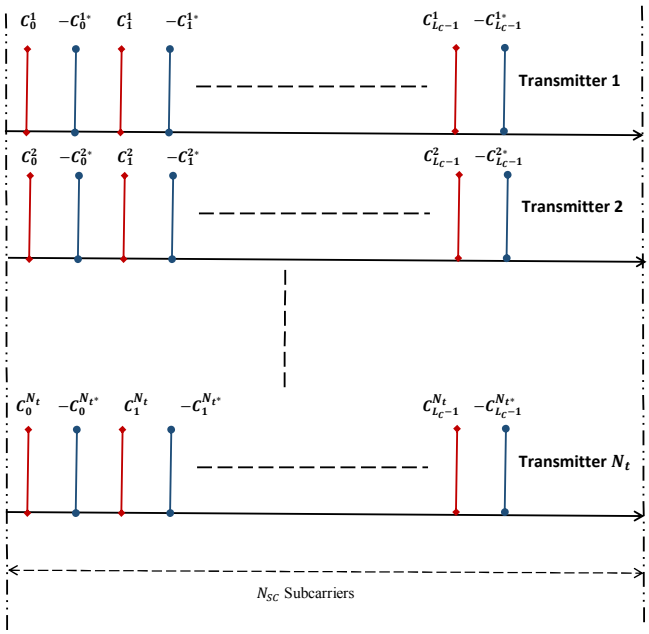


Fig. 8: First preamble structure in frequency domain over different transmit antennas

B. Second preamble structure

The second preamble structure is based on the first preamble structure, where the preamble is divided into two parts of length L_C each one. The first part contains a complete CAZAC sequence C, where the second part contains the minus conjugate of C denoted $-C^*$. Different CAZAC sequences are transmitted over different transmit antennas. This preamble structure is presented in Figure 9 in frequency domain.

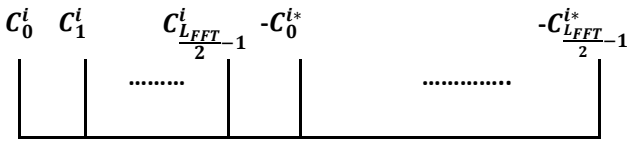


Fig. 9: General preamble structure in frequency domain for the second structure

The autocorrelation and cross-correlation functions of this preamble in time domain are presented in Figure 10. The autocorrelation function of this preamble has a secondary peak at 30% of the normalized autocorrelation peak.

Let X_u^i be the preamble sent on the i^{th} transmit antenna, the equation of this preamble in frequency domain is given by:

$$X_u^i(k) = \begin{cases} C^i(k) & if \quad 0 \leq k \leq L_C - 1 \\ -C^{i*}(k - L_C) & if \quad L_C \leq k \leq L_{FFT} - 1 \end{cases} \quad (14)$$

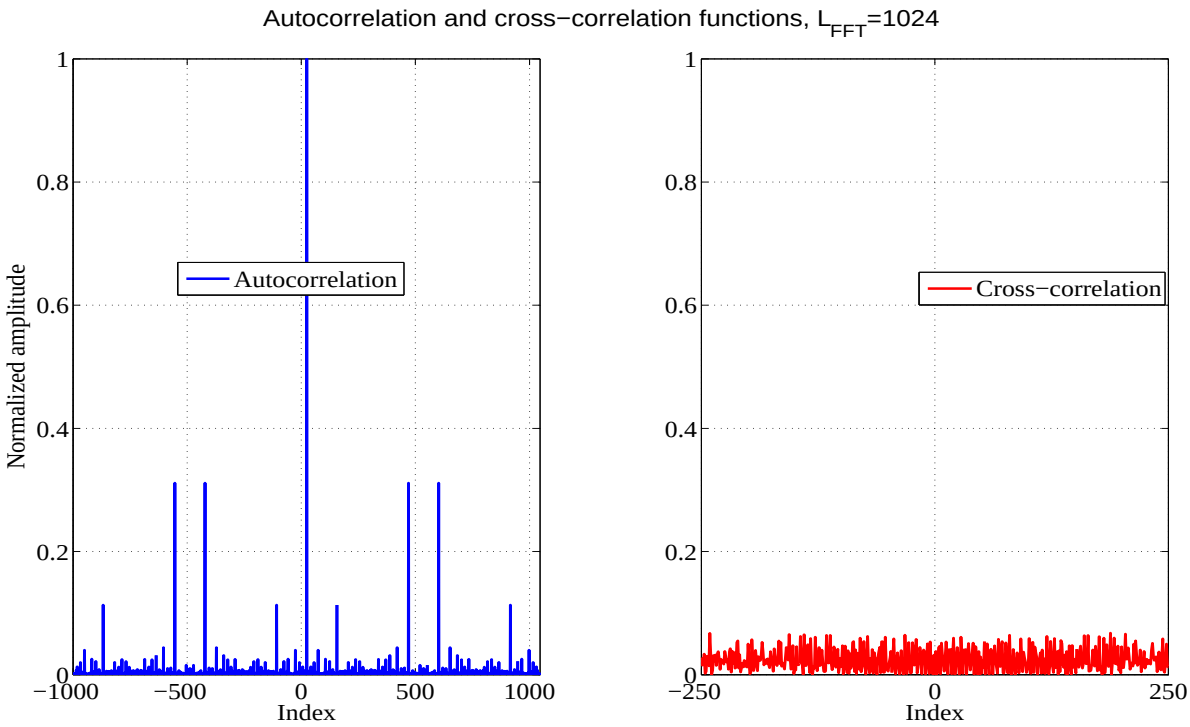


Fig. 10: Autocorrelation and cross-correlation functions of the second preamble structure ($L_{FFT} = 1024$)

VI. Simulations results

The simulations have been done, in both AWGN channel and multipaths fading channel, to evaluate the performance of our proposed preamble against previous proposed preamble.

A. Simulation parameters

A SISO-OFDM (Single Input-Single Output) and MIMO-OFDM systems up to 8×8 transmit and receive antennas were simulated. An OFDM system with 512 and 1024 sub-carriers ($L_{FFT} = 512, L_{FFT} = 1024$ resp.) was considered in Rayleigh multipath fading channel with 6 paths sample-spaced with T_s (Sampling Time), suggested by the IEEE 802.11 Working Group [28]. The simulation parameters are summarized in the following Tables I and II.

TABLE I: Simulation parameters

Simulation Parameters	**Value**
MIMO	up to 8×8
FFT/IFFT Length	1024 & 512
Cyclic Prefix Length	$L_{FFT}/4$
Channel Type	Multi-path Rayleigh and AWGN channel
Sequences	CAZAC
Length of orthogonal code L_C	$L_{FFT}/2$
Number of synchronization symbol	1
Number of channel taps between different antennas	6
SNR over all the OFDM Frame	from 0dB to 25dB

TABLE II: The average power profile of the multipath Rayleigh channel model

Propagation delay between different multipath	[0.T_s, 1.T_s, 2.T_s, 3.T_s, 4.T_s, 5.T_s]
The power of each multipath	[0.8111, 0.1532, 0.0289, 0.0055, 0.0010, 0.0002]

B. Timing synchronization

In order to detect the timing synchronization peak, a correlation function $\mathcal{R}_{r_j,seq_j}$ between the received signal r_j [equation (5)] and the local sequence seq_j at the receive antenna R_j is applied in time domain. Due to the distribution of CAZAC sequence C and $-C^*$ in each preamble, the correlation between received signal and the local sequence may give a high peak's value. The correlation function is calculated as:

$$\mathcal{R}_{r_j,seq_j}(n) = \sum_{n=1}^{L} [r_j(n) * seq_j(n-\tau)] \qquad (15)$$

where n is the index of the sample.

By correlating the received signal r_j with a local sequence seq_j at each receive antenna R_j, a timing synchronization estimate ($\hat{ind}_n$) is given by:

$$\hat{ind}_n = \underset{n}{\operatorname{argmax}}\{\|\mathcal{R}_{r_j,seq_j}(n)\|\} \qquad (16)$$

The $\hat{ind}_n$ is the timing estimate where n is considered as the coarse timing synchronization point. By shifting the FFT

window, we can find the fine timing synchronization or the beginning of the OFDM symbols on each frame.

C. Simulation results for the first preamble structure

The first preamble structure, presented in Section V, were simulated using the simulation parameters in Tables I and II. The probability of successful timing frame synchronization P_{SYNC} is evaluated in Figures 11 and 12. They show the acquisition probability for different OFDM systems (SISO-OFDM, MIMO-OFDM up to 8×8), using CAZAC sequences. The lengths of preamble are 1024 and 512, respectively ($L_{FFT} = 1024, L_{FFT} = 512$ resp.).

Figure 11 shows that the system has achieved a synchronization probability over 95% at an $SNR = -5dB$ for both SISO-OFDM and MIMO-OFDM 2×2 systems. For a MIMO-OFDM 4×4 system the $P_{SYNC} > 97\%$ at an $SNR = 0\,dB$. On the other hand, for MIMO-OFDM 8×8 system, the acquisition probability P_{SYNC} reaches 90% at an $SNR = 0\,dB$.

In Figure 12 and for $L_{FFT} = 512$, it should be mentioned that the acquisition probability P_{SYNC} is greater than 99% for both SISO-OFDM and MIMO-OFDM 2×2 systems at an $SNR = -5\,dB$, and $P_{SYNC} > 90\%$ for MIMO-OFDM 4×4 system at an $SNR = 0\,dB$. For MIMO-OFDM 8×8 system the P_{SYNC} reaches 80% at an $SNR = 3\,dB$.

For MIMO-OFDM systems, the limitations of all timing synchronization methods, are the number of transmit antennas and the length of the synchronization preamble.

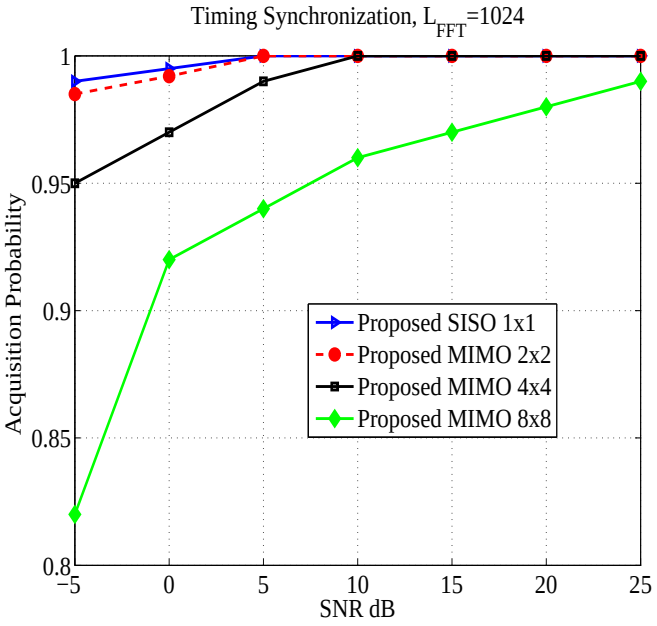


Fig. 11: Timing synchronization performance of the first proposed approach ($L_{FFT} = 1024$)

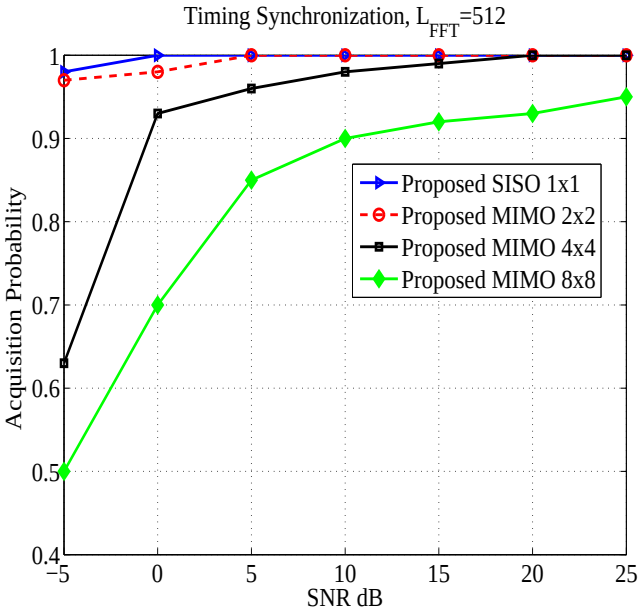


Fig. 12: Timing synchronization performance of the first proposed approach ($L_{FFT} = 512$)

The simulation results of our proposed method (Figure 11) shows a good performance in terms of timing synchronization acquisition probability for MIMO-OFDM system up to 8×8, with a preamble length equal to 1024. A degradation of performances in term of P_{SYNC} (Figure 12) is observed when the length of preamble is smaller than the length used in Figure 11. The degradations of P_{SYNC}, for a large number of transmit antennas, are due to the length of the synchronization preamble. Otherwise, at the receiver, as the length of the preamble is longer then the probability of timing synchronization is higher.

In Figure 13, we compared the performance between our proposed approach and the synchronization scheme in [23]. Based on simulation parameters presented in Tables I and II, our proposed method has a good performance against those obtained by [23]. Furthermore, this figure shows that the acquisition probability P_{SYNC} of our proposed method is greater than 90% for a $SNR > 3\,dB$ for both MIMO-OFDM 2×2 and 3×3 systems, while the method proposed in [23], shows that the acquisition probability is between 50% and 75% for a $SNR > 3\,dB$. On the other hand, and for both MIMO-OFDM 2×2 and 3×3 systems, our method presents a perfect timing synchronization for a $SNR > 10\,dB$.

D. Simulation results for the second preamble structure

This section presents the simulation results for the second preamble structure presented in Section V. This simulation are performed using the simulation parameters in Tables I and II.

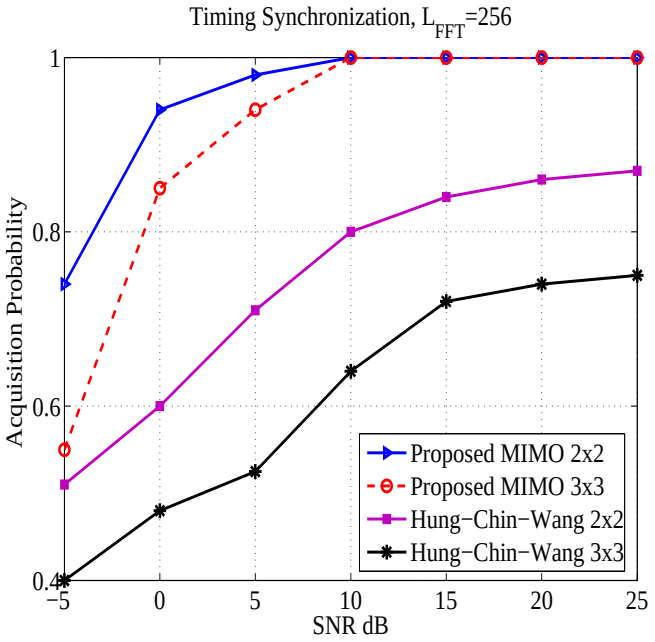


Fig. 13: Comparisons between the first proposed approach and subband-based preamble proposed in [23]

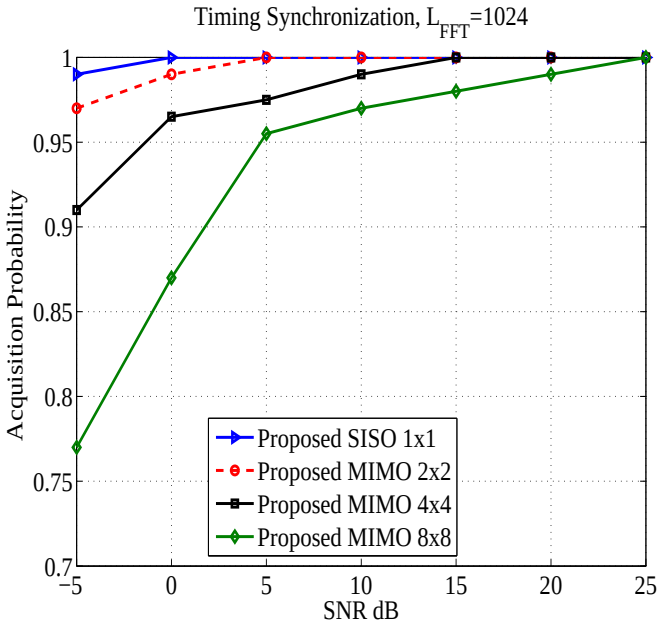


Fig. 14: Timing synchronization performance of the second proposed approach ($L_{FFT} = 1024$)

The timing acquisition probability P_{SYNC} for different size of FFT ($LFFT = 1024$ and $LFFT = 512$) is shown in Figures 14 and 15, respectively.

Figure 14 shows that for a $L_{FFT} = 1024$ and a $SNR \geq -5\,dB$, the acquisition probability P_{SYNC} is greater than 95% for both SISO-OFDM and MIMO-OFDM 2×2 systems. Otherwise, the synchronization probability is perfect from a $SNR \geq 5\,dB$ for the sam systems. On the other hand, for a MIMO-OFDM 4×4 system the P_{SYNC} is greater than 96% for a $SNR \geq 0\,dB$, and its equal to 1 for a $SNR \geq 15\,dB$. Therefore, the acquisition probability for MIMO-OFDM 8×8 system, reaches 90% for an $SNR > 2\,dB$.

In Figure 15, the simulation results are performed with an $L_{FFT} = 512$. This figure shows that the acquisition probability P_{SYNC} is greater than 99% for both SISO-OFDM and MIMO-OFDM 2×2 systems at an $SNR = 0\,dB$, and $P_{SYNC} > 90\%$ for MIMO-OFDM 4×4 system at an $SNR = 0\,dB$. For MIMO-OFDM 8×8 system the P_{SYNC} reaches 90% at an $SNR = 13\,dB$.

Other simulations are performed to compare our second proposed approach with the method proposed in [23], using the same simulation parameters of Tables I and II. Results are shown in Figure 16. In this figure, the size of the preamble is $L_{FFT} = 256$. As shown in this figure, the timing synchronization acquisition probability for our proposed approach is better than the proposed method in [23].

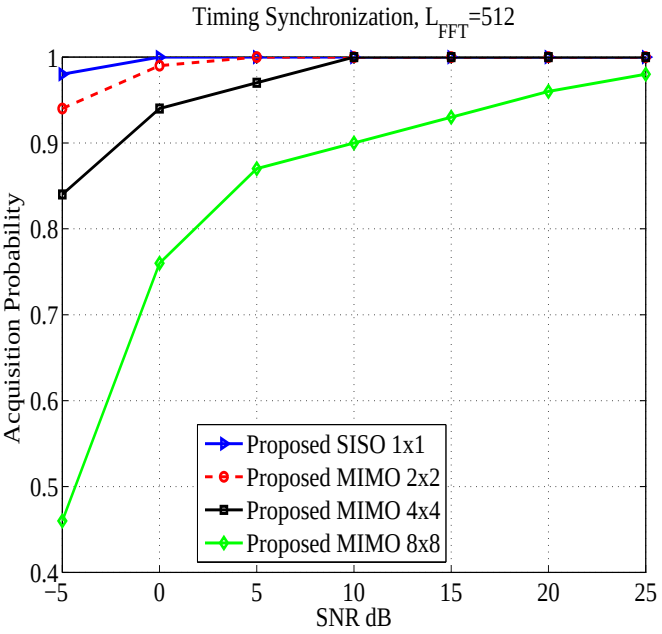


Fig. 15: Timing synchronization performance of the second proposed approach ($L_{FFT} = 512$)

E. Performance of our proposed approach

In this section, we present a comparaison between our proposed approach, in order to find the most efficient among them. In the Figure 17, we obtained acquisition probability of our proposed approaches for MIMO-OFDM system 4×4, and using the simulation parameters of Tables I and II.

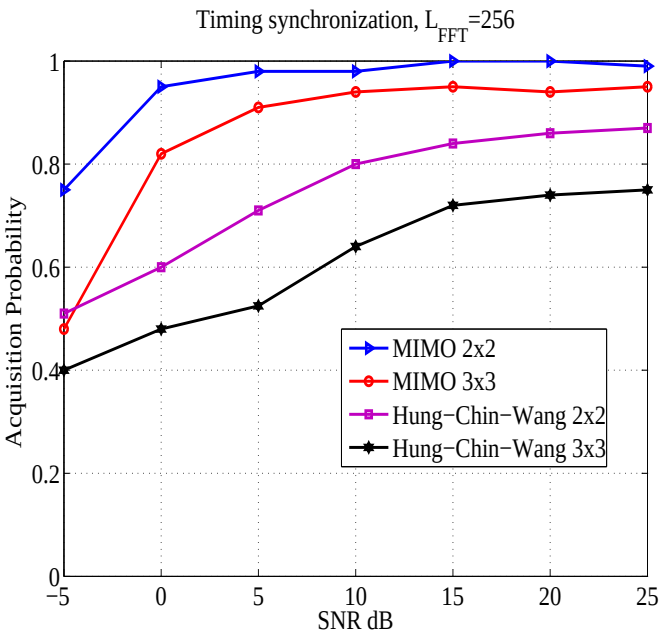


Fig. 16: Comparisons between the second proposed approach and subband-based preamble proposed in [23]

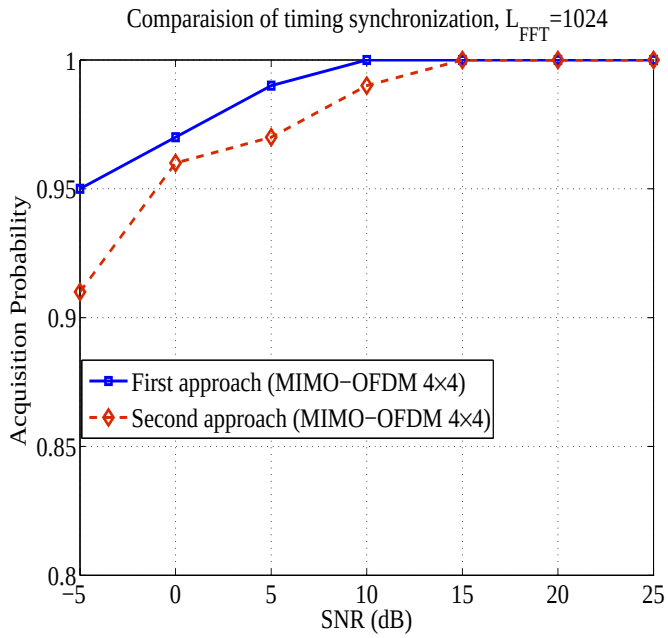


Fig. 17: Comparisons between the first and the second proposed approach for MIMO-OFDM system 4×4

The two approaches present a good timing synchronization acquisition probability at a low SNR. On the other hand, the acquisition probability of the first approach is perfect for an $SNR \geq 10\,dB$, while its equal to 99% for the second approach for the same SNR. Therefore, the first approach presents a high performances versus the second approach. This performance is due to the good correlation function of the synchronization preamble.

VII. Conclusion

One of the main challenges in a MIMO-OFDM system is the detection of the beginning of the received OFDM frames, this step called timing synchronization. The most of the timing synchronization methods are preamble based. In this paper, we proposed a novel compact preamble structure, in order to detect the timing frame synchronization. This structure is divided into two approaches. At the transmitter, a synchronization preamble is appended at the beginning of each OFDM frame. This preamble is based on CAZAC sequences, where those sequences have a good autocorrelation and cross-correlation functions.

At the receiver, and due to this combination, the correlation function between received signal and a local sequence shows a good frame detection with a large number of transmit antennas up to 8×8. In comparison to the subband preamble based proposed by [23], our timing synchronization approaches present a better timing frame synchronization at a low SNR. Finally, we present the performance of our approaches, we find a few degradation of performance between our two approaches due to the correlation functions of different preambles structure. Our approaches can be implemented in MIMO-OFDM systems, with a large number of transmit antennas.

Even though there is still a large number of open research problems in the area of MIMO-OFDM systems. The proposed methods in this paper are presented for timing synchronization for such systems. Hence, it will be interesting to see the performance of our approach for frequency synchronization. Future work will assess, also, the performance of our approach in other communications systems.

Acknowledgment

This work was supported by the GET of Lebanese University and IETR of Rennes-France.

References

[1] A. Rachini, A. Beydoun, F. Nouvel, and B. Beydoun, “A novel compact preamble structure for timing synchronization in mimo-ofdm systems using cazac sequences,” in International Conference on Communications, Computation, Networks and Technologies (INNOV), 2013, pp. 1–6.

[2] IEEE Standard 802.11n: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput, Institute of Electrical and Electronics Engineers, Oct. 2009.

[3] R. Prasad, OFDM for Wireless Communications Systems. Norwood, MA, USA: Artech House, Inc., 2004.

[4] B.R.Saltzberg and S.B.Weinstein, “Fourier transform communication systems,” in Computer Machinery Conference, Pine Mountain, 1969.

[5] A. Peled and A. Ruiz, “Frequency domain data transmission using reduced computational complexity algorithms,” in Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP ’80., vol. 5, Apr. 1980, pp. 964–967.

[6] D. Dromard and D. Seret, ”Architecture des réseaux”, 2nd ed. Pearson, Jun. 2013.

[7] H. Jafarkhani, Space-Time Coding: Theory and Practice, 1st ed. New York, NY, USA: Cambridge University Press, 2010.

[8] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wireless Personal Communications, vol. 6, 1998, pp. 311–335.

[9] I. E. Telatar, "Capacity of multi-antenna gaussian channels," European Transactions On Telecommunications, vol. 10, no. 6, Dec. 1999, pp. 585–595.

[10] S. Alamouti, "A simple transmit diversity technique for wireless communications," vol. 16, no. 8, Oct. 1998, pp. 1451 –1458.

[11] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," Information Theory, IEEE Transactions on, vol. 44, no. 2, 1998, pp. 744–765.

[12] V. Tarokh, A. Naguib, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," Communications, IEEE Transactions on, vol. 47, no. 2, Feb. 1999, pp. 199 –207.

[13] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for ofdm," Communications, IEEE Transactions on, vol. 45, no. 12, Dec. 1997, pp. 1613 –1621.

[14] U. Rohrs and L. Linde, "Some unique properties and applications of perfect squares minimum phase cazac sequences," in Communications and Signal Processing, 1992. COMSIG '92., Proceedings of the 1992 South African Symposium on, Sept. 1992, pp. 155 –160.

[15] B. Yang, K. Letaief, R. Cheng, and Z. Cao, "Burst frame synchronization for ofdm transmission in multipath fading links," in Vehicular Technology Conference, 1999. VTC 1999 - Fall. IEEE VTS 50th, vol. 1, 1999, pp. 300 –304 vol.1.

[16] J.-J. van de Beek, M. Sandell, M. Isaksson, and P. Ola Borjesson, "Low-complex frame synchronization in ofdm systems," in Universal Personal Communications. 1995. Record., 1995 Fourth IEEE International Conference on, Nov. 1995, pp. 982 –986.

[17] L. Li and P. Zhou, "Synchronization for b3g mimo ofdm in dl initial acquisition by cazac sequence," in Communications, Circuits and Systems Proceedings, 2006 International Conference on, vol. 2, Jun. 2006, pp. 1035 –1039.

[18] R. Frank, S. Zadoff, and R. Heimiller, "Phase shift pulse codes with good periodic correlation properties (corresp.)," Information Theory, IRE Transactions on, vol. 8, no. 6, Oct. 1962, pp. 381 –382.

[19] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," Information Theory, IEEE Transactions on, vol. 45, no. 5, 1999, pp. 1456–1467.

[20] A. Rachini, A. Beydoun, F. Nouvel, and B. Beydoun, "Timing synchronisation method for mimo-ofdm system using orthogonal preamble," in Telecommunications (ICT), 2012 19th International Conference on, 2012, pp. 1–5.

[21] W. Jian, L. Jianguo, and D. Li, "Synchronization for mimo ofdm systems with loosely synchronous (ls) codes," in Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on, Sept. 2007, pp. 254 –258.

[22] G. Gaurshetti and S. Khobragade, "Orthogonal cyclic prefix for time synchronization in mimo-ofdm," International Journal of Advanced Electrical and Electronics Engineering (IJAEEE), vol. 2, 2013, pp. 81–85.

[23] H.-C. Wang and C.-L. Wang, "A compact preamble design for synchronization in distributed mimo ofdm systems," in Vehicular Technology Conference (VTC Fall), 2011 IEEE, Sept. 2011, pp. 1–4.

[24] R. Gold, "Optimal binary sequences for spread spectrum multiplexing (corresp.)," Information Theory, IEEE Transactions on, vol. 13, no. 4, Oct. 1967, pp. 619–621.

[25] J. G. Proakis, Digital Communications, 4th ed. McGraw-Hill, 2000.

[26] R. Frank, S. Zadoff, and R. Heimiller, "Phase shift pulse codes with good periodic correlation properties (corresp.)," Information Theory, IRE Transactions on, vol. 8, no. 6, Oct. 1962, pp. 381 –382.

[27] P. Fan and M. Darnell, Sequence Design for Communications Applications. John Wiley & Sons Ltd., 1996.

[28] B. O'Hara and A. Petrick, The IEEE 802.11 Handbook: A Designer's Companion. Standards Information Network IEEE Press, 1999.

A VAD/VOX Algorithm for Amateur Radio Applications

William B. D. Forfang*, Eduardo Gonzalez†, Stan McClellan‡, and Vishu Viswanathan‡

* Freescale Semiconductor Inc., Austin TX, USA will.forfang@freescale.com
† FlexRadio Systems, Austin TX, USA ed.gonzalez@flexradio.com
‡ Ingram School of Engineering, Texas State University, San Marcos TX, USA
stan.mcclellan@txstate.edu, vishu.viswanathan@txstate.edu

***Abstract*—In amateur radio applications, voice activity detection (VAD) algorithms enable hands-free, voice-operated transmissions (VOX). In this paper, we first review a recent hybrid VAD algorithm, which was developed by combining features from two legacy speech detection algorithms long used in amateur radio applications. We then propose a novel VAD algorithm whose operating principles are not restricted to those of legacy approaches. The new method employs two key features. The first feature, called sub-band variance ratio, is the ratio of energies calculated over a low-frequency region and over the rest of the spectrum of the input audio signal. The second feature, called temporal formant density, is a running N-frame sum of the number of low-bandwidth formants over a low-frequency region. Both features are shown to yield low values for non-speech segments and relatively high values for speech segments. A two-state decision logic that uses these two features is employed to make frame-by-frame VAD decisions, which are then used in the VOX function for amateur radio transmissions. The proposed new method is compared against the hybrid method using both a simple objective measure involving comparisons against manually derived true VAD data and a subjective pairwise comparison listening test, over audio signal data from amateur radio transmissions at various signal-to-noise ratios. The results from these comparison tests show that the new method provides a better overall performance than the hybrid method. In summary, a new VAD/VOX algorithm for amateur radio applications is proposed that offers performance benefits over existing methods.**

***Keywords*—voice activity detection; VAD; voice-activated switch; voice-activated transmission; VOX.**

I. Introduction

As technology has progressed, speech processing algorithms have found ubiquitous deployment within consumer electronics. In many of these algorithms, Voice Activity Detection (VAD) plays an important role in increasing overall performance and robustness to noise. Amateur radios, digital hearing aid devices, speech recognition software, etc. are common examples of speech processing applications that employ VAD [1]–[3]. Precise discrimination between speech and non-speech allows, for example, an algorithm to capture, characterize, and update an accurate noise profile for adaptive noise cancellation [4]. The integrity of silence compression in discontinuous transmission schemes also relies upon the accuracy of VAD algorithms. Speech coding, speaker recognition, and speech enhancement are all examples of VAD applications [4]–[10].

VAD schemes with basic energy level detection can provide satisfactory performance at high signal-to-noise ratios (SNRs), but often perform poorly in noisy conditions. More robust VAD methods have been developed, which consider statistical features beyond average energy such as long-term spectral divergence [3] or multiple-observation likelihood ratio tests [11]. In 2012, Gonzalez and McClellan [1] published a VAD scheme that performs well in noisy environments while maintaining low computational complexity. Their algorithm specifically targets voice-activated transmission (VOX) applications.

A VOX is an amateur radio application that allows hands-free switching between the operating modes of a transceiver. A radio transceiver with a push-to-talk operating scheme requires a physical 'transmit' button to be pressed and held for the duration of the transmission, whereas a VOX automatically activates 'transmit' mode upon detection of an operator's voice. It then disables 'transmit' mode after observing a sufficient interval of non-speech.

In designing the VAD algorithm, Gonzalez and McClellan emulated a well-known legacy hardware approach to VOX, and then rectified its deficiencies by incorporating ideas from a complementary digital approach. The resulting hybrid algorithm was then tested in the context of amateur radio transmissions and was found to exhibit a higher robustness to noise than its legacy constituents, without an increase in computational cost.

Motivated by the success of the hybrid algorithm, this research investigates the design of a new VOX/VAD algorithm whose operating principles are not restricted to those of legacy devices. Instead, features for speech detection in the new algorithm are extracted from linear prediction coefficients and spectral sub-band energy analysis. Both algorithms target amateur radio applications, and are therefore compared to one another using audio stimuli captured from amateur radio transmissions. Quantitative evaluations of their relative performance were performed across multiple SNRs with both objective and subjective methods. For the remainder of this discussion, the algorithm by Gonzalez and McClellan will be referred to as the "hybrid algorithm," and the new algorithm will be referred to simply as the "new algorithm." Performance comparisons show that the new algorithm has better overall

performance than the hybrid algorithm.

It is worth emphasizing that our goal in this work has been to develop a VAD algorithm specifically for use with VOX transmission for amateur radio application. In this application, a speech transmission decision is made only after a frame has been declared as speech, and the decision to stop speech transmission is made only after a sufficient number of frames have been declared as non-speech. Therefore, for effective VOX transmission, it is not necessary that VAD decisions are made accurately every single frame. Because of the inherent delay in switching from transmit to no-transmit decision mode, occasional frame VAD errors are not significant and do not impact the effectiveness of the VOX transmission system. For this reason, the new algorithm is not compared to industry standard VAD implementations.

In Section II, the theoretical backgrounds behind the hybrid algorithm and the new algorithm are discussed. The legacy VAD approaches used in designing the hybrid algorithm are explained. The operating principles behind the new algorithm are described mathematically and shown graphically. In Section III, two methods used to compare the algorithms' performances are explained. One method is comprised of an objective performance measurement, and the other method is a subjective listening test. The results of the comparisons are given and discussed. The new algorithm was found to achieve equal or better performance under the majority of tested conditions. In Section IV, areas of further research are presented.

II. Theory of Operation

In a VAD paradigm an input audio signal can be generalized to fall into one of three categories at any given instant: noise (Eq. (1)), noise and voiced speech (Eq. (2)), or noise and unvoiced speech (Eq. (3)).

$$y(n) = q(n), \tag{1}$$

$$y(n) = q(n) + v(n), \tag{2}$$

$$y(n) = q(n) + u(n), \tag{3}$$

where $q(n)$ represents noise, $v(n)$ is voiced speech, and $u(n)$ is unvoiced speech. Distinguishing audio signals best modeled by Eq. (1) from those characterized by Eq. (2) or Eq. (3) is the goal of VAD.

In a general sense, speech is a statistically non-stationary signal. That is, its statistical description changes with time. However, a finite number of speech samples observed over a sufficiently short time period will exhibit wide-sense stationary statistical behavior [12]. To exploit such behavior for speech processing, an observed frame of audio samples must be short enough for meaningful statistical analysis but long enough to capture vocal features of interest. A detailed description of the two VAD/VOX algorithms is presented in the following subsections.

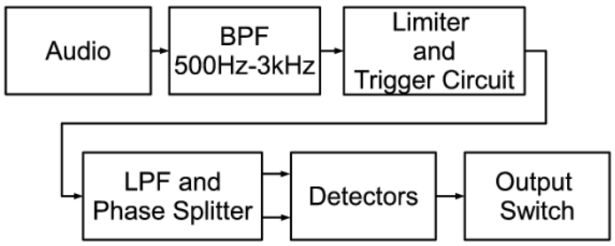


Figure 1. A high-level diagram of the MICOM algorithm, reprinted with permission from [1].

A. Hybrid Algorithm

The hybrid algorithm's operational theory proceeds from a combination of techniques found in two legacy VAD schemes. The first scheme is a hardware-driven approach developed by Motorola in the 1970s [13]. This circuit, which we refer to as the "MICOM" implementation, was popular with amateur radio enthusiasts since it provided a simple and easily implemented speech detection subsystem. The MICOM VOX continuously monitors a specified channel, suppressing non-speech noise in the idle channel while allowing detected speech signals to activate transmission. MICOM-like circuits exploit the syllabic rate of human speech (about 3 syllables per second) and include a detector for short-term frequency modulation, which is characteristic of voiced speech. The main components of the MICOM implementation include a high gain amplifier, a trigger circuit to produce constant width pulses, a 3.25 Hz low-pass filter, comparators, and timing circuitry to create hysteresis on the output "voicing" signal.

To implement MICOM features into the hybrid algorithm, a SPICE variant (Multisim [14]) was used to analyze and accurately decompose MICOM's functional components. These functional components were then duplicated using a simulation package (Simulink [15]) to model the subsystems via signal processing algorithms. Fig. 1 depicts a high-level block diagram of the MICOM algorithm.

Detailed descriptions of the individual blocks in Fig. 1 are given in [1], [13]. Briefly, the bandpass filter BPF extracts the voiceband part of the input audio signal; the Limiter and Trigger Circuit amplifies all non-zero samples to extreme saturation levels as a means of zero-crossing detection and generates a steady stream of uniform-width pulses, one per zero-crossing; the LPF and Phase Splitter block uses a 3.25 Hz low-pass filter to extract the syllabic rate envelope and a phase splitter to separate the signal into "top phase" and "bottom phase" voltages; the Detectors declare a detection event if either of these phase voltages is above a manually-set threshold; and finally, the Output Switch causes a single detection event to lead to a one-second holdover, using a timing capacitor, thereby avoiding "drop-outs" in the middle of active speech.

The second technique that influenced the design of the hy-

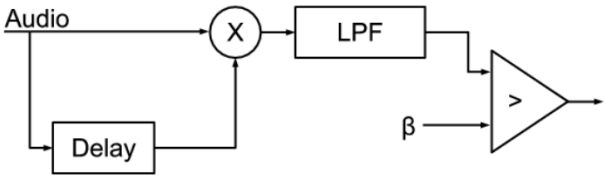


Figure 2. A high-level diagram of the Harris algorithm, reprinted with permission from [1].

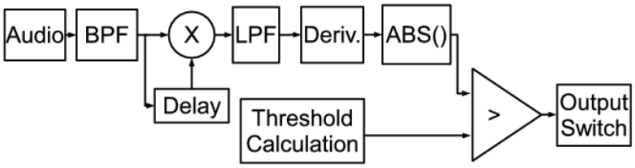


Figure 3. A high-level diagram of the hybrid algorithm, reprinted with permission from [1].

brid algorithm is a software-driven single-lag-autocorrelation process published by Harris Corp. in 1991 [16], which we refer to as the "Harris" algorithm. The Harris algorithm has several useful features for robust speech detection. However, in a complete implementation it may be lacking key features that are provided very effectively by aspects of the MICOM system. A block diagram of the Harris system is shown in Fig. 2.

The system incorporates a fixed delay and a multiply operation, which essentially computes a running autocorrelation at a single predetermined lag, according to Eq. (4).

$$ACF(l) = \sum_{n} X_n \bar{X}_{n-l} \tag{4}$$

The output from this delay and multiply operation is fed into a simple low-pass filter implemented as an accumulator. The resulting low-frequency component of the running autocorrelation is then compared to a threshold β to determine the presence of speech. The effect of the Harris approach is to detect strong, stable correlations around the predetermined lag value, which is related to pitch frequency.

Although the MICOM and Harris systems have advantages, they both also have shortcomings. The MICOM circuit is robust and simple to implement in an analog system, but some subtleties of modeling analog phenomena make it less stable and more difficult to implement directly in a discrete time system. The Harris algorithm performs well in detecting the onset of speech, but is inconsistent during active speech segments. The detector output has many false negatives (namely, non-speech) within active speech, and the resulting audio is choppy and incomprehensible. When the threshold is lowered to prevent these drop-outs, the same results occur during silence intervals since the noise creates a high enough output to repeatedly trigger a detect event.

The idea of detecting strong correlations around a predetermined lag value used in the Harris approach is valuable, but by itself it does not provide a reliable system. The hybrid implementation described here uses aspects of the MICOM system to address these problems.

A high-level diagram of the hybrid algorithm is shown in Fig. 3.

Each of the hybrid algorithm blocks is explained below:

- Bandpass Filter (300-700 Hz): The BPF provides the same function as the BPF in the MICOM circuit but the voiceband is decreased so that processing is done on more selective data.
- Delay and Multiply: Extracts short term periodicities in filtered audio. The chosen delay of 50 samples at a sampling frequency of 8000 Hz provides smooth operation and sufficient sensitivity.
- MICOM Low-pass Filter: Instead of using a simple accumulator, the 3.25 Hz low-pass filter from the MICOM circuit is used to extract syllabic rate information from the delay and multiply block. This filter also provides a much sharper cut-off, eliminating unwanted frequency components that interfere with the estimation in the Harris algorithm.
- Derivative and Absolute Value: The derivative converts the slowly changing output of the LPF into a more defined and faster changing waveform, which increases the tolerance and sensitivity of the threshold. Since the output of the LPF contains information about the changes in syllabic rate, like the phase splitter subsection of the MICOM circuit, both positive and negative deviations are important. The absolute value allows a single threshold to consider both deviations.
- Threshold Calculation: Removes the need for manual setting of the threshold value. To accomplish this, whenever speech is not detected, the energy of the noise is continuously calculated and the baseline threshold is established according to this changing energy level. This allows detection in varying noise floors.
- Modified MICOM Output Switch: Forces a holdover period following a detection event. Instead of using a 1.0s holdover (as in the MICOM circuit) the hybrid algorithm uses a 0.25s holdover, which minimizes drop-outs during active speech without overly extending the detection period.

The performance of the hybrid algorithm was compared against the MICOM system and the Harris algorithm in [1]. The hybrid algorithm was shown to exceed the performances of the Harris algorithm as well as the MICOM approach in both stability and robustness to noise. Fig. 4 shows performances of the Harris, MICOM, and hybrid VOX implementations in a low-noise condition. Although Fig. 4 seems to display a fairly "clean" or lab quality original signal, the signal is actually a speech utterance captured from an amateur radio

transmission, and contains some objectionable, non-speech noise. In the figure, several error conditions are labeled. Note the highly erratic performance of the Harris approach in voiced segments ("A"), but also the ability of the Harris approach to reliably (albeit aggressively) determine non-speech segments ("B"). Also note the inaccurate speech/non-speech decisions of the MICOM approach ("C"). The hybrid approach typically produces accurate voicing indicators with acceptable overhang, and without aggressive penetration into non-speech segments. There are a few exceptions (e.g., a missed onset at "D"). However, this level of performance is quite acceptable for real-time implementation, which avoids clipping, slow-attack, and other behaviors that are objectionable to amateur radio operators. For a more detailed discussion of the comparison among the hybrid algorithm, the Harris algorithm, and the MICOM implementation, the reader is referred to [1]. The performance comparison results presented in [1] show that the hybrid algorithm performs significantly better than the other two algorithms and provides robust and stable speech detection performance in realistic operational conditions.

B. New Algorithm

The operating principle behind the new algorithm relies on the predictability of formant locations during voiced speech. A spectral estimate of audio samples can be analyzed to exploit this property. The frame length must be kept small enough to minimize computational latency, but large enough to resolve the temporal events of interest. A 30 millisecond buffer with a 50% overlap was found to fall appropriately within this envelope of efficiency and speed, yielding a new data set every 15 milliseconds. Two features, or speech indicators, are implemented in the new algorithm. Section II-B1 explains the first feature, which is a ratio of energy levels at predetermined spectral locations. A summary of the second feature, which tracks formant activity, is given in Section II-B2. Finally, the logic that determines speech presence given the available feature data is discussed in Section II-B3.

1) Feature 1, Sub-band Variance Ratio: To analyze the spectral energy, the 30 millisecond frame is multiplied by a Hamming window and a 512 point fast Fourier transform (FFT) is applied. The absolute value of the FFT can be seen in Fig. 5, in which the shaded range (with width W, and distance from the origin d) represents the test area for active speech information.

To test for increased spectral activity in this range, the ratio of the variance within the shaded region to the variance of the remaining spectrum is compared to a threshold value. The derivation of this ratio, Γ_{var} can be seen in Eq. (5) - Eq. (7).

$$\sigma_1^2 = \frac{1}{W-1}\sum_{k=1}^{W}(s_1(k)-\mu_{s1})^2, \tag{5}$$

$$\sigma_2^2 = \frac{1}{N-W-1}\sum_{k=1}^{N-W}(s_2(k)-\mu_{s2})^2, \tag{6}$$

$$\Gamma_{\text{var}} = \frac{\sigma_1^2}{\sigma_2^2}, \tag{7}$$

where $s_1(k)$ is the shaded portion of the spectral estimate; $s_2(k)$ is the spectrum with $s_1(k)$ removed and the remainder concatenated together. μ_{s1} and μ_{s2} are the means of $s_1(k)$ and $s_2(k)$, respectively.

To maximize the selectivity of the Γ_{var} metric, the statistical distance between distributions of Γ_{var} values during speech and during non-speech should be maximized. A script was created, which adjusts the W and d values incrementally and creates distributions of Γ_{var} values for speech and non-speech audio after each adjustment. It then calculates the statistical distances between the distributions, using the well-known Bhattacharyya distance measure, and stores them into a two-dimensional array. By doing this many times, for many combinations of W and d values, a three-dimensional space can be analyzed in which peaks represent W and d combinations that yield high statistical separability of classes. Fig. 6 shows the described three-dimensional space for a range of W and d values obtained from processing test data from amateur radio transmissions with varying SNRs.

Through this analysis, the determined values for the W and d parameters were 781 Hz and 219 Hz, respectively. Fig. 7 displays the Γ_{var} function (bottom graph) and the utterance from which it was extracted (top graph). The utterance was sourced from an amateur radio broadcast and is mostly clean speech. Clearly, Γ_{var} is relatively small for non-speech segments and relatively large during speech segments.

2) Feature 2, Temporal Formant Density: The second feature of interest is also motivated by spectral distribution analysis, but is realized in a different manner. As described by Eq. (8), the coefficients (a_i) of a 10th order forward linear predictor (LPC) are calculated for each windowed frame of time-domain data $y(n)$. The general idea behind Eq. (8) is that a given speech sample at time n can be approximated as the linear combination of the past 10 samples (10th order) weighted by their respective coefficients, a_i.

$$\overline{y} = \sum_{i=1}^{10} a_i y(n-i) \tag{8}$$

The predictor coefficients (a_i) for each frame are computed via autocorrelation and Levinson-Durbin recursion [12]. Once these coefficients are found, they are treated as the coefficients of a polynomial whose roots are then solved for. These complex conjugate pairs of roots are expressed as $r_0 e^{\pm i\phi_0}$, where i is the square root of -1, r_0 is the root magnitude, and ϕ_0 is the root angle. From these roots, Eq. (9) and Eq. (10) are used to estimate formant center frequencies (F) and bandwidths (BW), respectively [12]; f_s denotes the sampling frequency in Eq. (9) and Eq. (10). These estimates are valid for high Q formants [12], which is the case for our formant analysis.

$$F = \frac{f_s}{2\pi}\phi_0 \tag{9}$$

$$BW = \frac{-f_s}{\pi}\ln(r_0) \tag{10}$$

For a more intuitive understanding of the formant estimation process, Fig. 8 graphically summarizes the steps from raw

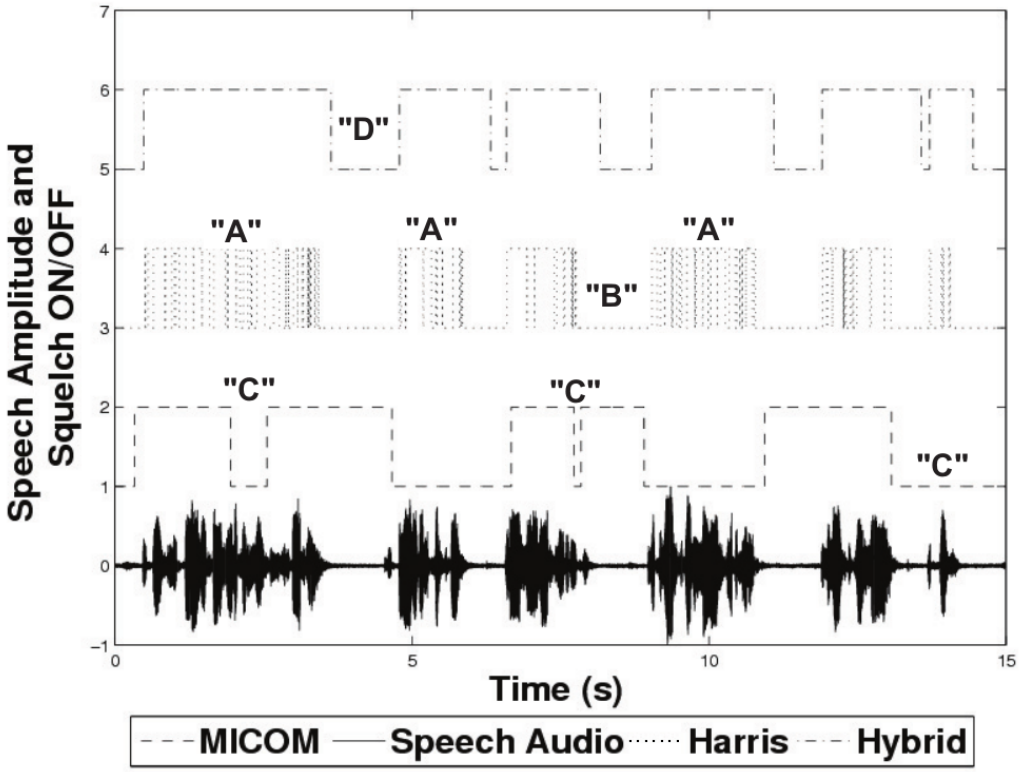


Figure 4. Performance of all three voice detection algorithms in a low noise, natural environment. The utterance was captured from an amateur radio transmission, and contains some non-speech noise. Annotations "A" through "D" indicate detection errors in each algorithm.

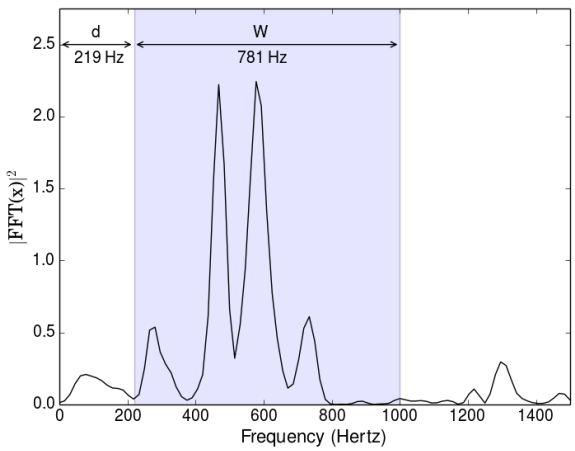


Figure 5. Spectral estimate of a 30 millisecond frame of audio data. The variance of the signal within the gray area with width W, and distance from origin d, is divided by the variance of the remaining spectrum. This quotient, Γ_{var}, signifies the level of spectral activity within the shaded area relative to the remaining spectrum.

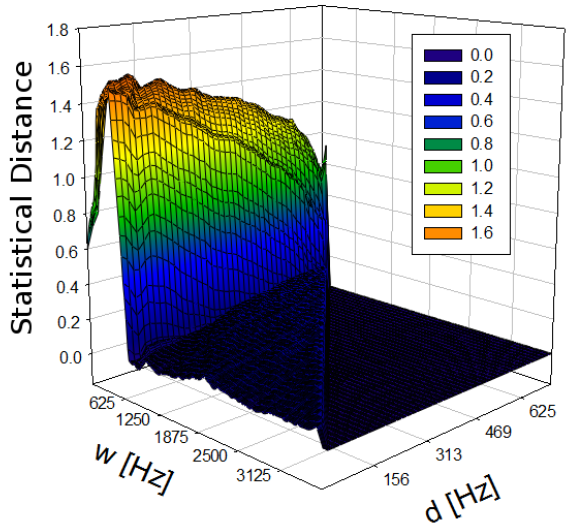


Figure 6. This plot shows the statistical distance between distributions of Γ_{var} values while processing speech audio versus non-speech audio for a large combination of W and d values. The highest peak in this space, which corresponds to a d value of 219 Hz and a W value of 781 Hz, represents a configuration that yields the highest statistical separability between speech/non-speech classes for the given test data.

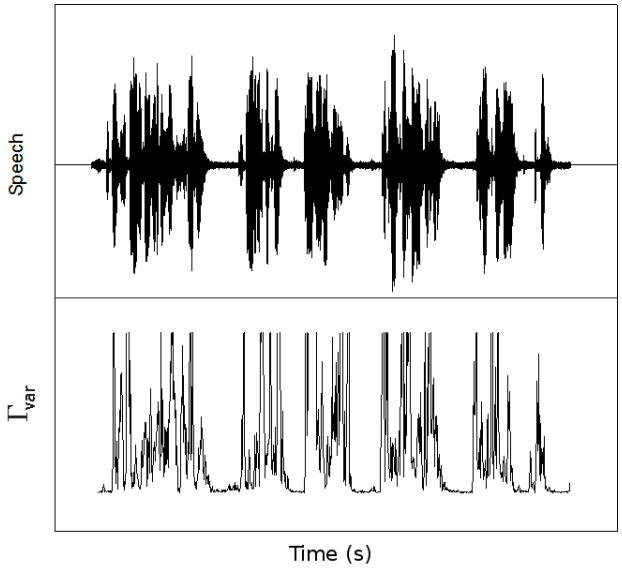


Figure 7. A recorded utterance sourced from an amateur radio transmission (upper) and its resulting Γ_{var} function (lower). The Γ_{var} function is truncated above an arbitrary value along the y-axis for clearer representation. The graph shows that the Γ_{var} feature yields relatively high values during speech segments and relatively low values during segments of non-speech.

audio data to formant estimation via Eq. (8) and Eq. (9). First, the raw audio data for a 30 millisecond frame (Fig. 8(A)) is multiplied by a Hamming window. From the weighted data, the autocorrelation and Levinson-Durbin recursion methods are used to calculate an LPC model for the given frame, as defined by Eq. (8). The LPC coefficients are treated as polynomial coefficients whose roots are then extracted and are plotted in polar form in Fig. 8(B). Eq. (9) is then used to translate root angles to formant center frequencies. The resulting formant estimations are plotted over the LPC spectrum in Fig. 8(C), along with the raw audio spectrum overlaid for comparison.

By comparing each root in Fig. 8(B) (in order of increasing angle) to their respective formant estimates in Fig. 8(C) (in order of increasing frequency) it is clear that roots farther from the origin correspond to estimated formants with smaller bandwidths.

By analyzing the formant locations and their bandwidths, an estimate of the spectral energy distribution can be made for each frame of time-domain data. In Fig. 9 and Fig. 10, two separate audio frames are displayed above scatter-plots of their respective formant estimates.

In these scatter plots, the y-axis represents formant bandwidth and the x-axis represents formant frequency. The shaded rectangles represent a decision space, within which formants will tend to lie when voiced speech is present. Adopting a similar approach to that used for determining the W and d values in the Γ_{var} analysis of Section II-B1, we chose the decision space dimensions to include formants with center frequencies between 100 and 1000 Hz, and bandwidths under 100 Hz. In an LPC spectrum, these formants would manifest as relatively lower frequency, pronounced spectral peaks. Notice that two formants fall within the shaded region during the particular speech frame displayed in Fig. 10 while no formants do so in Fig. 9. The speech frame in Fig. 10 is extracted from a voiced segment of the word 'principle,' recorded over an amateur radio transmission; the non-speech frame in Fig. 9 is background noise extracted from the same transmission.

By monitoring the number of formants within the decision space and computing a running sum of this number over the last N frames (labeled ρ_f below), a *temporal density* of formants is computed. That is, the number of formants that have landed within the decision space over the last N frames is computed to detect the presence of speech. By choosing N = 10, some hysteresis is introduced into ρ_f. In Fig. 11, the ρ_f feature is plotted above the utterance from which it was extracted. Clearly, non-speech segments correspond to values of ρ_f closer to zero while speech segments correspond to ρ_f values closer to 10.

3) Two-State Logic: To combine the above-mentioned features (Γ_{var} and ρ_f) into a speech detection scheme, decision thresholds were chosen statistically. Histograms of feature values during both speech and non-speech were calculated over a variety of utterances and noise levels. Given some overlap in these distributions, two thresholds for Γ_{var} were chosen to indicate a high or low likelihood of speech presence, labeled T_Γ^1 and T_Γ^0, respectively. This creates three possible sample spaces for Γ_{var} values at a given instance: likely speech ($\Gamma_{\text{var}} >= T_\Gamma^1$), likely non-speech ($\Gamma_{\text{var}} <= T_\Gamma^0$), or inconclusive ($T_\Gamma^0 < \Gamma_{\text{var}} < T_\Gamma^1$). Given the higher statistical distances between distributions of ρ_f values during speech and non-speech, only one threshold T_ρ is used.

Finite-state logic is then employed to allow state changes only when both features indicate that the current frame of audio differs from the previous frame of audio. With three possible Γ_{var} interpretations and two possible ρ_f interpretations, six combinations of feature indications can be realized, but only two merit a state change. The state logic is summarized in the diagram of Fig. 12, where **a** represents detected speech from both features, **b** represents detected non-speech from both features. For hands-free VOX switching applications, the transceiver would begin transmissions when the algorithm moves to the "speech" state. It would then end transmissions when the algorithm moves back to the "non-speech" state, and remains there for a sufficient duration.

III. Performance Comparison

A human subject's perceptual evaluation of audio stimuli is the outcome of a complex physiological process. In this process, the quality aspects of the audio are considered along with the subject's expectations, mood, context, etc. When assessing the quality of audio stimuli, the criteria upon which opinions are formed are difficult to characterize, even for the assessor. Forming a mathematically predictive model to forecast such psychoacoustic assessments is therefore difficult. In this research, a simple objective metric was designed to estimate the relative performance of the algorithms. Although

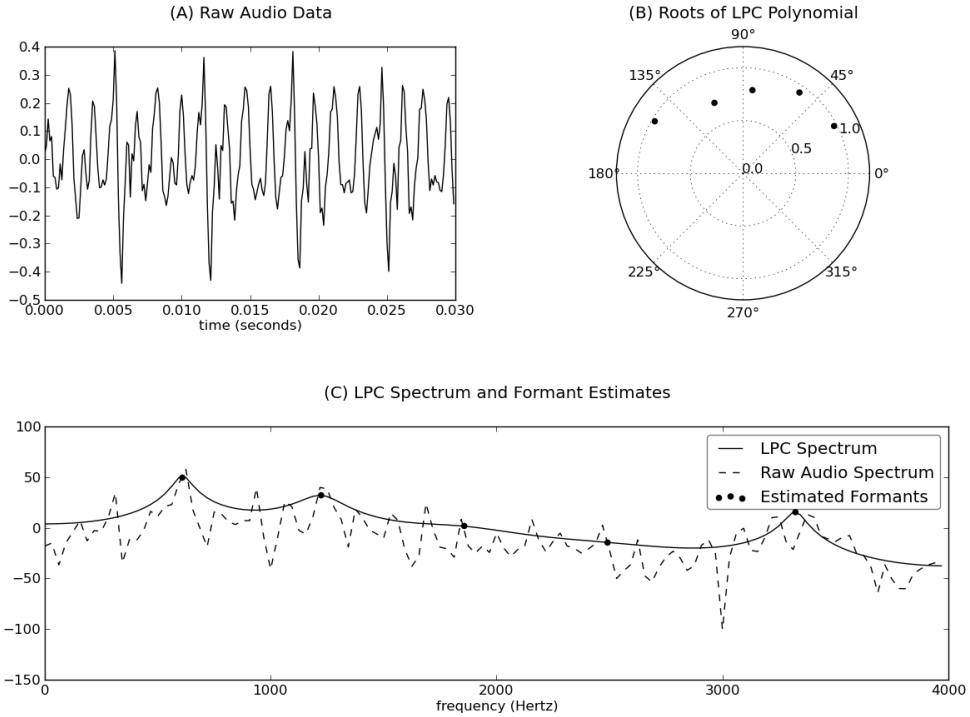


Figure 8. A graphical summary of the mathematical transition from raw audio to LPC formant estimates. A 30 millisecond sample of audio data is plotted in (A). After computing the LPC coefficients from the weighted data, the LPC polynomial roots are extracted and plotted in polar form (B). The formant center frequencies are estimated from the roots with Eq. (9) and plotted over the LPC spectrum in (C). The raw audio spectrum is also displayed in (C).

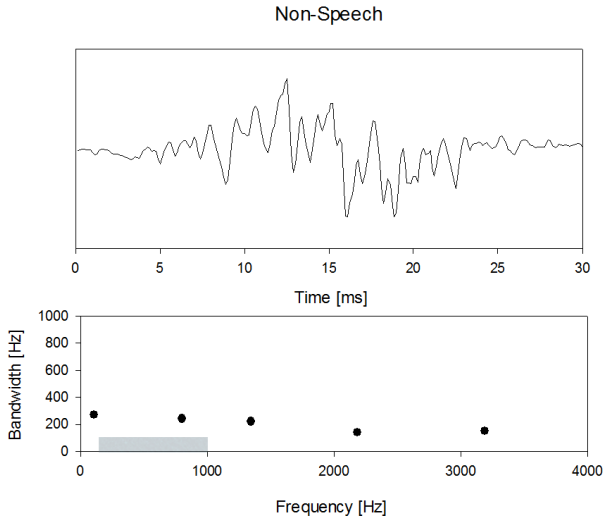


Figure 9. A 30 millisecond frame of non-speech audio (above) and its corresponding formant estimates (below). The y-axis of the lower plot is the estimated formant bandwidth, as defined by Eq. (10), and the x-axis is frequency. The shaded area signifies the decision space for speech-data. Notice that no formants lie within the shaded area.

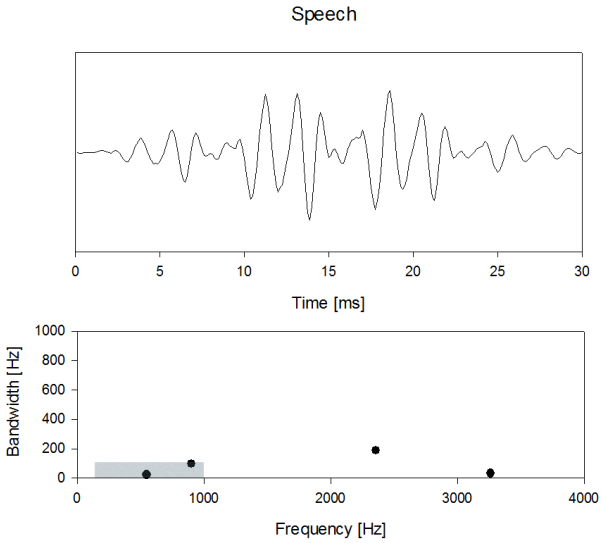


Figure 10. A 30 millisecond frame of voiced speech data (above) and its corresponding formant estimates (below). The y-axis of the lower plot is the estimated formant bandwidth, as defined by Eq. (10), and the x-axis is frequency. Notice two formants fall within the perimeter of the decision space, indicating an increased likelihood of speech presence.

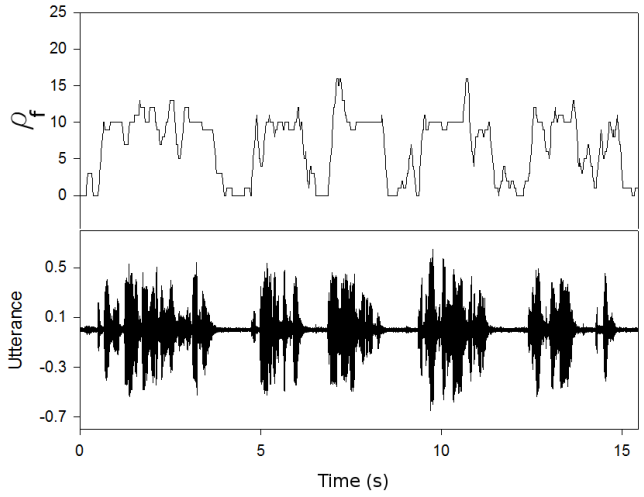


Figure 11. A running temporal formant density (ρ_f) (upper graph) extracted from an amateur radio speech recording (lower graph). It is clear from the graph that higher (ρ_f) values correspond to speech segments and lower (ρ_f) values correspond to non-speech segments.

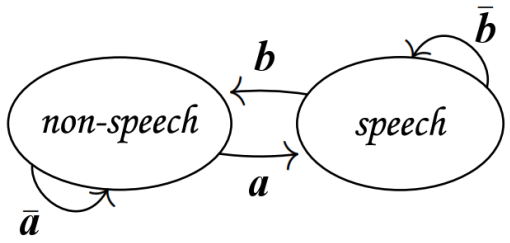


Figure 12. Two-state logic used for making VAD decisions. The audio is deemed either speech or non-speech by the algorithm depending on its previous state and current inputs. Path $\mathbf{a}$ is taken if both features indicate speech presence, path $\mathbf{b}$ is taken if both features indicate non-speech presence. Path symbols $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ refer the logical complements of $\mathbf{a}$ and $\mathbf{b}$.

this metric was not intended to predict subjective assessments, it provided valuable feedback during algorithm design at little cost.

In addition, a subjective evaluation using a paired comparison listening test was administered. Untrained subjects were briefly informed on the purpose and scope of the VAD algorithms, and then presented with a series of audio samples to evaluate. The samples were organized into pairs, and the subjects were asked to choose the preferred audio from each pair. The following sections provide greater detail into these test methods and their results.

A. Objective Comparison

To objectively evaluate a given VAD performance, its resulting activity mask was first compared to a manually derived true marker mask as seen in Fig. 13. In this figure, the audio is plotted under the resulting VAD activity mask. The absolute value of the difference between the VAD mask and the true marker mask is shaded in gray. Taking the integral of the

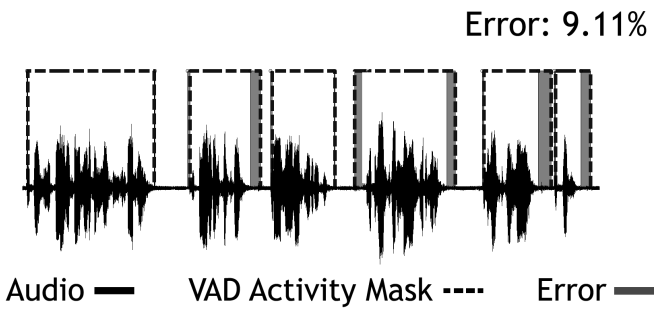


Figure 13. Audio with VAD activity mask overlaid. Gray shaded areas represent errors relative to true markers. The integral of the shaded area divided by the audio sample length gives the error percentage.

shaded areas and normalizing by the audio sample length gives a percentage error relative to the true markers, which was used to gauge the VAD accuracy.

To employ this test metric, the algorithms first processed relatively clean speech captured from radio transmissions and were scored accordingly. Noise was then added systematically to the speech files and the algorithms were graded objectively over a variety of SNRs. Fig. 14 plots the quantitative results from two radio speech transmissions processed at three different SNRs (30 dB, 15 dB, and 0 dB). Admittedly, the database used for evaluation was relatively small but was found to be adequate for our purposes. In Section IV, we suggest test methods that use a larger speech database.

In Fig. 14(a), both algorithms perform similarly. In Fig. 14(b) the hybrid algorithm fails in 'open' mode during the 0 dB test while the new algorithm remains functional. That is, the hybrid algorithm reports detected speech during the entirety of the test. A closer analysis of the algorithms' behavior during this particular test can be seen in Fig. 15.

The objective metric in Fig. 14 proves suitable for distinguishing between large performance disparities (such as those seen in Fig. 15). However, small differences in VAD performances may go unnoticed in this basic test paradigm. To the human auditory system, these small differences may exhibit high perceptual significance depending upon their context within an utterance. Therefore, the objective metric may be limited in its ability to accurately predict perceptual evaluations. To investigate this possibility, a paired comparison listening test was conducted.

B. Subjective Comparison

Although the objective metric described in Section III-A offers a relative evaluation, its comparison proved inconclusive for all but extreme disparities in algorithm performances. To pursue the possibility of subtle differences in algorithm performances (such as those in Fig. 14(a)) containing overlooked perceptual significance, a paired comparison listening test was designed and administered to 10 untrained participants.

Two utterances captured from radio transmissions were used as test audio for the subjective evaluations. Gaussian

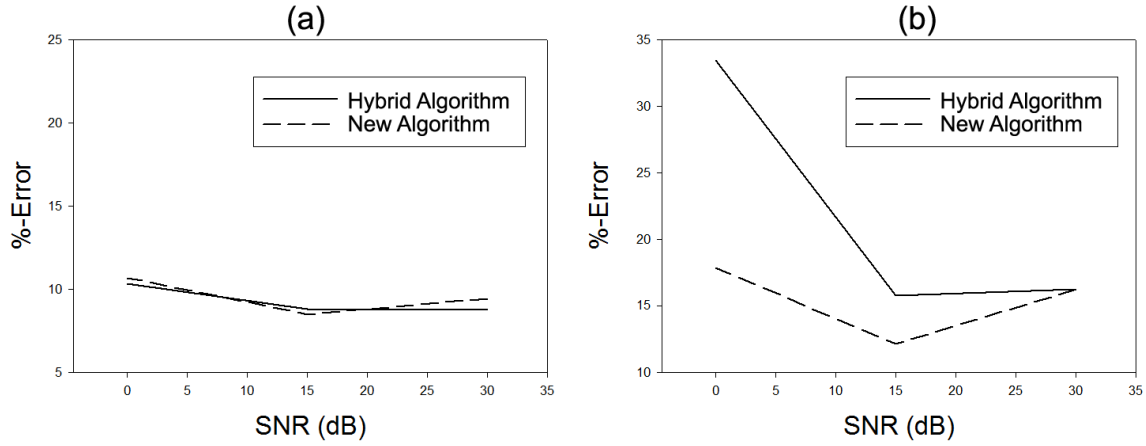


Figure 14. A comparison of algorithm performance swept over SNR. The raw test audio for the comparison in (a) was captured from a radio transmission containing clean speech. In (b) the raw test audio was captured from a radio transmission containing mostly clean speech with occasional hand clapping sounds. Noise was added systematically to both files to create the SNR sweep.

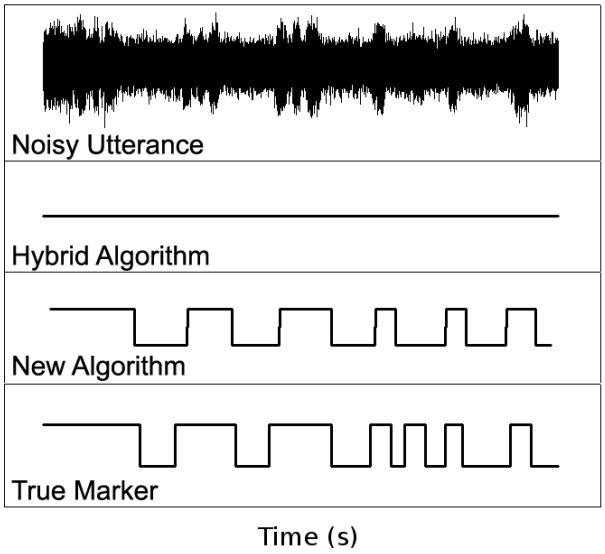


Figure 15. The noisy utterance used during the 0 dB objective measurement from Fig. 14(b) along with the algorithms' resulting VAD activity masks and the utterance's true marker mask. Here, the hybrid algorithm fails 'open,' falsely indicating speech throughout the utterance. The new algorithm maintains overall functionality.

white noise was added to produce three different SNRs per utterance. Signal power measurements for SNR calculations were taken only over the speech sections of the utterances, as determined by manually derived true markers. The test files were processed by the two VAD algorithms, which were configured to zero (mute) the sections they deemed non-speech and preserve the sections deemed speech. Eight pairs of the resulting processed audio files were played to listening test participants in the order presented in Table I.

Rows 6 and 8 in Table I contain pairs of audio samples with significantly different SNRs, which we included for control purposes. Test results from these pairs were used as a means

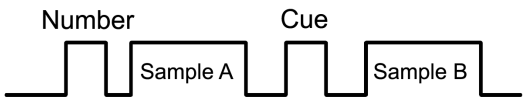


Figure 16. The format for each pair of audio stimuli administered in the paired comparison listening test. The pair number is first announced, followed by the two samples, which are separated by a cue tone for clarity.

to identify potentially unreliable listeners and to exclude their scores from statistical analysis.

The untrained test participants were provided a brief description of the definition and purpose of VAD algorithms before taking the test. An ordinal A/B multiple choice answersheet comprised the response format. Each of the 8 pairs of audio stimuli presented in Table I were played to the test participants in the fashion depicted by Fig. 16. The 'number' in Fig. 16 indicates which pair is about to be played; the 'cue' tone serves to separate the samples in each pair.

The mean of the responses was calculated at each tested SNR, and is plotted in Fig. 17. This data indicates that the new algorithm was preferred over the hybrid algorithm at SNRs above 4 dB. The hybrid algorithm was preferred slightly at 0 dB SNR. Thus, the new algorithm provides a better overall performance than the hybrid algorithm.

IV. Future Research Directions

We suggest below a few areas for further research. The first area is to investigate ways of reducing the computational complexity of the proposed new VAD/VOX algorithm. Recall that the temporal formant density feature used in the new algorithm requires solving for the roots of the LPC polynomial and calculating the number of low-bandwidth formants in the designated low-frequency region (see Section II-B2). Solving

TABLE I
LISTENING TEST DETAILS.

	SAMPLE A			SAMPLE B		
	ALGORITHM	SNR	UTTERANCE	ALGORITHM	SNR	UTTERANCE
1	hybrid	30 dB	1	new	30 dB	1
2	new	0 dB	2	hybrid	0 dB	2
3	hybrid	15 dB	1	new	15 dB	2
4	new	15 dB	2	hybrid	15 dB	2
5	hybrid	30 dB	2	new	30 dB	2
6	hybrid	15 dB	1	new	0 dB	1
7	new	0 dB	1	hybrid	0 dB	1
8	new	30 dB	2	hybrid	15 dB	2

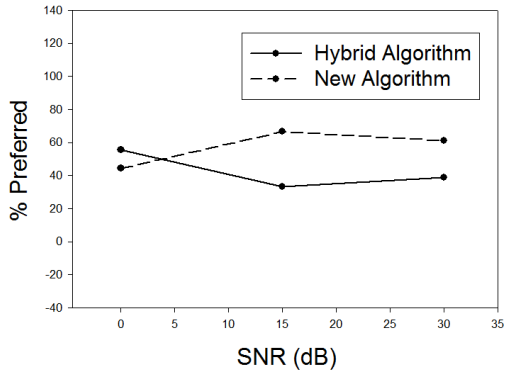


Figure 17. The averaged preference of test participants at different SNRs. The two preference functions sum to 100%. The listening test results reported a preference for the new algorithm when processing stimuli with SNRs over 4 dB. The hybrid algorithm was preferred at 0 dB.

for the polynomial roots is a computationally intensive task. Reference [17] proposes a way of directly calculating the number of formants in a given frequency range without first solving for the roots of the polynomial and shows significant computational savings. This approach may be modified to directly calculate the number of formants with bandwidths smaller than a specified value. The same paper also contains a procedure for calculating the bandwidths of formants in a specified frequency region.

We can also reduce the computational complexity of the sub-band variance ratio feature used in the new algorithm. Since the LPC analysis (used in the temporal formant density feature) provides the zeroth autocorrelation coefficient R(0), by Parseval's theorem, R(0) equals the energy under the entire power spectrum. By computing the power spectrum only over the shaded region in Fig. 5 and by using R(0), we can compute a ratio similar to the one given by Eq. (7). The approximation involves not using the means in Eq. (5) and Eq. (6), yielding an energy ratio rather than a variance ratio. The modest computational savings here arise from not having to compute the spectral values in the frequency range outside the shaded region.

As described in Section II, both sub-band variance ratio and temporal formant density features focus on the low-frequency region and as such are effective in identifying voiced frames. By the same token, the new method may likely misclassify unvoiced speech frames as non-speech. While these errors should not affect the effectiveness of our target VOX transmission application, it may be desirable to minimize misclassification of unvoiced speech as non-speech. To this end, a third feature may be developed by focusing on energy ratio and/or formant density in a designated high-frequency region; VAD decisions are then made using all three features.

Although the focus of this work was on VOX transmission for amateur radio application, it may be of interest to evaluate the effectiveness of the proposed new algorithm by comparing against industry standard VAD algorithms like the ones used in the ITU standard G.729 or 3GPP Adaptive Multi-Rate (AMR) standard [4]. The challenge here is to have access to a speech database with known ground-truth VAD decision data. We have identified one such database that is included in a 1998 TIA standard called TIA/EIA/IS 727 [18]. This database includes 5 male and 5 female clean speech sentences, noise files in four noise environments (which can be used to generate noisy speech files at different signal-to-noise ratios), and ground-truth VAD decision data for the ten clean speech files. Both objective and subjective tests may be performed in comparing the new algorithm against widely used industry standard VAD algorithms.

V. CONCLUSION

Motivated by the success of a hybrid VAD algorithm described in [1], a new algorithm targeting amateur radio applications was developed. Unlike the hybrid algorithm, whose design combines ideas from two well-known methods, the new algorithm was designed without restricting its operating principles to those of legacy approaches. The performance of the new algorithm was compared to the hybrid algorithm, both objectively and subjectively, in the context of amateur radio transmission data. The objective evaluations, which were computed by comparing the algorithm behavior to true VAD markers as described in Section III-A, indicate that the new algorithm achieves equal or higher performance than the hybrid algorithm under the tested noise conditions. The subjective evaluations, which were performed through the listening test described in Section III-B, show that the new algorithm was preferred over the hybrid algorithm by the majority of listeners, particularly at higher SNRs. In Section IV, future research ideas are suggested for (i) reducing the

computational complexity of the new algorithm, (ii) preventing the new algorithm from misclassifying unvoiced speech as non-speech, and (iii) enlarging the performance testing by using an industry standard speech database and by comparing the new algorithm against well-known industry standard VAD algorithms.

REFERENCES

[1] E. Gonzalez and S. McClellan, "A hybrid VOX system using emulated hardware behaviors," in *International Conference on Digital Telecommunications (ICDT)*, April 2012, pp. 105 – 110.

[2] T. V. den Bogaert, J. Wouters, S. Doclo, and M. Moonen, "Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 4, April 2007, pp. IV–565 – IV–568.

[3] J. Ramírez, J. Segura, C. Benítez, Á. de la Torre, and A. Rubio, "Efficient voice activity detection algorithms using long-term speech information," *Speech Communication*, vol. 42, no. 3-4, pp. 271–287, April 2004.

[4] J. Ramírez, M. Górriz, and J. Segura, "Voice activity detection. Fundamentals and speech recognition system robustness," in *Robust Speech Recognition and Understanding*, M. Grimm and K. Kroschel, Eds. Vienna, Austria: I-Tech, June 2007, pp. 1–22.

[5] S. G. Tanyer and H. Ozer, "Voice activity detection in nonstationary noise," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 478–482, July 2000.

[6] J. Séris, C. Gargour, and F. Laville, "VAD INNES: A voice activity detector for noisy industrial environments," in *50th Midwest Symposium on Circuits and Systems*, August 2007, pp. 377–380.

[7] B. V. Ilarsha, "A noise robust speech activity detection algorithm," in *International Symposium on Intelligent Multimedia, Video and Speech Processing*, October 2004, pp. 322–325.

[8] G. Evangelopoulos and P. Maragos, "Multiband modulation energy tracking for noisy speech detection," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 2024–2038, November 2006.

[9] L. Ye, W. Tong, C. Huijuan, and T. Kun, "Voice activity detection in nonstationary noise," in *IMACS Multiconference on Computer Engineering in Systems Applications*, October 2006, pp. 1573–1575.

[10] D. Wu, M. Tanaka, R. Chen, L. Olorenshaw, M. Amador, and X. Menendez-Pidal, "A robust speech detection algorithm for speech activated hands-free applications," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, March 1999, pp. 2407–2410.

[11] J. M. Górriz, J. Ramírez, J. Segura, and C. G. Puntonet, "An improved MO-LRT VAD based on bispectra Gaussian model," *Electronics Letters*, vol. 41, no. 15, pp. 877 – 879, July 2005.

[12] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Prentice Hall, 1978.

[13] *Service Manual for Motorola Micom HF SSB Transceiver*, Motorola, Inc., 1975, part No. 68-81025E95A, The "Constant SINAD" Squelch was used in the Motorola Micom HF SSB Transceiver. The MICOM squelch board part number is TRN6175.

[14] *Multisim 11.0*, National Instruments, July 2011.

[15] *Simulink 2011a*, MathWorks Inc., January 2011.

[16] M. Webster, G. Sinclair, and T. Wright, "An efficient, digitally-based, single-lag autocorrelation-derived, voice-operated transmit (VOX) algorithm," *Military Communications Conference, 1991. MILCOM'*, November 1991.

[17] R. C. Snell and F. Milinazzo, "Formant location from LPC analysis data," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 2, pp. 129 – 134, April 1993.

[18] *TDMA Cellular/PCS - Radio Interface - Minimum Performance Standards for Discontinuous Transmission Operation of Mobile Stations*, Telecommunications Industry Association, June 1998.

www.iariajournals.org

International Journal On Advances in Intelligent Systems

ICAS, ACHI, ICCGI, UBICOMM, ADVCOMP, CENTRIC, GEOProcessing, SEMAPRO, BIOSYSCOM, BIOINFO, BIOTECHNO, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS, ENERGY, COLLA, IMMM, INTELLI, SMART, DATA ANALYTICS

issn: 1942-2679

International Journal On Advances in Internet Technology

ICDS, ICIW, CTRQ, UBICOMM, ICSNC, AFIN, INTERNET, AP2PS, EMERGING, MOBILITY, WEB

issn: 1942-2652

International Journal On Advances in Life Sciences

eTELEMED, eKNOW, eL&mL, BIODIV, BIOENVIRONMENT, BIOGREEN, BIOSYSCOM, BIOINFO, BIOTECHNO, SOTICS, GLOBAL HEALTH

issn: 1942-2660

International Journal On Advances in Networks and Services

ICN, ICNS, ICIW, ICWMC, SENSORCOMM, MESH, CENTRIC, MMEDIA, SERVICE COMPUTATION, VEHICULAR, INNOV

issn: 1942-2644

International Journal On Advances in Security

ICQNM, SECURWARE, MESH, DEPEND, INTERNET, CYBERLAWS

issn: 1942-2636

International Journal On Advances in Software

ICSEA, ICCGI, ADVCOMP, GEOProcessing, DBKDA, INTENSIVE, VALID, SIMUL, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS, IMMM, MOBILITY, VEHICULAR, DATA ANALYTICS

issn: 1942-2628

International Journal On Advances in Systems and Measurements

ICQNM, ICONS, ICIMP, SENSORCOMM, CENICS, VALID, SIMUL, INFOCOMP

issn: 1942-261x

International Journal On Advances in Telecommunications

AICT, ICDT, ICWMC, ICSNC, CTRQ, SPACOMM, MMEDIA, COCORA, PESARO, INNOV

issn: 1942-2601